The Architecture of Metacognition: Insights into its Structural Organization and Neurobiological Substrates

– Kumulative Arbeit –

Inaugural-Dissertation

zur Erlangung der Doktorwürde

der

Philosophischen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität zu Bonn

vorgelegt von

Mirko Lehmann

aus

Bonn

Bonn, 2023

Gedruckt mit der Genehmigung der Philosophischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission:

Prof. Dr. Henning Gibbons	Vorsitzender
Prof. Dr. Ulrich Ettinger	Betreuer und Gutachter
Prof. Dr. Steffen Moritz	Gutachter
PD Dr. Ina Grau	Weiteres prüfungsberechtigtes Mitglied

Tag der mündlichen Prüfung: 24. Januar 2023

Contents

Acknowledgements

First and foremost, I would like to thank Ulrich Ettinger. It is by no means idle talk to say that I could not have wished for a better supervisor. Your knowledgeable and patient, amiable and trustful guidance throughout the long and sometimes arduous journey of my PhD has left me with nothing but appreciation. Our discussions have sparked my fascination for psychological research, and I will always be grateful to you for placing your trust in me and providing me with the opportunity to work as freely and independently as I did on such exciting projects.

Likewise, I want to thank Steffen Moritz for agreeing to be my second reviewer and fitting my thesis in what must surely be a very busy schedule. Of course, this also applies to Ina Grau and Henning Gibbons, who I thank for being part of my examination board.

These long days at the MRI scanner could easily have been a pretty dull affair. Quite the contrary, however, they soon became something to look forward to, courtesy to my fantastic MetaKeta study team. I would like to thank Sven, Achilles, Claudia and Laura in particular, who made these days entertaining while keeping the assessments on track at all times.

I also thank all the students who spent hours and hours recruiting participants and collecting research data in the wake of their Bachelor or Master theses. I thoroughly enjoyed my time with you and fondly remember many stimulating conversations.

Furthermore, I am grateful to say that I got really lucky with my colleagues. I was able to work with a group of incredibly diligent, yet also super-kind and supportive people who definitely made my PhD time more worthwhile. This goes especially to Rebekka, Kaja, Eliana, Kristof, Maria, Yaira, Celina, Philine and Jule – it was my pleasure getting to know all of you, and I certainly hope that we stay in touch.

A heartfelt “Thank you” goes to my beloved family and friends. I knew that I could always count on your support, and that is something that I do not take for granted.

Finally, I would like to thank Stella, my star, my roots and my wings. I couldn’t have fathomed doing this without you. You mean the world to me, and these years will forever be an important part of our story. Thank you for always being there for me and being my companion in this adventure, and hopefully, every adventure to come.

ABSTRACT

Metacognition, the capacity to monitor one's own mental states and processes, is integral to successful human behavior. However, key questions regarding the architecture of metacognition remain hitherto unresolved. For example, it is unclear whether the accuracy of metacognitive evaluations converges on a single, unitary resource across various tasks or whether there are separable metacognitive subsystems for different demands. This also poses a challenge to the investigation of the neurobiological processes underlying metacognitive operations, as insights into the structure of metacognition in one functional domain are thus not readily applicable to other domains. Furthermore, only little is known about how metacognition is underpinned at the level of neurotransmitter systems.

The research reported in this dissertation aimed to yield insights towards resolving these questions by investigating the architecture of metacognition in three original studies. Specifically, a large-scale behavioral study was conducted to analyze the pattern of individual differences in metacognitive ability across experimental domains and methodological approaches, in which several aspects were taken into account that are considered to be critical to the investigation of this research question. It was found that a combination of domain-general (unity) and domain-specific (diversity) components most adequately describes the structure of metacognition. Moreover, two separate functional magnetic resonance imaging (fMRI) studies are described which examined the effects of a specific pharmacological challenge on the integrity of metacognitive processes in two different domains, episodic memory and visual perception: The *N*-methyl-D-aspartate (NMDA) glutamate receptor antagonist ketamine represents a promising candidate for pharmacological modulations of metacognition, as previous studies obtained indications of ketamine-induced alterations in related functions. Convergent evidence of both studies indicates an involvement of the glutamatergic neurotransmitter system in metacognition, as the precision of meta-level evaluations was found to be attenuated during acute ketamine administration, which may partially be compensated by re-representations of maintained object-level information. Overall, the studies presented in this dissertation provide novel contributions to the scientific understanding of the architecture of metacognition.

ZUSAMMENFASSUNG

Metakognition, die Fähigkeit zur Überwachung eigener mentaler Zustände und Prozesse, stellt eine wesentliche Voraussetzung für erfolgreiches menschliches Verhalten dar. Zentrale Fragen hinsichtlich der Architektur der Metakognition sind jedoch bisher ungeklärt. So ist beispielsweise offen, ob die Genauigkeit metakognitiver Bewertungen über verschiedene Aufgaben hinweg auf einer einzigen, gemeinsamen Grundlage basiert oder ob es möglich ist, trennbare metakognitive Subsysteme für unterschiedliche Anforderungen zu identifizieren. Dies stellt auch eine Herausforderung für die Untersuchung der neurobiologischen Prozesse dar, die metakognitiven Vorgängen zugrunde liegen, da Einblicke in die Struktur der Metakognition in einer funktionellen Domäne somit nicht ohne Weiteres auf andere Domänen übertragbar sind. Ferner ist nur wenig über die Grundlagen der Metakognition auf Ebene der Neurotransmitter-Systeme bekannt.

Die in dieser Dissertation beschriebenen Forschungsstudien hatten zum Ziel, Erkenntnisse zur Beantwortung dieser Fragen hervorzubringen, indem die Architektur der Metakognition in drei Originalarbeiten untersucht wurde. Zum einen wurde eine umfassende Verhaltensstudie zur Analyse des Musters individueller Unterschiede in der metakognitiven Fähigkeit über experimentelle Domänen und methodische Zugänge hinweg durchgeführt, in der mehrere Aspekte Berücksichtigung fanden, die als besonders wichtig für die Untersuchung dieser Forschungsfrage gelten. Es zeigte sich, dass die Struktur der Metakognition am angemessensten durch eine Kombination aus domänenübergreifenden (Einheit) und domänenspezifischen (Vielfalt) Komponenten repräsentiert wird. Darüber hinaus werden zwei separate Studien unter Verwendung von funktioneller Magnetresonanztomographie (fMRT) beschrieben, in denen die Auswirkungen einer spezifischen pharmakologischen Intervention auf die Integrität metakognitiver Prozesse in zwei verschiedenen Domänen, dem episodischen Gedächtnis und der visuellen Wahrnehmung, untersucht wurden: Für pharmakologische Modulationen der Metakognition gilt der *N*-methyl-D-Aspartat (NMDA)-Glutamat-Rezeptor-Antagonist Ketamin als vielversprechend, u.a. da frühere Studien Hinweise auf Ketamin-induzierte Veränderungen bei verwandten Funktionen erbrachten. Die übereinstimmenden Ergebnisse beider Studien legen eine Beteiligung des glutamatergen Neurotransmittersystems an der Metakognition nahe, da ermittelt wurde, dass die Genauigkeit von Metaebenen-Bewertungen während akuter Ketamin-Verabreichung eingeschränkt war, was jedoch möglicherweise durch die Re-Repräsentation von aufrechterhaltener Objektebenen-Information anteilig kompensiert werden kann. Insgesamt leisten die in dieser Dissertation beschriebenen Studien neue Beiträge zum wissenschaftlichen Verständnis der Architektur der Metakognition.

I. GENERAL INTRODUCTION

I.1 INTRODUCTION TO METACOGNITION

I.1.1 DEFINITION AND DISAMBIGUATION

To understand why the introspective ability for metacognition is of particular interest to experimental psychological research, it is useful to refer to the Swedish natural scientist Carl Linnaeus. In his influential publication *Systema Naturae* (1735), widely regarded as the basis of modern zoological and botanical taxonomy, Linnaeus meticulously annotated characteristics of numerous different species. When it came to *Homo sapiens*, however, he wrote down only "Nosce te ipsum" – “those that know themselves" (Fleming, 2021), quoting an inscription above the entrance to the temple of Apollo at Delphi. Self-awareness, perhaps best appreciated as a form of meta-representation of oneself and one’s cognitions, arguably constitutes an integral part of the human condition and may be part and parcel of what makes humans human in the first place (Grimaldi *et al.*, 2015; Kepecs *et al.*, 2008; J. D. Smith *et al.*, 2009).

This notion is intimately linked with what has been known as "metacognition" in experimental psychology since the 1970s. The term was coined primarily by the American developmental psychologist John H. Flavell, who originally defined metacognition as "knowledge and cognition about cognitive phenomena" (Flavell, 1979, p. 906), which he explored mainly in learning contexts. Expanding upon this, researchers have further defined metacognition as the knowledge about one’s own cognition and the use of this knowledge to regulate ongoing processes (Efklides & Vlachopoulos, 2012; Szczepanik *et al.*, 2020), or more broadly as the human ability to assess the accuracy of knowledge states (Barrett *et al.*, 2013) and to think about and monitor one’s own mental states and processes (Dunlosky & Metcalfe, 2008; S. Kim *et al.*, 2021; Molenberghs *et al.*, 2016). These broader conceptualizations are often simplified by defining metacognition as “thinking about thinking” (Reynolds & Wade, 1986) or “cognition about cognition” (Desender *et al.*, 2017) and typically extend far beyond learning and education. Ultimately, the capacity for metacognition is now recognized as a central element of a self-referential checks-and-balances system that enables successful human action across a range of contexts, with particular importance in complex social environments (Fleming, Dolan, *et al.*, 2012; Frith, 2012; Heyes *et al.*, 2020).

It is worthwhile to examine the conceptual evolution of metacognition more closely, as it helps to disentangle the construct from related functions. In his original model, Flavell (1979) conceptualized the monitoring of different cognitive enterprises as a function of the actions of and interactions among several classes of phenomena, two of which are of a metacognitive nature: *Metacognitive knowledge*, which represents task and strategy variables as well as knowledge structures about the multidimensional substrates of thoughts and actions of different agents, and *metacognitive experiences* (also termed metacognitive regulation), which refers to conscious cognitive or affective experiences

associated with ongoing behavior (Flavell, 1979). Subsequently, Nelson and Narens (1990) added a *monitoring* component to this conceptualization, which represents the use of metacognitive knowledge and experiences to guide behavior (Weil *et al.*, 2013). Their central premise was the assumption of two interacting levels (Figure 1): the object-level, which contains cognitions about objects, and the meta-level, containing a dynamic representation of the object-level composed of (meta-)cognitions about object-level cognitions. These levels interact in two ways: (1) in the form of information flow from object-level to meta-level (monitoring of ongoing cognitions) and (2) via control in the sense of behavioral adjustment initiated from meta-level to object-level (Nelson & Narens, 1990). Consequently, the epistemological status of metacognition is that of a meta-level representation of an object-level cognition; empirically, however, metacognition is operationalized mostly as "behavior about behavior" (Fleming, Dolan, *et al.*, 2012).

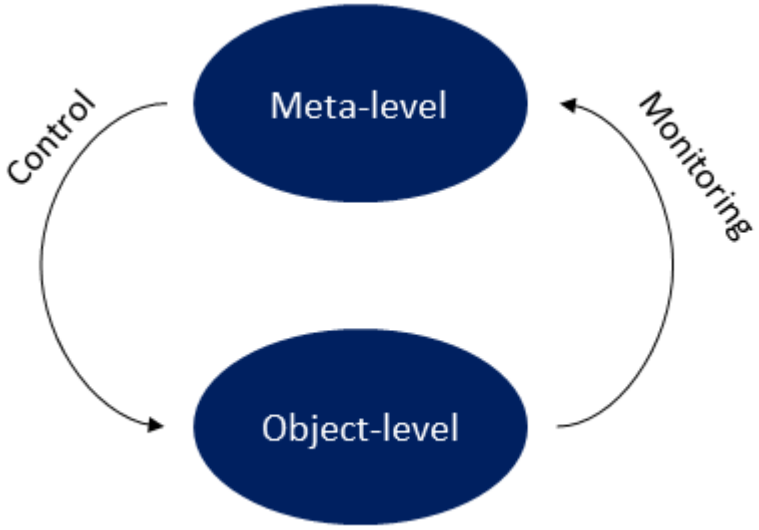

Figure 1. Theoretical mechanism proposed by Nelson and Narens (1990), wherein a meta-level can be distinguished from an object-level, and these two levels interact via monitoring and control pathways. Adapted from "Metamemory: A Theoretical Framework and New Findings" by T. Nelson and L. Narens, 1990, in G. Bower (Ed.), *The Psychology of Learning and Motivation*, p. 126; copyright (1990), with permission from Elsevier.

Metacognitive evaluations may further differ in the level to which they refer (Fleming, Dolan, *et al.*, 2012). For instance, the tripartite conceptualization of metacognition proposed by Metcalfe and Son (2012) distinguishes between anoetic metacognition (judgments about an event or object in the external world), noetic metacognition (judgments of an internal representation without explicit self-reference), and autonoetic metacognition (judgments of an internal representation with self-reference). However, anoetic judgments are primarily used to assess task performance and would not

be assigned to the meta-level in Nelson and Narens' (1990) model, since they refer to tangible *contents* (and thus object-level cognitions) rather than meta-level processes (Metcalfe & Son, 2012).

This issue serves as a demonstration that the ensemble of what is subsumed under the term is notoriously vague, so it is unsurprising that metacognition has been referred to as a fuzzy umbrella term (Baird *et al.*, 2015; Norman *et al.*, 2019; Scott & Levy, 2013). Distinguishing metacognition from related constructs has proven a challenging enterprise, which may be due to inconsistent use of terminology or differing goals and priorities within different subfields of psychology. These obstacles notwithstanding, it is possible to identify factors for disambiguation. The distinction between metacognition and the concept of *cognitive control*, for instance, arises from the fact that although (explicit) metacognitive representations are frequently supplied for use in cognitive control mechanisms, these can also be exerted implicitly and unconsciously (Shea *et al.*, 2014; Van Gaal *et al.*, 2010). Likewise, the terms metacognition and *intelligence* should not be used interchangeably: Although the monitoring of internal cognitive states is a defining characteristic of intelligent adaptive behavior (Hertzog & Robinson, 2005), the constructs are typically placed on different levels (J. H. H. Song *et al.*, 2021), as intelligence manifests itself primarily on the object-level. The distinction from the concepts of *mentalizing* or the *theory of mind* (ToM) appears particularly relevant, as both are based on recursive reasoning and trace back to at least partly common neural bases (D'Argembeau *et al.*, 2007; Vaccaro & Fleming, 2018). Mentalizing can be regarded as a subtype of metacognitive processes that come into play when metacognition is applied to others, such as to predict their behavior (Frith, 2012), and appears to represent a simulation of what oneself would do if one were in another person's position (Jenkins *et al.*, 2008; Vaccaro & Fleming, 2018). However, it has also been argued that the monitoring mechanism underlying metacognitive self-awareness is independent of ToM (Nichols & Stich, 2003). Finally, it is relevant to distinguish metacognition from *consciousness*, which altogether represents a broader, superordinate construct associated with a wider plethora of concepts than metacognition – e.g. wakefulness, arousal and the subjective qualia of mental representations (Clark *et al.*, 2019; Grimaldi *et al.*, 2015).

Nevertheless, consciousness is intimately linked to meta-representation (Shea & Frith, 2019), and higher-order cognitive monitoring has been argued to represent a necessary prerequisite for human consciousness (H. C. Lau & Rosenthal, 2011; Pasquali *et al.*, 2010; Timmermans *et al.*, 2012). In their computational account of metacognition, Timmermans *et al.* (2012) proposed a two-level network architecture – not unlike the theoretical mechanism outlined by Nelson and Narens (1990) – in which the second-order network receives information about the outcomes of the first-order network and applies this information to form confidence judgments. Consciousness is conceived in the model as a representational-redescriptive process along a complex hierarchy of three predictive loops (inner loop, self-other loop and action-perception loop) that gradually improve their efficacy to re-represent first-order mental states (Timmermans *et al.*, 2012). Thereby, the authors adopt a position in the tradition

of Karmiloff-Smith (1994), according to which metacognition should not be regarded as a consequence of content becoming conscious and therefore available to higher-order processes, but as a necessary prerequisite for consciousness: In order for content to become conscious, a system needs to be able to represent its internal states to itself, i.e., knowledge *in* the system has to become knowledge *for* the system (Timmermans *et al.*, 2012).

I.1.2 IMPORTANCE AND PATHOLOGICAL IMPLICATIONS

It is well established that the capacity for metacognitive monitoring constitutes a fundamental puzzle piece to successful human behavior (Metcalfe & Shimamura, 1994). For instance, it enables navigation in highly dynamic, complex environments by guiding the error-prone process of decision-making, as errors are made identifiable through efficient metacognitive evaluations (Allen *et al.*, 2016; Lapate *et al.*, 2020; Nisbett & Wilson, 1977). This beneficial effect is not limited to the individual: Given certain conditions at the group-level, metacognition improves shared decision-making by means of communicating one's own decision certainties (Bahrami *et al.*, 2010; Koriat, 2012). Generally, the functional benefits of metacognition may be most pronounced in the area of complex abilities, such as planning and reasoning, possibly as an extension of a system originally evolved to subserve sociocultural functions (Carruthers, 2009; Fleming, Dolan, *et al.*, 2012; Fletcher & Carruthers, 2012). Consequently, metacognition has been demonstrated to be elemental to the success of social interactions (Frith, 2012; Heyes *et al.*, 2020), conceptualized e.g. as a system of explicit supra-personal cognitive control in multi-agent processes (Shea *et al.*, 2014). Recently, Kuchling *et al.* (2022) argued that in evolutionary terms, a system of meta-level representations is in fact more energetically efficient than purely object-level cognition when natural selection operates at multiple timescales. While it is worth noting that engaging in metacognition can also have detrimental effects, as it may interfere with performance or give rise to unhelpful comparisons with others (Norman, 2020), the overwhelming body of research nonetheless suggests that metacognitive processes are fundamental to numerous areas of interest (Barrett *et al.*, 2013): This comprises the control of conscious thought (Schooler *et al.*, 2011), executive functions (Roebers, 2017), effective learning (Dienes & Seth, 2010; Veenman *et al.*, 2006), development (Weil *et al.*, 2013), prospective memory (Rummel & Meiser, 2013), the evaluation of phenomenal qualities of internally generated experiences (Pearson *et al.*, 2011), and processes underlying gradual perceptual awareness (Brogaard, 2011; Kanai *et al.*, 2010; Overgaard *et al.*, 2006).

It follows that when metacognition is pathologically impaired, as can occur e.g. as a result of brain tissue damage (Fleming *et al.*, 2014; Pannu & Kaszniak, 2005), there are widespread effects on a number of functional domains. Likewise, impairments in metacognition are a characteristic feature of several

psychiatric disorders, as the integrity of the metacognitive system is fundamental for the realization that some of one's thoughts and feelings are expressions of an illness (Lysaker, Dimaggio, *et al.*, 2011). Moreover, cognitive insight in neuropsychiatric disorders is underpinned by a neural network that bears close resemblance to the brain network of metacognition (David *et al.*, 2012; see Chapter I.2.1). The metacognitive profiles of patients with a psychiatric diagnosis are therefore of major interest in clinical settings (Massé & Lecomte, 2015), and it has even been postulated that each psychopathological condition is characterized by a unique set of metacognitive deficits, the alleviation of which could be elemental to achieve a successful psychotherapeutic outcome (Semerari *et al.*, 2003).

Importantly, metacognitive deficits appear to contribute to various aspects of the symptomatology of *schizophrenia*, with metacognitive distortions representing a cornerstone in the development and maintenance of psychotic symptoms in schizophrenia spectrum disorders, in particular in the memory domain (Charles *et al.*, 2017; Eisenacher *et al.*, 2015; Eisenacher & Zink, 2017; Moritz *et al.*, 2003). Overall, suboptimal metacognition can be postulated as a contributory factor to positive (Eichner & Berna, 2016; Eisenacher & Zink, 2017), negative (Hamm *et al.*, 2012; Lysaker *et al.*, 2005, 2015) and disorganized (Lysaker *et al.*, 2007) symptomatology of schizophrenia; the limitations seem to be most pronounced in patients with formal thought disorder (Köther *et al.*, 2012). Not only are patients with schizophrenia limited in their ability to compose complex ideas about themselves and others and to apply these ideas to meet psychosocial demands (Lysaker & Dimaggio, 2014); specifically, it has been shown that individuals with schizophrenia exhibit overconfidence (i.e., inordinately increased decision confidence) in incorrect responses, while being also less confident in correct responses when compared to control groups (Moritz *et al.*, 2005; Moritz & Woodward, 2006). This reduced *confidence gap* between correct and incorrect answers is indicative of diminished metacognitive sensitivity (cf. Chapter I.1.3).

Although some studies on metacognitive deficits in schizophrenia may have taken insufficient care to consider the confounding influence of task performance on estimates of metacognitive capacity (Rouy *et al.*, 2021), therapeutic intervention programs explicitly targeting metacognitive processes appear to represent a promising addition to the existing repertoire of psychosis therapies, e.g. as a complement to classical cognitive-behavioral therapy (Moritz & Woodward, 2007) or to pharmacotherapy (Moritz *et al.*, 2014). For instance, the efficacy of metacognitive training in psychosis (Moritz *et al.*, 2014; Vitzthum *et al.*, 2014), which focuses on the exposure of psychotic thought biases and dysfunctional problem-solving behaviors, has been confirmed in several independent meta-analyses (e.g. Lopez-Morinigo *et al.*, 2020; Sauvé *et al.*, 2020), and there are positive indications for application in patients with *borderline personality disorder* (BPD; Schilling *et al.*, 2018), especially as an early therapeutic target, since the modulation of metacognitions has a pronounced impact on therapeutic alliance (Dimaggio *et al.*, 2019). Much like schizophrenia, BPD is characterized by a specific pattern of metacognitive deficits that may

be at the core of the problems that BPD patients encounter in their everyday lives (Lysaker *et al.*, 2017; Vega *et al.*, 2020).

Therapeutic approaches focusing on the modulation of metacognitive elements have also been developed for *attention deficit hyperactivity disorder* (Solanto *et al.*, 2010), *post-traumatic stress disorder* (Wells & Sembi, 2004), and recurrent *depression* (Wells *et al.*, 2009) – the extent to which the nature of the respective therapies is a *meta*-cognitive one, however, is the subject of controversy (Moritz *et al.*, 2018; Wells & Fisher, 2011). Furthermore, the crucial role of metacognitive beliefs in the development and maintenance of generalized, repetitive, and seemingly uncontrollable worry in the context of *generalized anxiety disorder* (GAD; Wells, 1995, 1999, 2005) or *addictive behaviors*, such as excessive drinking (Spada *et al.*, 2007, 2014; Spada & Wells, 2008, 2009), was highlighted in several models. Dysfunctional metacognitive beliefs and thought control strategies also seem to contribute to the core symptomatology of *obsessive-compulsive disorder* (OCD; Irak & Tosun, 2008; Park *et al.*, 2020; Wells & Papageorgiou, 1998). Generally, psychiatric symptom dimensions related to major depression and anxiety are associated with lower overall confidence (Drueke *et al.*, 2022; Moses-Payne *et al.*, 2019) and increased metacognitive accuracy, whereas a dimension comprising intrusive thoughts and OCD-related behavior could be associated with the reverse pattern, i.e., higher metacognitive bias (see Chapter I.1.3) and lower metacognitive accuracy (Rouault, Seow, *et al.*, 2018).

I.1.3 MEASUREMENT OF METACOGNITION

Empirical measures of metacognition differ with respect to the level of resolution at which they examine metacognition. For instance, the model of (metacognitive) awareness outlined by Toglia and Kirk (2000) describes the interaction between “online” awareness, activated as a regulating and monitoring process within a given situation or task, and longer-term, multi-layered metacognitive assumptions about oneself (motivational dispositions, self-agency expectations etc.) that exist across a number of situations and may effectively be understood as a latent disposition guiding behavior across different requirements (Toglia & Kirk, 2000). Consequently, “online" measures are those methods that are obtained during specific situations or tasks, i.e., they are characterized by their explicit relation to a particular metacognitive action in immediate temporal association with task performance (Veenman *et al.*, 2006), typically in a laboratory setting. In contrast, "offline" measures target the attitudes, experiences, and beliefs individuals have about themselves as (meta-)cognitive beings and do not refer to a single specific demand.

Due to their trans-situational nature, offline measures are also considered as measures of "metacognition in everyday life" and are conventionally obtained in the context of self-report measures

such as questionnaires or interviews (Akturk & Sahin, 2011; Rouault, McWilliams, *et al.*, 2018; Saraç & Karakelle, 2012). Therefore, all known drawbacks of self-report measures must be considered, such as untruthful or socially desirable responding, unreliable introspections, memory distortions, desire for positive self-representation, or comprehension issues (Akturk & Sahin, 2011; Baker & Cerro, 2000; Norman *et al.*, 2019). Bearing this in mind, it may be constructive to relate the self-report measure to or supplement it with an informant rating (Fitzgerald *et al.*, 2017). However, particularly in terms of test economy, there are considerable advantages to self-report questionnaires as offline measures of metacognition, as they can easily be administered to groups and evaluated quickly and objectively (Tobias & Everson, 1996, as cited in Akturk & Sahin, 2011). While several questionnaires have been developed which tap specific aspects of metacognitive monitoring by collecting standardized responses to various items, an additional complicating component in the offline measurement of metacognition is that none of the available questionnaires contain unambiguous representations of metacognitive behavior and processes in everyday situations, as they were developed in different contexts with different goals of application (see Chapter II.2 for further detail). For instance, some of the questionnaires focus on the role of metacognitive styles in the onset and maintenance of pathologically relevant patterns of thought and action, whereas others emphasize learning processes and are optimized for educational contexts. This raises the difficulty of obtaining a reflection of metacognitive processes that is as realistic and applicable to everyday life as possible (Saraç & Karakelle, 2012).

Upon consideration of model accounts, such as the Toglia and Kirk (2000) model in which metacognition is conceptualized as a trait, i.e., as a time-consistent disposition, it is often to a certain degree presupposed in psychological research that online and offline measures of metacognition draw upon a common resource (Veenman *et al.*, 2006). Convergent validity of the two sets of methodological approaches (or "method-groups") would be highly desirable in order to take advantage of the aforementioned benefits of offline self-report measures, which in this case could serve as a proxy for performance-related metacognition measures in a specific task (Baker & Cerro, 2000; Schellings & Van Hout-Wolters, 2011). The extent to which this is consistent with empirical data will be explored in detail in Chapter I.1.4.

For online measures of metacognition, on the other hand, it is import to address another relevant distinction, which concerns the relative timing of metacognitive evaluation: If a metacognitive judgment is obtained prior to task performance, e.g. if a prediction is to be made about what information will be available in memory at a given time in the future, one would refer to such as a judgment of *prospective metacognition* (Grimaldi *et al.*, 2015; Miyamoto *et al.*, 2021). The counterpart – r*etrospective metacognition* – represents the default approach in online metacognition research and signifies judgments about a (typically recent) past experience (Grimaldi *et al.*, 2015). Although prospective metacognition judgments can also be obtained in classical decision-making paradigms using either

global or item-by-item pre-performance estimates (Fleming & Dolan, 2012; Rouault *et al.*, 2019), assessments of prospective metacognition are common only within the memory domain (Jang *et al.*, 2020; Mazancieux, Dinze, *et al.*, 2020; Metcalfe, 1986), where they are implemented e.g. via so-called judgments of learning or feelings of knowing (Fleming & Dolan, 2012).

A key challenge for both method-groups is to objectively assess the concept of metacognition, being inherently subjective and based on introspective processes (Grimaldi *et al.*, 2015). In most online assessments of metacognition, this is accomplished by prompting participants to indicate their subjective decision confidence in the immediate follow-up to a cognitive or perceptual judgment. This requires participants to adjust their secondary, metacognitive response to the primary response. To account for the dimensionality of subjective decision certainty (Overgaard *et al.*, 2006; Sandberg *et al.*, 2010), confidence ratings – unlike the typically binary first-order decision – are commonly provided in different gradients (Grimaldi *et al.*, 2015), which may e.g. take the shape of discrete fixed levels or percentages ranging from complete uncertainty (0%) to complete certainty (100%). It has been suggested that such measures are too abstract and that participants may underestimate their confidence or even withhold conscious knowledge, as they are given too little motivation to reveal it (Persaud *et al.*, 2007). Post-decision wagering has therefore been proposed as an alternative, whereby participants bet high or low amounts of monetary value on the correctness of a previous decision, depending on their subjective sense of certainty (Persaud *et al.*, 2007). However, such wagering measures were demonstrated to suffer from low sensitivity to intermediate confidence ranges and are also strongly confounded with interindividual differences in risk aversion (Fleming & Dolan, 2010; Grimaldi *et al.*, 2015; Sandberg *et al.*, 2010). Overall, post-decisional confidence ratings seem to allow for the assessment of the widest range of subjective states of awareness (Wierzchoń *et al.*, 2012). Nonetheless, it needs to be noted that all available online measures of metacognition that rely on collecting self-reported information during a task lead to some degree of reactivity, i.e., the act of collecting confidence ratings can impact both first-order performance and metacognitive processes (Double & Birney, 2019; Lei *et al.*, 2020).

Going forward, it is necessary to establish some terminology; see Table 1 for a glossary of the most important terms introduced in this chapter. The primary task in relation to which metacognition is assessed is typically referred to as the *Type I task*. The accuracy by which individuals perform on that task is analogously termed *Type I performance* (Maniscalco & Lau, 2014). The fact that this Type I performance is typically obtained through binary categorizations, i.e., the classification of a stimulus as belonging to one or the other stimulus category (e.g. old versus new, left versus right) is relevant for the application of signal detection theory (SDT) methodology (Galvin *et al.*, 2003; Green & Swets, 1966). Among several other advantages detailed below, SDT methodology allows to account for whether the distinction was made between two qualitatively different (e.g. stimulus present versus stimulus absent)

or two qualitatively similar (e.g. upper or lower stimulus was included in study phase) stimulus categories. The metacognitive confidence judgment that follows the Type I task on a given trial then constitutes the *Type II task*, and the task-related ability of individuals to discriminate between their own correct and incorrect Type I responses by means of confidence judgments is referred to as *Type II performance*, or alternatively *metacognitive sensitivity*.

Table 1. Glossary of important terms in the context of measuring metacognition.

Term	Description	Relevant metric
Offline measures	Generalized estimates of metacognitive ability across situations and demands in the absence of a specific task	-
Online measures	Specific estimates of metacognitive performance obtained during an ongoing demand, e.g. an experimental task	-
Retrospective measures	Metacognitive reports obtained after a corresponding first-order rating	-
Prospective measures	Metacognitive reports obtained prior to a corresponding first-order rating	-
Type I task	The first-order object-level task in relation to which metacognition is assessed	-
Type I response bias	An individual's propensity to report one first-order stimulus category over another	-
Type I performance/ Type I sensitivity	An individual's ability to correctly discriminate between first-order stimulus categories	d’
Type II task	The metacognitive confidence judgment part of the trial sequence	-
Type II response bias/ Metacognitive bias	An individual's propensity to report high or low confidence ratings	-
(Absolute) Metacognitive sensitivity*	An individual’s ability to discriminate between their own correct and incorrect judgments on the Type I task by means of confidence ratings	meta-d’
Relative metacognitive sensitivity/ Metacognitive efficiency*	An individual’s ability to discriminate between their own correct and incorrect judgments on the Type I task by means of confidence ratings, corrected for their Type I performance	meta-d’/d’

Note: * both absolute and relative metacognitive sensitivity regularly serve as online measures of Type II performance, which will be used synonymously with metacognitive performance, metacognitive accuracy, and metacognitive ability.

For the purpose of quantifying Type II performance as accurately as possible, a number of confounding factors which may exert a confounding influence on the parameters of metacognitive sensitivity must be taken into account, which is why the simple assessment of the across-trial correlation between Type I accuracy and Type II confidence is regarded as insufficient (Grimaldi *et al.*, 2015; Masson & Rotello, 2009). For instance, in comparatively simple Type I tasks, it is easier to discriminate between correct and incorrect response trials, whereas in difficult Type I tasks, raw metacognitive sensitivity scores are bound to be lower (Fleming & Lau, 2014). Such confounding of Type I and Type II performance is problematic given both behavioral and neural dissociations between performance and metacognition (Fleming & Dolan, 2012; Zylberberg *et al.*, 2012). In addition, response biases both at the Type I and the Type II level can distort estimates of Type II performance unless they are adequately accounted for methodologically: This refers to an individual's Type I (propensity to report one stimulus category over another) and Type II (propensity to report high or low confidence levels) response criteria (Barrett *et al.*, 2013); the latter is frequently labeled *metacognitive bias* and is itself the subject of various investigations (Fleming & Lau, 2014). These confounding factors are intrinsically accommodated in the SDT-based meta-d' framework; in Chapter II.1.1, the quantification of online metacognitive ability (commonly labeled *metacognitive efficiency* when referring to an index corrected for confounding factors) in this framework will be elaborated in further detail.

The overwhelming majority of Type I tasks in relation to which metacognition is assessed pertain to either the perceptual or memory domain (Rouault, McWilliams, *et al.*, 2018; Samaha & Postle, 2017). However, within these domains, numerous differentiations can be made: Metacognition in the perceptual domain (i.e., “metaperception”) has been studied using paradigms involving both moving (Rausch *et al.*, 2015; Wokke *et al.*, 2017) and static stimuli, which may differ e.g. in their contrast or orientation (C. Song *et al.*, 2011), but most commonly involve the assessment of the magnitude of given sets (Arbuzova *et al.*, 2021; Fleming *et al.*, 2016; Moeller *et al.*, 2016). In the memory domain, retrospective online measures of “metamemory” are typically acquired in relation to retrieval of words (Baird *et al.*, 2013; Sadeghi *et al.*, 2017), albeit other stimulus types such as abstract shapes (Morales *et al.*, 2018) and faces (Busey & Arici, 2009; Pannu *et al.*, 2005) have also been subject to investigation. In addition, a few customary domain-specific features in the methodological implementation are of note: For instance, the individual task performance of a person in the perceptual domain is regularly constrained through so-called staircase procedures, which adapt the difficulty of a task to the participant’s ability. This allows for keeping average performance of all participants at a fixed, constant level (Macmillan & Creelman, 2005; Wright *et al.*, 2012) and thereby accounts for the confounding influence of Type I sensitivity on metacognitive performance (Allen *et al.*, 2017; Fitzgerald *et al.*, 2017). For the memory domain, it is characteristic that encoding and retrieval of mnestic stimuli typically take place within the same study session (Busey & Arici, 2009; Honey, Honey, Sharar, *et al.*, 2005). Encoding

of stimuli may either occur in the absence (Fleming *et al.*, 2014; McCurdy *et al.*, 2013) or in the presence of a specific task instruction, such as pleasantness ratings of words (Honey, Honey, Sharar, *et al.*, 2005) or likeableness ratings of faces (Bower & Karlin, 1974). Two-choice recognition paradigms are more common for retrieval than free recall paradigms, as they facilitate trial-by-trial measurements of metacognitive confidence and are associated with varying degrees of familiarity (Souchay *et al.*, 2013).

The clear focus of metacognition research on metaperceptual and metamnestic processes, however, comes with the risk that relevant properties of metacognition more closely associated with other task domains remain undiscovered or unconnected to the epistemic repertoire of metacognition research. For instance, the disparity of literatures regarding the fields of error monitoring and metacognitive confidence has been pointed out, as both are frequently considered different types of self-evaluation (Fleming, Dolan, *et al.*, 2012; Fleming & Daw, 2017) despite strong evidence for shared neural mechanisms, as indicated by a distinct and graded confidence-related modulation of the electrophysiological error positivity (Boldt & Yeung, 2015). Indeed, relevant methodological dissimilarities emerge in that decision confidence studies frequently manipulate the quality of evidence (e.g. the ambiguity of a perceptual discrimination), whereas in error awareness research, it is typically the quantity of evidence which is manipulated, e.g. by imposing time pressure on participants (Yeung & Summerfield, 2012). Another reason for error monitoring and metacognitive confidence to inhabit such disparate research landscapes may lie in the assumption of a more implicit system of conflict and error monitoring, whereas metacognition is traditionally associated with explicit monitoring and control (Ridderinkhof *et al.*, 2004; Weil *et al.*, 2013). However, once the experimental setup requires participants to report errors and e.g. provide a metacognitive judgment about the accuracy of their response, the monitoring process tapped in such behavioral inhibition tasks is of an explicit nature. Ultimately, error awareness may be understood as a genuine metacognitive process integral to successful human action (Osman, 2010), since humans have metacognitive access to a multitude of performance parameters (Di Gregorio *et al.*, 2020), as outlined in more detail below.

Said disparity may also partially be attributed to the longstanding lack of an appropriate computational framework that subsumes both metacognitive confidence and error awareness (Fleming & Daw, 2017; Yeung & Summerfield, 2012). The fairly novel Bayesian framework of second-order metacognitive computation by Fleming and Daw (2017) aims to provide such a unified account of confidence and error monitoring. To apprehend this model, it is useful to briefly review the controversy on the informational basis of metacognitive processes. According to the direct-translation hypothesis, confidence ratings rely on the same sources of information that are used for primary task performance (Higham *et al.*, 2009; Kepecs & Mainen, 2012). This perspective is incorporated into the structure of many SDT approaches for quantifying metacognition, which map a continuum from absolute decision uncertainty to absolute decision certainty (Maniscalco & Lau, 2012). Other studies have pointed out that performance and

confidence may also draw upon different sources of information, as sensory evidence continues to be accumulated after an action or decision (Gajdos *et al.*, 2019; Klein *et al.*, 2013; Moran *et al.*, 2015). This suggests the formation of confidence to at least partly be based on post-decisional processes (Hilgenstock *et al.*, 2014; Yeung & Summerfield, 2012), which can include the appraisal of a given decision to be incorrect (Moreira *et al.*, 2018) and therefore accommodate a sense of error information that extends past absolute decision uncertainty (Fleming & Daw, 2017). Cases of supra-optimal metacognitive sensitivity (i.e., going over and beyond what would be expected by the associated task performance) are also incompatible with the direct-translation hypothesis (Fleming & Daw, 2017). However, post-decisional models cannot explain that confidence judgments are more strongly dependent on a signal's sensory reliability than task performance, since both sets of information were already available prior to the Type I decision (Boldt *et al.*, 2017). Fleming and Daw's (2017) model of second-order computation considers metacognitive operations as second-order inferences on a coupled, yet functionally distinct decision system. Accordingly, the decision and confidence variables are represented as two correlated hidden states, which ultimately implies that the metacognitive system infers the behavior of a different agent, particularly when internal states about decisions and metacognitive evaluations are separated across space and time (Fleming & Daw, 2017). Likewise, the finding that actions themselves influence confidence judgments is naturally accommodated in the second-order account, whose model architecture, along the lines of post-decisional models, posits a continuum from absolute certainty of having committed an error to absolute certainty in a correct response (Fleming & Daw, 2017).

In the field of error awareness research, paradigms with high susceptibility to the emergence of erroneous responses are typically employed, which can then be used to assess awareness of these errors. Unlike metacognition research with its traditional focus on perception and memory domains, this is mainly achieved via behavioral inhibition paradigms in the domain of attention-to-action. One such paradigm is the antisaccade task (Bialystok *et al.*, 2006; Nieuwenhuis *et al.*, 2001), an oculomotor paradigm in which the gaze is to be directed to the exact opposite position of an appearing cue stimulus, thereby requiring participants to inhibit the reflexive response of looking at the cue, with errors not being uncommon (Hutton & Ettinger, 2006). The error awareness task by Hester *et al.* (2005, 2012), a modified Go/No-Go task with two different No-Go conditions, is another example for a paradigm in which inhibition errors are provoked, with respect to which metacognitive insight into the error detection process can be investigated.

I.1.4 GENERALITY VERSUS SPECIFICITY OF METACOGNITION

The question of whether the capacity for metacognitive monitoring can be construed as modular and thereby fragmented into a number of different subsystems, or whether it represents a single, trait-like disposition applied across demands, contexts and domains (Figure 2), has been subject to decade-long controversy.

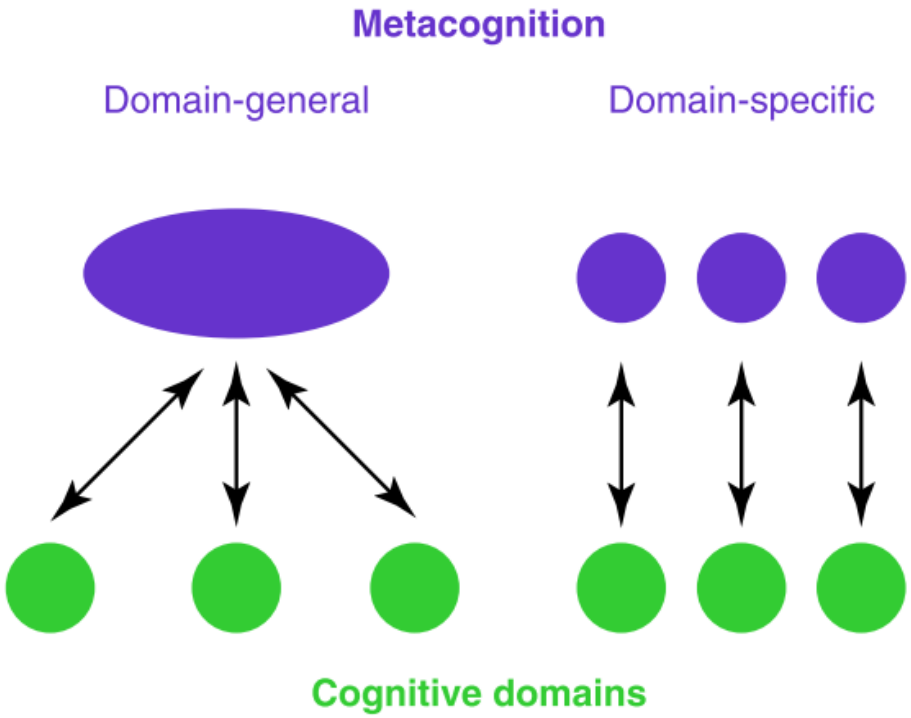

Figure 2. Conflicting conceptualizations whether metacognitive judgments are determined by a general, unitary resource across domains (left) or whether metacognition relies entirely on domain-specific components (right). Republished from "Human Metacognition Across Domains: Insights from Individual Differences and Neuroimaging" by M. Rouault, A. McWilliams, M.G. Allen and S.M. Fleming, 2018, *Personality Neuroscience, e17*, p. 2; in accordance with Cambridge University Press Copyright Policy, permission is granted under the Creative Commons Attribution-NonCommercial-NoDerivatives license (https://creativecommons.org/licenses/by-ncnd/4.0/).

One might plausibly assume that metacognition – in analogy to another major higher-order construct, intelligence (Spearman, 1904) – converges on a single G-factor that underlies an individual's ability to metacognitively monitor their ongoing processes and decisions and should hence be tapped by any kind of real measurement of metacognition (Mazancieux, Fleming, *et al.*, 2020). However, it might be an equally plausible assumption that metacognitive judgments are generated as a function of the underlying demands, i.e., that separate monitoring systems have been formed for e.g. different elements in the pathway of human information processing, such as perception, attention, or memory. This would be consistent with the view that modularity is deeply embedded into the structural

organization of the central nervous system and, ultimately, of human personality and behavior (Grigsby & Schneiders, 1991). Consequently, it needs to be determined whether the structure of metacognition (i.e., the organization of its functional properties) is characterized by a high degree of generality (i.e., individual differences covary between different points of measurement in a substantial manner) or by a high degree of specificity (i.e., very small or insignificant correlations among different estimates of metacognitive accuracy).

If high unity in metacognitive processes could be assumed, inferences about performance in one domain or requirement could be considered as estimates of performance in other domains or requirements as long as these share instrumental characteristics – which may, however, represent a key obstacle for observing commonalities between measures of metacognition across method-groups. As will become apparent from Chapter I.2.1, there is a fairly high convergence of findings from neuroimaging studies on metacognition, suggesting a central role of prefrontal or posterior-medial frontal structures (Fleck *et al.*, 2006). For example, Morales *et al.* (2018) were able to predict confidence in a perceptual task based on activation patterns in a memory task and vice versa; hence, confidence appears to covary with task-independent neural representations. Similarly, Rouault *et al.* (2022) described a shared brain system for forming confidence judgments in the domains of semantic memory and duration perception, namely ventromedial aspects of the prefrontal cortex (PFC). However, there is also an indication of domain-specific signals at the neural level (McCurdy *et al.*, 2013), as there seem to exist distinct connectivity patterns for different Type I modalities within the anterior PFC (aPFC), depending on whether confidence judgments succeeded perceptual or mnestic decisions (Baird *et al.*, 2013), which could express the marking of metacognitive representations with task-specific information (Baird *et al.*, 2014, 2015; Rouault, McWilliams, *et al.*, 2018). Nevertheless, the neural profiles of metamemory and metaperception, as laid out in a meta-analysis by Vaccaro and Fleming (2018), show substantial overlaps, as will be elaborated on in greater depth in Chapter I.2.1. Further evidence for a domain-general resource underlying the accuracy of metacognitive judgments could be derived from Carpenter *et al.* (2019), who suggested transfer effects across domains by providing feedback on metacognitive performance; note, however, that two more recent studies were unable to obtain convergent findings (Haddara & Rahnev, 2022; Rouy *et al.*, 2022).

A characteristic feature of many studies advocating in favor of a domain-general account is a high degree of similarity between the studied tasks with respect to instrumental or stimulus characteristics, and significant cross-task correlations have most frequently been observed within the perceptual domain, e.g. across sensory modalities (Faivre *et al.*, 2018). Meanwhile, Ais *et al.* (2016) reported a correlation in metacognitive sensitivity only for two of four perceptual tasks, whereas metacognitive bias proved to be highly consistent across all four paradigms. McCurdy *et al.* (2013) and Palmer *et al.* (2014) provided two of only few reports of a significant correlation between metacognitive accuracy in perception and

episodic memory tasks; in a recent study by Mazancieux, Fleming, *et al.* (2020), in which metacognition was investigated across four task domains (visual perception, episodic memory, semantic memory, and working memory/attention), the intercorrelation between perception and episodic memory was not confirmed. Remarkably, however, the correlations of all other task pairings reached significance, which suggests the substantial contribution of a domain-general resource. In general, the line of reasoning in favor of domain-generality in task-based metacognition is consistent with the notion that it is more efficient to classify performance within a global, task-independent framework that can be used to inform future actions or decisions (de Gardelle & Mamassian, 2014; Rouault, McWilliams, *et al.*, 2018).

On the other hand, overgeneralization across different contexts and demands could also prove to be maladaptive, as evidenced in psychiatric disorders (Rouault, McWilliams, *et al.*, 2018), which would be less of an issue if the monitoring system was ultimately idiosyncratic to a specific first-order requirement. And indeed, evidence for domain-specificity of task-based metacognition emerged in several studies. For instance, Kelemen *et al.* (2000) observed no significant commonalities in metacognitive accuracy across different types of metamemory judgments, and a number of studies failed to obtain significant associations between metaperception and metamemory (Baer *et al.*, 2021; Baird *et al.*, 2013, 2014, 2015; Fitzgerald *et al.*, 2017; Morales *et al.*, 2018; Ruby *et al.*, 2017; Sadeghi *et al.*, 2017). Consequently, a systematic review and meta-analysis by Rouault, McWilliams *et al.* (2018) across behavioral studies yielded significant correlations across perceptual metacognition tasks, but no significant cross-domain association of metaperception and metamemory. Within the memory domain, metacognitive accuracy was shown to vary depending on whether retrospective or prospective judgments were collected (Kelemen *et al.*, 2000; Maki & Swett, 1987), with lower sensitivity in prospective tasks (Siedlecka *et al.*, 2016). Recently, Mazancieux, Dinze, *et al.* (2020) found that metacognitive efficiency was correlated between episodic and semantic memory tasks only for retrospective, but not for prospective memory judgments. The study by Fitzgerald *et al.* (2017) occupies a rather distinctive position among the aforementioned studies arguing in favor of a domain-specificity account, since in this study, behavioral inhibition in the domain of attention-to-action was measured in addition to perception and memory tasks by means of a modified error awareness task (Hester *et al.*, 2005). There was no relevant correlation in metacognitive performance across the three investigated domains, although a correlation for response latencies across the tasks was argued to indicate shared programming mechanisms for different kinds of metacognitive responses.

Meanwhile, the investigation of the generality or specificity of metacognition across different method-groups represents a no less challenging undertaking. As can be inferred from the model of Toglia and Kirk (2000), task-based expressions of metacognitive awareness are assumed to influence and be influenced by superordinate, longer-term metacognitive assumptions about oneself. Consequently, the supposedly latent factors should therefore at least partially manifest within task-based metacognitive

variation. However, methodological hurdles are particularly imposed on making these latent factors measurable, as the gaining of realistic insight into one's own monitoring processes is clearly not independent of metacognitive ability, not least since these processes might be beyond verbal accessibility for many individuals (Kircher *et al.*, 2007; Lysaker, Buck, *et al.*, 2011).

This might contribute as to why only few studies have investigated this link thoroughly, obtaining heterogeneous findings: Several studies hinted at links between task-based metacognition and self-reported mental health deficits (Rouault, Seow, *et al.*, 2018) and a tendency towards radical political beliefs (Rollwage *et al.*, 2018). With respect to the Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994), which was developed primarily to tap metacognitive strategies within an educational context, an indication of a weak relationship between the instrument and retrospective metacognitive confidence judgments was provided by Sperling *et al.* (2004). However, Schraw and Dennison (1994) found that only prospective pre-test ratings, but not retrospective judgments were related to the "Knowledge about Cognition" component of the MAI. Likewise, no relationship emerged between cognitive self-insight as measured by the Beck Cognitive Insight Scale (BCIS; Beck *et al.*, 2004) and metacognitive performance in a perceptual task (Fleming, Huijgen, *et al.*, 2012). The short form of the Metacognitions Questionnaire (MCQ-30; Cartwright-Hatton & Wells, 1997; Wells & Cartwright-Hatton, 2004), which is applied quite regularly to measure metacognition offline despite being developed primarily to map clinically relevant monitoring tendencies and metacognitive beliefs in the context of GAD (Wells, 1995, 1999, 2005), was associated with metacognitive performance in a matrix reasoning task, albeit only in those individuals with high levels of (task-based) metacognitive ability (Grossner *et al.*, 2021).

Of additional interest in the offline approach to measuring metacognition is the discrepancy between self-reported and other-reported (meta-)cognitive functioning, with a low discrepancy typically considered as an indicator of accurate metacognition, while a high discrepancy would suggest a bias in the sense of self-underestimation or overestimation (Harty *et al.*, 2013). Accordingly, Harty *et al.* (2013) described an overall negative correlation between the discrepancy measure in the Cognitive Failures Questionnaire (CFQ; Broadbent *et al.*, 1982) and online error awareness, with younger adults (non-significantly) underestimating and older adults significantly overestimating their memory functioning and attentional control. It was found that the more participants underestimated themselves relative to informant reports, the higher their error awareness, which would lend support to the somewhat intuitive conclusion that a comparatively higher level of insight is accessible to self-report. Meanwhile, Fitzgerald *et al.* (2017), who investigated the question of generality and specificity of metacognition not only across experimental domains, but also across method-groups, found limited error awareness in individuals who underreported memory failures and attentional lapses in everyday life relative to informant reports on the CFQ to be associated with metacognitive ability in the perceptual domain

(Fitzgerald *et al.*, 2017). Metacognitive performance in the perceptual task could further be predicted by the informant rating rather than the self-rating, so it would appear more likely that an association of offline estimates with online metacognitive performance is possible on the basis of external assessment, which is assumed to be grounded to a greater extent in objectively observable behavior (Pronin, 2008).

However, for a majority of the aforementioned studies that examined the structure of metacognition in terms of its generality or specificity across domains or method-groups, a series of methodological problems emerge. For instance, in several studies, the laboratory measures of metacognition were not quantified via SDT methods such as the meta-d' framework, which leads to the inherent problems noted in Chapter I.1.3, such as insufficient control over the confounding influence of response bias and task performance. While the meta-analysis by Rouault, McWilliams, *et al.* (2018) examining the intercorrelation of different task-based estimates of metacognition only included studies that applied SDT methodology, the authors point out that the outcome of the meta-analysis may nonetheless be biased in favor of the domain-specificity hypothesis. Generally, the lack of cross-domain associations of online metacognitive performance may be due to differences in metacognition metrics as well as differences in task requirements between studies, e.g. two-alternative-forced-choice (2AFC) versus Yes/No judgments, whereby the establishment of a reliable connection across tasks and domains becomes an even more intricate endeavor (A. L. F. Lee *et al.*, 2018).

Beyond that, a different, yet equally important question must be raised over many studies advocating the domain-specificity hypothesis on the basis of small or non-significant correlations, which concerns statistical power. For instance, Fitzgerald *et al.* (2017) obtained data from no more than thirty participants, and although application of a refined method must be acknowledged as stimulus characteristics were matched across all tasks to control for the confounding influence of task properties on metacognitive accuracy, it is nonetheless likely that the sample size might have been insufficient to detect a stable correlation. Accordingly, in the relatively few cases in which sample sizes consistent with recommendations for testing individual differences within measures of the same construct were employed (Schönbrodt & Perugini, 2013), indications of at least a proportionate domain-general resource were obtained markedly more often (e.g. Mazancieux, Fleming, *et al.*, 2020).

A promising procedure when combined with a sufficiently large sample, which none of the cited studies implemented, is the application of a latent variable approach. Similar to the matching of local task properties by Fitzgerald *et al.* (2017), this represents an attempt to deal with the so-called *task impurity problem* (Burgess, 1997), whereby differences in task requirements may obscure correlations between tasks and domains. The correlational approach adopted in the aforementioned studies is particularly susceptible to the task impurity problem, whereas the latent variable approach represents a viable alternative, as the degree of commonalities and differences in metacognitive processes across different

experimental domains is examined at the level of population parameters (i.e., latent variables) rather than single tasks (Schermelleh-Engel *et al.*, 2003). Finally, in addition to the classical notion of the two conflicting perspectives of generality versus specificity, the latent variable approach also allows to map the combination of both phenomena. Beyond a fully unified and a fully modularized structure, a so-called unity and diversity structure, i.e., a fairly balanced coexistence of substantial domain-general and domain-specific signals, is also to be found in higher-order mental functions, as such an organization was demonstrated for executive functions by use of a latent variable approach (Miyake *et al.*, 2000; Miyake & Friedman, 2012). Ultimately, given the heterogeneous findings regarding generality and specificity across task domains and method-groups presented above, such a structure might plausibly be assumed for metacognition, as well.

I.2 BIOLOGICAL SUBSTRATES OF METACOGNITION

I.2.1 NEURAL SUBSTRATES

A major scientific focus of the past thirty years has been set on the identification of the neural correlates of specific facets of consciousness (Crick & Koch, 1990; LeDoux *et al.*, 2020). The widespread availability of neuroimaging techniques has provided a major impulse for consciousness research and has allowed piecing together a mosaic of the neural basis of metacognitive processes. A brief summary of the current scientific understanding of this mosaic will be given in the following, focusing on studies employing functional magnetic resonance imaging (fMRI).

Convergent evidence points to the existence of several "hot zones" of the neural correlates of consciousness (NCCs), roughly subdividing into a more anterior frontoparietal network (Baars, 2005; Vaccaro & Fleming, 2018) essential for explicit reports and online monitoring of ongoing action sequences, and a more posterior cortical hot zone encompassing temporo-parietal-occipital areas (Koch *et al.*, 2016). The latter is thought to provide the most promising anatomical candidates for full and content-specific NCCs supporting phenomenological distinctions, such as for the conscious perception of faces (Koch *et al.*, 2016). Processes located to the posterior hot zone may also be considered enabling factors for higher-order cognitive functions such as metacognition, as they provide sensory input to metacognitive evaluations (Bang & Fleming, 2018). In particular, the posterior parietal cortex (PPC) has been linked fairly intimately with metacognitive functioning: For instance, Humphreys and Lambon Ralph (2017) demonstrated domain-general involvement of the lateral PPC in higher-order cognitive functions, while Elman *et al.* (2012) observed parietal activity during metacognitive judgments in episodic memory retrieval. When contrasted with recognition-related activation patterns, higher

relative activation during confidence assessments was found in medial and lateral parietal regions, in line with an association of said areas with internally directed cognition (Chua *et al.*, 2006, 2009).

Prefrontal regions were also activated in this contrast, consistent with the multitude of studies suggesting a major involvement of the PFC, an area generally associated with higher cognitive functioning (Frith & Dolan, 1996) and widely regarded as one of the core areas of confidence formation (Bang & Fleming, 2018). The involvement of both posterior and prefrontal cortical areas could arguably be characterized as the neural manifestation of a large-scale cognitive control network (Paul *et al.*, 2015) incorporating information by systems dedicated to monitoring one's own mental states, as monitoring and control are considered to be functions of partially overlapping brain systems (Boldt & Gilbert, 2022). Most importantly, however, these regions appear to interact via a hierarchical loop between object-level processes (as expressed by posterior cortical activation) and meta-level processes anchored within the PFC (Fleming & Dolan, 2012).

In general, the maintenance and integration of information about preceding decisions provided by regions involved in primary stimulus categorization in the PFC is regarded as crucial in paradigms aiming to tap retrospective metacognition, as the integrity of this information determines the accuracy of the metacognitive evaluation (Fleming & Dolan, 2012). It has been suggested that ventromedial (Gherman & Philiastides, 2018) and/or dorsolateral (Shekhar & Rahnev, 2018) aspects within the PFC read out the strength of the sensory evidence and relay it to the most frontopolar PFC subregion, the aPFC. This structure is strongly associated with the formation of self-reflective metacognitive judgments beyond performance-related processing (Miele *et al.*, 2011) and is known to integrate information from various sources, e.g. areas involved in hapticospatial imagery (Kaas *et al.*, 2007). Not only can the aPFC be placed at the top of the cognitive and perceptual decision-making hierarchy (Badre & D'Esposito, 2009; Rahnev, 2017); ultimately, it may be regarded as the seat of confidence judgment generation (Shekhar & Rahnev, 2018).

When broadening the perspective provided by standard fMRI activation studies to studies assessing interindividual variations in gray matter volume (GMV), it was found that GMV in aPFC was positively associated with the accuracy of metacognitive evaluations, whereas a negative association emerged for GMV in left inferior temporal gyrus (Fleming *et al.*, 2010). Prefrontal GMV and metacognitive accuracy were also positively linked in patients with moderate or severe traumatic brain injury (Grossner *et al.*, 2018), while patients with focal prefrontal lesions were significantly less likely than controls to consciously detect their own correct responses on a visual backward masking task (Del Cul *et al.*, 2009). In contrast, it should be noted that Molenberghs *et al.* (2016) observed a negative association between aPFC activation and metacognitive accuracy on a social and cognitive reasoning task, which they argued to indicate positive affect-related biasing of metacognitive judgments towards overconfidence.

Nevertheless, a large scale coordinate-based meta-analysis by Vaccaro and Fleming (2018) further corroborated the association of frontal and prefrontal areas with metacognitive evaluations: Synthesizing findings from 47 neuroimaging studies (i.e., employing structural or functional magnetic resonance imaging with associated behavioral measurements) on metacognition, they identified a domain-general frontoparietal network anchored in posterior medial frontal cortex, dorsolateral and ventromedial PFC, precuneus, insula/inferior frontal gyrus (IFG), and ventral striatum across all metacognition-related activations. By taking into account study-specific differences such as Type I domains, the type of metacognitive judgment under investigation (judgment-related activation, confidence-related activation, and predictors of metacognitive sensitivity), as well as the temporal focus in metamemory tasks (retrospective versus prospective), they were able to identify specific activation patterns for metaperceptual and metamnestic processes as well as for parametrically confidence-level-related contrasts. The right-hemispheric anterior dorsolateral PFC emerged as specialized in metaperceptual activations; interestingly, it had previously been argued that anterior and dorsolateral aspects of the PFC are more integral to metaperception than to metamemory (Fleming *et al.*, 2014), as discussed in detail in Chapter IV.1. For metamemory, on the other hand, a specific role of the bilateral parahippocampal cortex was suggested (Martín-Luengo *et al.*, 2021; Vaccaro & Fleming, 2018). The fairly consistent involvement of bilateral insula/IFG might be attributable to a specific function of this area for error monitoring processes (Bastin *et al.*, 2017).

With respect to the debate outlined in Chapter I.1.4 on the generality versus specificity of metacognition across domains, the adoption of a neuroscientific perspective therefore suggests domain-general contributions to metacognitive judgments (Vaccaro & Fleming, 2018). However, the isolated brain functional approach is unable to provide a thorough account on its own, as understanding the biological substrates of psychological behavior requires consideration of other parameters, such as the integrity of the underlying endogenous hormone and neurotransmitter systems. The application of reversible external pharmacological interventions allows to identify potential neurophysiological substrates of metacognitive processes which are themselves based on a complex interplay of small-scale mechanisms at the neuronal and neurotransmitter level; these will be referred to as *pharmacological substrates* in the following.

I.2.2 PHARMACOLOGICAL SUBSTRATES

Whereas the mechanisms underlying metacognition at the brain functional level have been explored fairly comprehensively in recent years, there is still a scarcity of studies addressing the question of the biological basis of metacognition at the level of neurotransmitter systems. The major small-molecule neurotransmitters active in the human brain can be classified into the groups of amines, biogenic amines, indoleamines, membrane solubles, and amino acids (J. H. Schwartz, 2002), the latter group comprising γ-aminobutyric acid (*GABA*), the primary inhibitory (Lydiard, 2003), and *glutamate*, the primary excitatory (Nedergaard *et al.*, 2002) neurotransmitter of the human central nervous system (CNS). While agonists of a given neurotransmitter activate signal transduction in the corresponding cell by binding to a specific receptor, antagonists inhibit the effect of an agonist (Frohlich & Van Horn, 2014; Krall *et al.*, 2015). For the latter, a further distinction can be made between competitive antagonists, which directly compete with endogenous ligands or agonists for the binding site (Swinney, 2004), non-competitive antagonists, which bind to an allosteric binding site (Neubig *et al.*, 2003), and uncompetitive antagonists, which block the influx of ions through a channel (Frohlich & Van Horn, 2014). Figure 3 provides an illustration of the afferent and efferent projections in the medial PFC (Bittar & Labonté, 2021) of several of the major neurotransmitter systems reviewed in the following, which allows for an optimal integration with the previous chapter, as the PFC represents a core neural substrate of metacognitive confidence formation (Vaccaro & Fleming, 2018). For a more specific representation of the key neural pathways of the most relevant neurotransmitter in the context of this dissertation – glutamate – at the whole-brain level, see T. Schwartz *et al.* (2012).

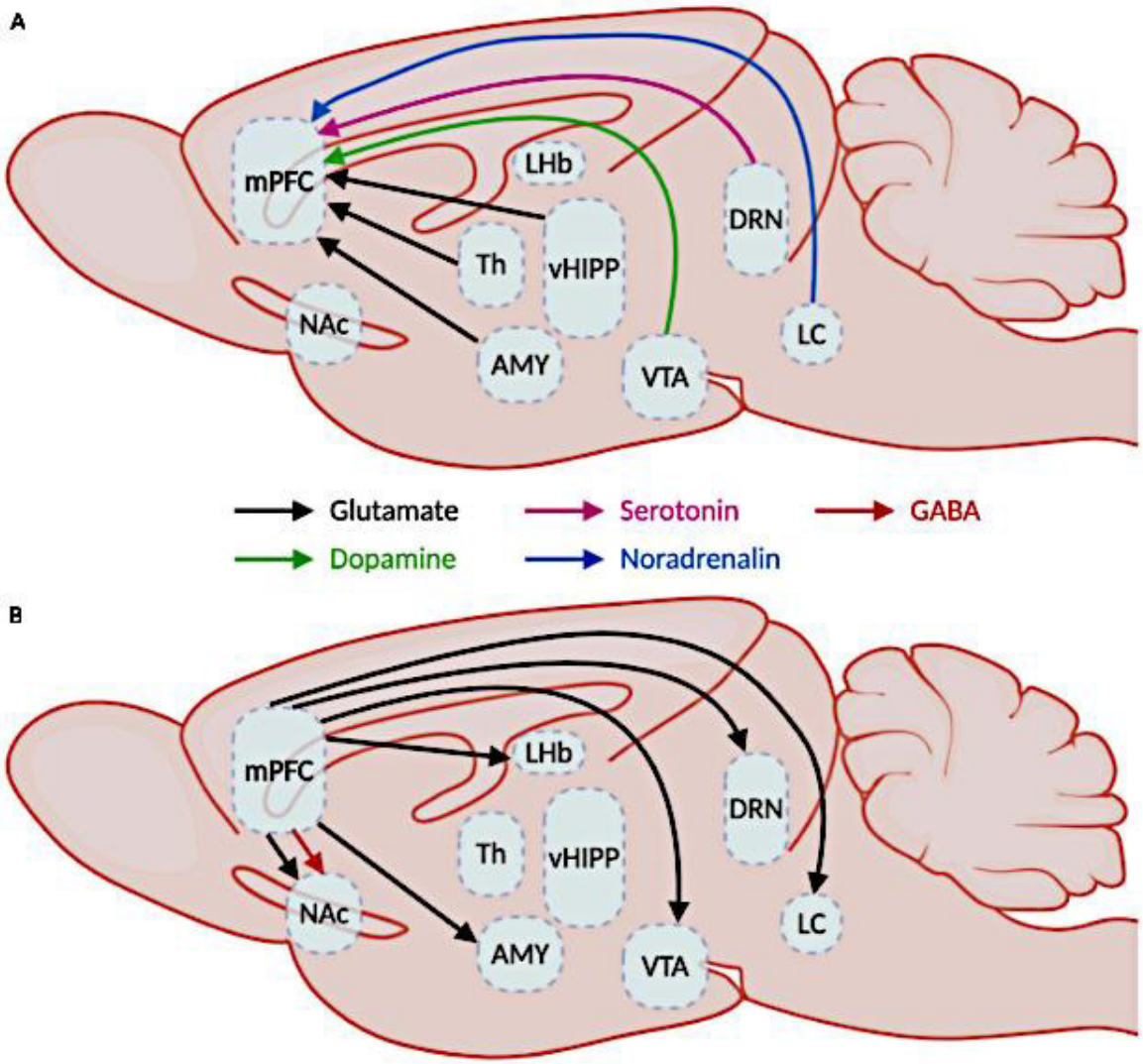

Figure 3. Afferent (A) and efferent (B) projections of major neurotransmitter systems in the medial prefrontal cortex (mPFC; schematic representation). Abbreviations: AMY, amygdala; DRN; dorsal raphe nucleus; LC, locus caeruleus; LHb; lateral habenula; NAc, Nucleus accumbens; Th, Thalamus; vHIPP; ventral hippocampus; VTA, ventral tegmental area. Republished from "Functional Contribution of the Medial Prefrontal Circuitry in Major Depressive Disorder and Stress-Induced Depressive-Like Behaviors" by T. Bittar and B. Labonté, 2021, *Frontiers in Behavioral Neuroscience, 15*(699592), p. 4; in accordance with Frontiers Copyright Policy, permission is granted under the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0).

Associations of the indoleamine molecule *serotonin* with metacognitive functioning have mainly been obtained using offline measures of metacognition: Multi-day application of selective serotonin reuptake inhibitors, which induce an elevation of synaptic serotonin levels (Artigas *et al.*, 2002), was shown to lead to a reduction in pathological aspects of self-reported metacognition in OCD patients (Besiroglu *et al.*, 2011; Park *et al.*, 2020). Furthermore, the 5-HTTLPR genetic polymorphism, which contains a regulatory variation determining the transcriptional activity of the SLC6A4 gene and thereby indirectly the serotonin concentration in the synaptic cleft (Canli & Lesch, 2007), was associated with self-reported variations in everyday monitoring functionality of healthy individuals (Alfimova *et al.*, 2017).

In contrast, the biogenic amine *dopamine* has been investigated in relation to online measures of metacognition in several studies. There is evidence for an association of dopamine with metacognitive bias: Dopaminergic activation via the agonist pergolide was shown to increase subjective confidence in seeing rapidly presented words (Lou *et al.*, 2011). Andreou *et al.* (2014) found a significantly reduced number of high-confidence incorrect responses in a visual memory task under the dopaminergic D2 receptor antagonist haloperidol, compared to both a dopaminergic agonist and placebo. Moreover, dopaminergic stimulation via oral administration of the L-dopa agonist was associated with selective improvement in self- and task-related components of metacognition (Joensson *et al.*, 2015). Haloperidol administration was also shown to lead to impaired metacognitive efficiency for 'New' responses in an episodic memory paradigm (Clos *et al.*, 2019a); however, as the blockade of presynaptic D2 autoreceptors via haloperidol is thought to induce a potentiation of phasic dopamine release (Y.-C. I. Chen *et al.*, 2005; Clos *et al.*, 2019b; Ford, 2014; Garris *et al.*, 2003), these findings cannot be considered as clear evidence for a constituent role of the dopaminergic system in metamnestic processes (Clos *et al.*, 2019a). After all, Hauser *et al.* (2017) found no effect of the D2/D3 receptor antagonist amisulpride on metacognitive performance in a perceptual decision-making task. Ultimately, the heterogeneity of these findings may point to potential domain-specific pharmacological mechanisms underlying perceptual and mnestic metacognition (Morales *et al.*, 2018).

Importantly, Hauser *et al.* (2017) also examined the association of the biogenic amine *noradrenaline* with task-based metacognition. They observed a selective increase in metacognitive performance following noradrenergic blockade via propranolol administration. This finding, which is remarkable in that it points to neuroenhancement by means of a pharmacological modulation, was attributed to the influence of uncertainty and arousal: Since the detection of errors triggers an orienting response mediated via a phasic noradrenergic burst, it results in a local reset and erasure of currently maintained information (Dayan & Yu, 2006; Hauser *et al.*, 2017). Experimental blockade of noradrenaline could prevent this reset of accumulated perceptual information and thus provide more complete information to the metacognitive process.

The corticosteroid *cortisol* has also been investigated in relation to the accuracy of online metacognitive processes. Induction of a physiological stress response via external application of hydrocortisone was shown to result in selective deterioration of metacognitive efficiency during a perceptual discrimination task (Reyes *et al.*, 2020), in line with the association of elevated cortisol levels with impaired metacognitive ability (Reyes *et al.*, 2015). This is attributable both to direct downregulatory effects of cortisol on prefrontal areas as well as interactions with noradrenergic activity. However, it is unclear to what extent this effect hinges on the type of sensory modality or the experimental domain in which metacognition is assessed (Reyes *et al.*, 2020).

An overall heterogeneous picture arises regarding GABAergic modulations of metacognitive operations. While some studies found a significant deleterious effect of lorazepam, a positive allosteric modulator at the $GABA_A$ receptor (Faßbender *et al.*, 2021), on the accuracy of online metamnestic judgments (Bacon *et al.*, 1998; Mintzer & Griffiths, 2003), other studies failed to report such an association (Izaute & Bacon, 2006; Massin-Krauss *et al.*, 2002). The GABAergic agonist triazolam (Carter, Kleykamp, *et al.*, 2013; Weingartner *et al.*, 1993) and the non-selective *acetylcholine* receptor antagonist (Wohleb *et al.*, 2016) scopolamine have also been reported to significantly impair an individual's ability to monitor the accuracy of their decisions (Mintzer *et al.*, 2010; Mintzer & Griffiths, 2003). Finally, the main psychoactive constituent of cannabis, Δ9-tetrahydrocannabinol, which targets receptors of the endocannabinoid system, was likewise associated with decreased metacognitive insight (Adam *et al.*, 2020)

As glutamate accumulation is closely linked to a regulatory mechanism within cognitive control circuits (Wiehler *et al.*, 2022), it appears sensible to investigate the role of glutamate with respect to the related function of metacognition. Intramuscular administration of different doses of ketamine, an uncompetitive antagonist at the glutamatergic *N*-methyl-D-aspartate (NMDA) receptor with known dissociative properties (Gitlin *et al.*, 2020), had no effect on metamemory in two studies (Carter, Kleykamp, *et al.*, 2013; Lofwall *et al.*, 2006). It is noteworthy that Lofwall *et al.* (2006) did not observe any alterations in conscious experiences under ketamine, raising the possibility that a potential influence of glutamate remained obscured due to insufficient dosage and/or the intramuscular method of application. More importantly, however, as in most studies published before the introduction of the meta-d' framework (Maniscalco & Lau, 2012; see Chapter II.1.1), both studies quantified metacognitive performance in a suboptimal manner: Whereas the gamma correlation coefficient offers the advantage that – unlike SDT measures – it does not rely on distributional assumptions (Nelson, 1984, 1986), it is unable to provide sufficient control over the confounding influence of response bias the Type I and Type II level and of Type I sensitivity on metacognitive performance (Masson & Rotello, 2009). It therefore appears plausible that due to a combination of methodological conditions, the relationship between glutamate and metacognition has not yet been adequately delineated. Finally, Carter, Reissig *et al.* (2013) described significant, yet unspecific effects of dextromethorphan (DXM) on metacognitive performance. Although DXM acts primarily at the Sigma-1 receptor, it also shows affinity for the NMDA receptor (C. Brown *et al.*, 2004) as an uncompetitive antagonist (Burns & Boyer, 2013) and evokes effects similar to those of dissociative anesthetics such as ketamine (Stahl, 2013), which may provide a preliminary indication for NMDA-associated modulations of metacognition.

I.3. ACUTE KETAMINE ADMINISTRATION

I.3.1 THE GLUTAMATERGIC NMDA RECEPTOR COMPLEX

The neurotransmitter glutamate accounts for the overwhelming majority of cortico-cortical neurotransmission and approximately 60% of total brain neurons (Kantrowitz & Javitt, 2010). When glutamate is released into the synaptic cleft as a result of neuronal stimulation (Jourdain *et al.*, 2007; Sanchez-Prieto *et al.*, 1996), it binds to different types of postsynaptic glutamate receptors. These can be classified into metabotropic (Conn & Pin, 1997), G-protein-coupled (Olivares *et al.*, 2012), and ionotropic receptors (Dingledine *et al.*, 1999). NMDA receptors belong to the group of ionotropic glutamate receptors with intrinsic cation-permeable and ligand-gated ion channels, which furthermore includes α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and kainate (Asztély & Gustafsson, 1996; Meldrum, 2000). While the primary agonist of NMDA receptors is glutamate, the hydrophilic amino acid glycine also binds to a specific glycine binding site (Cummings & Popescu, 2015; Yu & Lau, 2018). The requirement of a co-agonist for channel pore opening in tandem with the primary agonist constitutes a unique property of the NMDA receptor (NMDAR) compared to other ligand-gated ion channels, and the mechanism of co-agonist action is only incompletely understood (Laube *et al.*, 1997). However, micromolar background levels of glycine are already sufficient for channel activation (Frohlich & Van Horn, 2014). Subsequent to activation of the receptor by simultaneous depolarization and binding of glutamate and glycine (Mayer *et al.*, 1984), the magnesium (Mg^{2+}) ion blocking the NMDAR channel at resting membrane potential is removed, and the channel pore becomes permeable, in particular to calcium (Ca^{2+}) ions (Dingledine *et al.*, 1999; Foster & Fagg, 1987; Johnson & Kotermanski, 2006); see Figure 4 for an illustration of regular NMDAR transmission and NMDAR inhibition by the uncompetitive NMDAR antagonist ketamine.

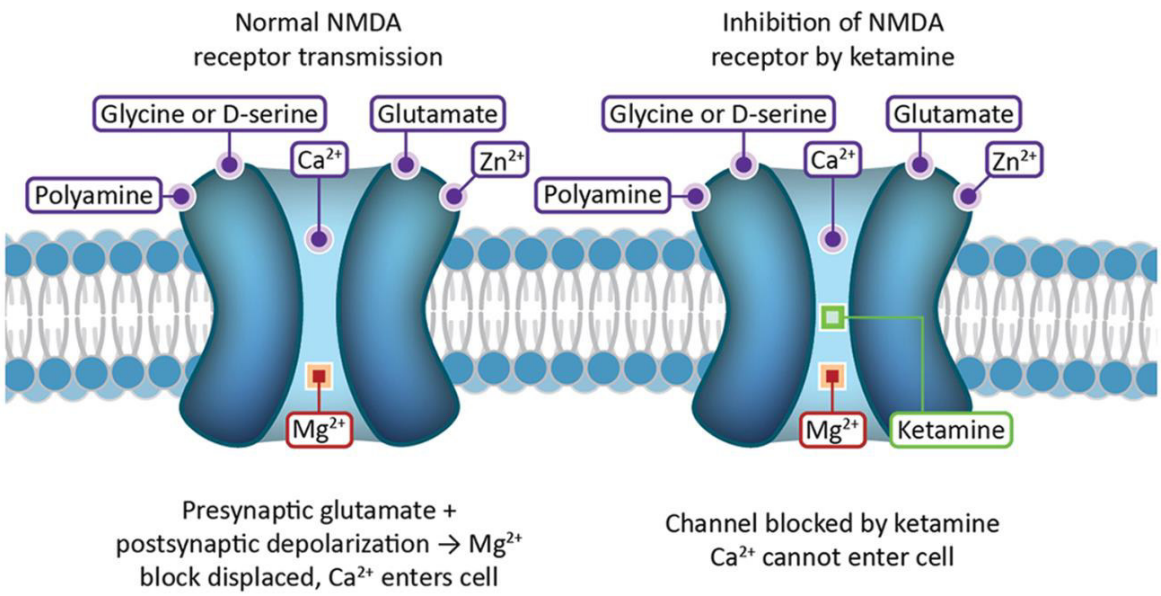

Figure 4. Ketamine as an uncompetitive antagonist at the NMDAR. Left panel: Regular NMDAR transmission in which the NMDAR channel becomes permeable to Ca^{2+} ions after presynaptic glutamate binding and postsynaptic depolarization. Right panel: Ketamine blocks the NMDAR channel at a phencyclidine (PCP) binding site and prevents influx of Ca^{2+} ions. Republished from "Reviewing the ketamine model for schizophrenia" by J. Frohlich and J.D. Van Horn, 2014, *Journal of Psychopharmacology, 28*(4), p. 289; copyright (2014), with permission from SAGE Publishing.

An early focus in psychopharmacological research on the NMDAR was directed at its role in learning and memory through long-term potentiation in the postsynaptic membrane (Collingridge, 1987), with synaptic plasticity promoted by positive feedback loops within a complex biochemical homeostasis (C. M. Norris *et al.*, 2006; Platel *et al.*, 2007). In this vein, the NMDAR has been suggested as a promising target for cognitive enhancement (Collingridge *et al.*, 2013). NMDA receptors are composed of two GluN1 and two GluN2 subunits (Tovar *et al.*, 2013) encompassing four domains (Mesbahi-Vasey *et al.*, 2017), with the transmembrane domain being of highest relevance for the action of uncompetitive NMDAR antagonists (Ogden & Traynelis, 2011; Wilding *et al.*, 2016; Zorumski *et al.*, 2016), which do not act directly at the glutamate binding site, but inside activated ion channels (H.-S. V. Chen & Lipton, 2006; Robinson & Keating, 2006). Since over-release of glutamate and/or overstimulation of NMDA receptors can lead to neurotoxic processes at postsynaptic neurons, NMDAR antagonists are of therapeutic interest both in the context of acute brain lesions (Fujikawa, 2015; A. Lau & Tymianski, 2010; Schauwecker, 2010) and in various neurodegenerative diseases (Doss *et al.*, 2014; Hardingham & Bading, 2010; Olivares *et al.*, 2012). Laboratory-synthesized uncompetitive NMDAR antagonists with high CNS permeability (Tyler *et al.*, 2017) such as phencyclidine (PCP) and ketamine were primarily optimized for their anesthetic effects in which respiration and heart rate remain largely unaffected, making it the agent of choice in patients with unstable hemodynamics (Sinner & Graf, 2008). However,

due to their dose-dependent dissociative and hallucinogenic effects, these compounds also came to be used as recreational drugs (Lodge & Mercier, 2015). Short-term effects include hallucinations, feelings of distorted reality, a sense of unity with the external environment, and out-of-body sensations sometimes referred to as the "K-hole" (Muetzelfeldt *et al.*, 2008). Long-term effects of recreational ketamine use are mostly unknown (Tyler *et al.*, 2017), but may include deficiencies in executive functions and working memory due to altered prefrontal neurotransmission (Narendran *et al.*, 2005). In addition to its primary role as an uncompetitive antagonist, an allosteric mechanism is also known to contribute to ketamine blockade of the NMDAR (Orser *et al.*, 1997). The antagonistic potency of ketamine at the NMDAR is further mediated by whether it is present in its racemic form, i.e., as a 1:1 mixture of both isomers (Moss, 1996), or as an (*S*)-enantiomer, the latter of which is assumed to lead to an approximately four times stronger blockade of NMDA receptors, thus maximizing the effects of the compound (Peltoniemi *et al.*, 2016). A recent study by Passie *et al.* (2021) suggested, however, that the (*S*)-enantiomer does not produce stronger psychotomimetic or dissociative effects than racemic ketamine, which was associated with lower negative psychopathology and stronger antidepressant effects than (*S*)-ketamine.

I.3.2 KETAMINE EFFECTS ON COGNITION AND CONSCIOUSNESS

A variety of studies has demonstrated pronounced debilitating effects of ketamine administration on cognitive functions, such as episodic (Malhotra *et al.*, 1996; Morgan *et al.*, 2004) and working memory (Ma *et al.*, 2018), selective attention (Oranje *et al.*, 2000), sense of agency (Moore *et al.*, 2011), temporal (Coull *et al.*, 2011) and sensory perception (Øye *et al.*, 1992), perceptual feature integration (Meuwese *et al.*, 2013), and cognitive flexibility (Krystal *et al.*, 1994). Furthermore, ketamine is considered an adequate and well-established pharmacological model of schizophrenia (Adell *et al.*, 2012) with a good safety record in experimental and clinical settings (Schmechtig *et al.*, 2013; Wolff & Winstock, 2006), as it temporarily and reversibly induces some of the cognitive and emotional deficits as well as key elements of the positive (e.g. delusions, sensory alterations) and negative (e.g. emotional and social withdrawal) symptomatology of schizophrenia in healthy individuals (Javitt, 2007; Krystal *et al.*, 1994; Lahti *et al.*, 2001). Acute ketamine administration also has deleterious effects on motor, in particular oculomotor, functions, including markedly increased antisaccade error rates and latencies (Condy *et al.*, 2005) as well as robust deficits in smooth pursuit eye movements (Steffens *et al.*, 2016; Weiler *et al.*, 2000). A few findings, however, suggest cognitive enhancement via ketamine, such as improved visual search in rhesus monkeys (Shen *et al.*, 2010) or improvement of visual and working memory in individuals with treatment-resistant depression (Y. Lee *et al.*, 2016; Zhang & Ho, 2016).

The latter finding could be argued to represent a manifestation of dose-dependent antidepressant ketamine effects at the cognitive level; after all, the rapid and sustained antidepressant property of ketamine has become a primary focus of research over the course of the last years (Abdallah *et al.*, 2018; Krystal *et al.*, 2013; Monteggia & Zarate, 2015; Murrough *et al.*, 2013). Current consensus holds that NMDAR blockade is unlikely to be at the heart of these effects, but rather the concomitant increase in AMPA receptor density and/or function, which activates downstream signaling pathways that ultimately restore synaptic strength in PFC and hippocampus (Akinfiresoye & Tizabi, 2013; Aleksandrova *et al.*, 2017; Koike *et al.*, 2011; Zhou *et al.*, 2014). Preliminary evidence also suggests an abstinence-facilitating effect of ketamine in the treatment of substance use disorders (Jones *et al.*, 2018). Whereas the immediate analgesic effects of ketamine appear to be mediated predominantly by a combination of opioid system sensitization and aminergic anti-nociception, its hypnotic effects may be caused by a combination of immediate channel blockade of NMDA and hyperpolarization-activated cyclic nucleotide-gated ion channels involved in synaptic integration and learning or memory processes (Nolan *et al.*, 2007; Sleigh *et al.*, 2014). Eventually, decreased frontoparietal top-down connectivity at anesthetic doses of ketamine could express a loss of self-awareness under ketamine during which individuals have vivid, dreamlike experiences (Baird *et al.*, 2019; U. Lee *et al.*, 2013).

Another feature of ketamine that has spawned scientific interest pertains to the compound's psychotomimetic effects, which are of particular relevance to the study of human consciousness (Nutt, 2014; Shushruth, 2013). In addition to the cognitive and perceptual distortions and impairments outlined above, the psychoactive effects of ketamine include ego transcendence and feelings of disconnection from both body and environment (Vlisides *et al.*, 2018). Consequently, ketamine is known to induce a conscious state distinctly altered from the “normal” human waking state (Anis *et al.*, 1983; Passie *et al.*, 2021; Vlisides *et al.*, 2018) sometimes referred to as a psychedelic (Bowdle *et al.*, 1998), but more consistently as a dissociative state (Corssen & Domino, 1966; Muetzelfeldt *et al.*, 2008; Pallavicini *et al.*, 2019). Pursuant to the *Theory of conscious states* by Carhart-Harris *et al.* (2014), this “primary” altered state of consciousness is characterized by elevated neural entropy (i.e., high disorder/uncertainty) and is speculated to have preceded the development of normal human waking consciousness (“secondary consciousness”). The latter is associated with streamlined cognition and a highly organized neurodynamic, in which the brain operates below a point of criticality, i.e., at comparatively low entropy within an evolutionary adaptive equilibrium between order and disorder. Pharmacological induction of a psychedelic-like state of consciousness is argued to collapse the highly organized activity within synchronously active brain regions attributed to the default-mode network (Andrews-Hanna, 2012), leading to a shift towards primary consciousness with unrestricted cognition and phenomenology (Carhart-Harris *et al.*, 2014). This should be accompanied by a deficit in self-referential and self-monitoring processes: Although cognition and phenomenology may be restrained

to certain facets in normal secondary consciousness, these are optimized to ensure the integrity of processes that serve self-monitoring – and thereby metacognitive – functions (Carhart-Harris *et al.*, 2014).

And indeed, using an antisaccade task in macaque monkeys, Skoblenick and Everling (2014) described a ketamine-associated deterioration in the metacognitive process of error monitoring. This was interpreted as evidence for the hypothesis put forward by Bickel and Javitt (2009) that NMDAR dysfunction not only underlies the psychotic symptoms, but also – and particularly – the self-monitoring deficits of schizophrenia, which could be induced in a comparable fashion by ketamine in healthy volunteers within the domain of verbal monitoring capacity (Stone *et al.*, 2011). Likewise, brain activity in regions involved in self-monitoring was found to be reduced during a ketamine challenge (Ionescu *et al.*, 2018). Thus, one might plausibly assume that under methodologically optimized conditions, core metacognitive processes might ultimately be altered under ketamine, which would suggest a key role for the NMDAR in human metacognition.

I.4 GOALS OF THE PRESENT THESIS

As outlined in the previous chapters, there is still a plethora of unresolved questions regarding the structural organization of metacognition as well as its biological, in particular its pharmacological, substrates. Emerging insight into the architecture of metacognition with regard to the organization of its functional properties, for instance, suggests that the contribution of domain-general signals may have been spuriously obscured in earlier investigations (Mazancieux, Fleming, *et al.*, 2020; Rouault, McWilliams, *et al.*, 2018). Consequently, this dissertation addresses the primary question whether it is possible to identify a domain-general monitoring ability with a comprehensive set of state-of-the-art paradigms and instruments, which would imply metacognition to be a trait-like, unitary disposition, or whether there are metacognitive subsystems for different types of demands, consistent with a domain-specific account. This was investigated in Study 1 using a combination of correlational and latent variable approaches within a large laboratory sample, which also made it possible to map the coexistence of both phenomena correspondent to a unity and diversity structure. For this purpose, a total of three task domains with two paradigms each were investigated, which encompassed the domains of (visual) perception and episodic memory as well as the somewhat underrepresented attention-to-action domain to map the complexity of metacognitive processes in a thorough manner. Secondly, the structure of metacognition was investigated not only across experimental domains, but also across method-groups by relating these online measures to a set of different offline questionnaires, each focusing on different facets of metacognitive enterprises in real-world situations.

Beyond the fairly well-established neural correlates of metacognition, it is furthermore necessary to expand the scientific understanding of the neurobiological underpinnings of metacognition by delineating how metacognitive processes are underpinned at the neurotransmitter level. For reasons outlined above, the glutamate system, and more specifically glutamatergic modulations via the NMDAR, represents a particularly promising candidate to gain insight into the neuropharmacological basis of metacognition. With the existing body of knowledge on metacognitive deficits and impaired self-monitoring in schizophrenia spectrum disorders (Bickel & Javitt, 2009; Moritz *et al.*, 2005, 2006), and in agreement with the glutamate hypothesis of schizophrenia (Javitt, 2012; Kantrowitz & Javitt, 2010), such an association seems reasonable to assume. Although initial efforts were undertaken to investigate this relationship in two previous studies (Carter, Kleykamp, *et al.*, 2013; Lofwall *et al.*, 2006), it seems sensitive to revisit this inquiry under improved methodological conditions, namely the application of SDT methodology to quantify metacognitive performance; even more so as the finding of DXM-induced alterations in metacognition provides a tentative indication of modulations of metacognition that may involve the NMDAR (Carter, Reissig, *et al.*, 2013). In addition, the NMDAR antagonist ketamine is already known to impair verbal self-monitoring and the metacognitive process of error detection (Skoblenick & Everling, 2014; Stone *et al.*, 2011). Finally, it is well-established that ketamine leads to an altered state of consciousness, and since metacognition represents a function of consciousness, it could also be assumed to induce an altered state of metacognition (Carhart-Harris *et al.*, 2014).

Due to heterogeneous results of previous studies on the structure of individual differences in metacognition, but also due to partly conflicting insights at the level of pharmacological interventions (Clos *et al.*, 2019a; Hauser *et al.*, 2017), it must be assumed that metaperception and metamemory constitute relevantly distinct processes, which were therefore investigated in two separate experiments: Study 2 examined the effects of a psychotomimetic dose of ketamine on metacognition and the associated brain function in an episodic memory framework, whereas Study 3 provided a congruent implementation within the domain of perceptual decision-making. Thus, in addition to the central question of the potential involvement of the NMDAR in metacognition, the question of whether its involvement is consistent across domains is likewise of primary concern. Indeed, if a common pharmacological underpinning of metaperception and metamemory was to be identified, this could constitute a compelling piece of evidence for the domain-general hypothesis of metacognition. Since both studies involved fMRI, the question of (*i*) if and how potential pharmacological effects were mediated at the neural level and (*ii*) the identification of common and shared neural correlates of the respective paradigms were of further relevance. In this regard, both whole-brain activation patterns and measures of context-dependent functional connectivity were of interest, which could shed additional light on the interaction of the known neural substrates of metacognition (Vaccaro & Fleming, 2018) and the known ketamine effects on brain activation (Ionescu *et al.*, 2018).

II. METHODOLOGY

The investigation of the research goals outlined in Chapter I.4 comprised behavioral (task-based) and psychometric (questionnaire-based) measurements and analyses, such as those related to first-order (task) and second-order (metacognitive) performance parameters and their interrelations across tasks and domains, as well as pharmacological administration and functional magnetic resonance imaging.

II.1 BEHAVIORAL MEASUREMENT

II.1.1 QUANTIFICATION OF METACOGNITION

In experimental laboratory paradigms, metacognition is typically tapped by participants reporting their subjective confidence contingent on a previously given first-order judgment, as detailed in Chapter I.1.3. The simplest and perhaps most intuitive approach to quantifying metacognitive performance is to calculate the across-trial correlation between the accuracy of Type I and Type II judgments, where a high positive correlation would be interpreted to reflect high metacognitive performance (Fleming & Lau, 2014; Grimaldi *et al.*, 2015). Over many years, non-parametric correlation indices such as the phi (Guilford & Perry, 1951) or the gamma coefficient (Goodman & Kruskal, 1954) were routinely applied to this end, which are, however, fraught with the problem that they do not quantify metacognitive performance independently of metacognitive bias (Galvin *et al.*, 2003; Masson & Rotello, 2009). This is necessary because individuals who e.g. tend to report low confidence judgments may nonetheless be highly sensitive to differences in their Type I performance.

For this reason, SDT-based approaches provide a suitable alternative, since they intrinsically account for differences in confidence placement criteria (Barrett *et al.*, 2013). As application of SDT methodology assumes two distributions of internal responses, a Type I observer criterion may be identified, based on which all signals below the criterion are assigned to one distribution (stimulus category 1) and all signals above it are assigned to the other distribution (stimulus category 2; Grimaldi *et al.*, 2015). Analogously, it can be assumed that observers set criteria for their second-order confidence judgments as well, below and above which low or high levels of confidence are assigned to the decision in favor of the respective stimulus category (Grimaldi *et al.*, 2015; Maniscalco & Lau, 2014). A common non-parametric SDT method for quantification of metacognitive performance represents the calculation of each participant's Type II receiver operating characteristic (ROC), where the number of correct positive (Hit Rate) and incorrect positive (False Alarm Rate) responses are calculated for each criterion and plotted, ultimately yielding the ROC curve (Fleming & Lau, 2014). The higher the area under the Type II ROC (AUROC2), the higher an individual's metacognitive accuracy; see Figure 5 for an illustration.

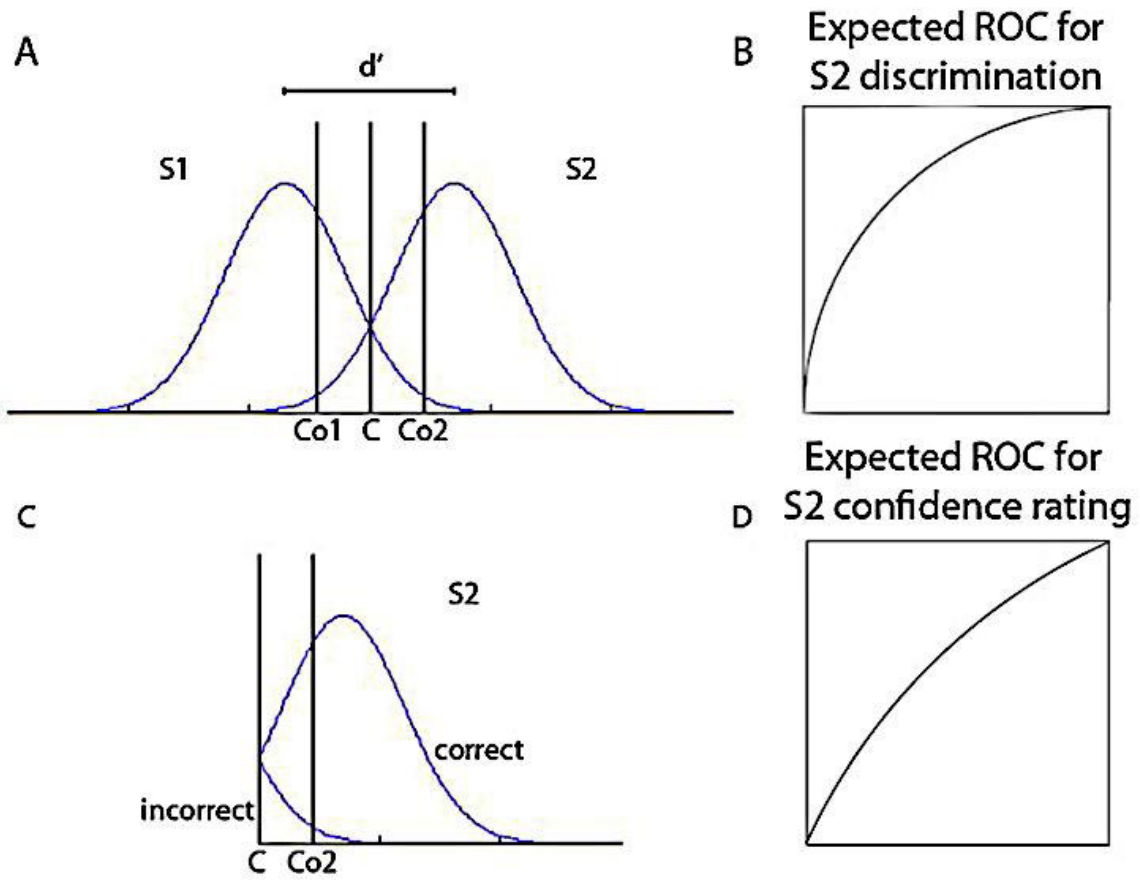

Figure 5. Application of signal detection theory to quantify performance (panels A-B) and metacognition (C-D). A.) Internal response distributions for stimulus categories S1 and S2. Sensations below Type I criterion C are considered belonging to S1, sensations exceeding C are considered belonging to S2. Co1 and Co2 express confidence placement criteria; responses lower than Co1 and higher than Co2 are endorsed with high confidence, all those in between with low confidence. B.) Type I Receiver operating characteristic (ROC) curve resulting from shifting C along the x-axis. C.) Segment for S2 response. D.) Type II ROC curve resulting from shifting Co2 along the x-axis. Republished from "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" by P. Grimaldi, H. Lau and M.A. Basso, 2015, *Neuroscience and Biobehavioral Reviews, 55*, p. 91; copyright (2015), with permission from Elsevier.

While Type II ROC analyses offer the advantage that they are not based on explicit distributional assumptions, they are nevertheless not independent from the confounding influence of Type I performance and Type I response bias on metacognitive sensitivity (Galvin *et al.*, 2003; Higham *et al.*, 2009). The meaningfulness of Type II ROC analyses is thus limited unless e.g. an adaptive staircase procedure is applied to control for differences in performance (Fleming & Lau, 2014). The meta-d' framework offers a model-based alternative that facilitates control over these confounding influences (Maniscalco & Lau, 2012). It assumes that the information available for the Type I task is typically exhaustive of the information available for the Type II task, which would be consistent with the direct-translation hypothesis (Higham *et al.*, 2009; Kepecs & Mainen, 2012). A key advantage of the method is

that the first-order (d') and second-order (meta-d') sensitivity measures are provided on the same unit: Meta-d' can be conceived as the sensory evidence available to the metacognitive process in signal-to-noise ratio units, whereas d' represents the sensory evidence available for first-order decision-making in signal-to-noise ratio units (Fleming & Lau, 2014). The meta-d' approach assumes an ideal observer model of the link between Type I and Type II SDT, according to which a metacognitively optimal or ideal observer is hypothesized to be a person rating confidence with the maximum possible metacognitive sensitivity (Fleming & Lau, 2014). At the heart of the meta-d' approach is the relationship between Type I d' and Type II meta-d': If meta-d' < d', the observer is metacognitively suboptimal, i.e., not making use of the complete information available to the Type I process. It therein becomes apparent why measures relating meta-d' to d' are labeled measures of metacognitive *efficiency*: the degree to which meta-d' is smaller than d' reflects the degree to which the observer is metacognitively inefficient (Maniscalco & Lau, 2012). On the other hand, there may also be cases of supra-optimal metacognition (meta-d' > d'), which can be explained by post-decisional models or the model of second-order computation (Fleming & Daw, 2017).

Metacognitive efficiency can be expressed e.g. as a difference measure (meta-d'–d') or more commonly as a ratio (meta-d'/d'), yielding an index of metacognitive performance adjusted for the confounding influence of Type I performance and Type I and Type II response bias on metacognition (Maniscalco & Lau, 2012, 2014) that has become the method of choice for quantifying metacognitive performance in the majority of studies following the introduction of the meta-d' framework. The original approach by Maniscalco and Lau (2012) is typically implemented using a maximum likelihood estimation of parameters, retrieving the value of a metacognitively ideal observer's d' that would produce a participant's empirically observed confidence rating data with the highest probability. The evaluation can also be conducted in a response-specific manner, which may be desirable in cases of a qualitative difference (e.g. stimulus present versus absent) between Type I stimulus categories; meta-d' is then determined separately for the Type I response categories (Maniscalco & Lau, 2014). Over recent years, some extensions or alternatives – such as the dynamic evidence accumulation framework (Desender *et al.*, 2021, 2022) – have been proposed, the most well-received being the HMeta-d method by Fleming (2017), which implements a hierarchical Bayesian approach (M. D. Lee & Wagenmakers, 2014) instead of maximum likelihood estimation. It aims to address several problems of the conventional meta-d' method, as it is argued that robust parameter estimates are achieved only at high trial counts with the conventional method, which furthermore comes with the risk of increased statistical noise, as it merely provides point estimates of meta-d', and which regularly requires the application of edge correction to avoid zero cell counts of confidence ratings (e.g. for high confidence error trials) in the calculation of sensitivity parameters (Fleming, 2017). Further advantages of the hierarchical Bayesian approach concern the fact that group-level hypotheses may be tested in a straightforward manner, as the

posterior distribution of a difference parameter between groups can be estimated directly, and that the model can be extended to determine a correlation coefficient between metacognitive efficiency scores obtained in two different tasks within the same sample (Fleming, 2017). Guggenmos (2021) also found that the test-retest reliability of efficiency parameters can be substantially increased by application of the hierarchical method. However, the Bayesian approach is associated with an increased Type I error (false positive) rate compared to the conventional meta-d' estimation (Rausch & Zehetleitner, 2022).

II.1.2 BEHAVIORAL PARADIGMS

The behavioral paradigms in which metacognitive performance is determined typically implement a two-choice response structure on a trial-by-trial basis. However, it is also possible to tap metacognitive performance in a classical error processing paradigm and quantify Type I and Type II performance using SDT methodology: In the Go/No-Go error awareness task by Hester *et al.* (2005, 2012), participants respond to different types of stimuli according to a prespecified instruction, with so-called Go stimuli indicating execution and No-Go stimuli indicating inhibition of this response (Figure 6A). Type I performance (first-order d'), i.e., a person's ability to discriminate between Go and No-Go trials, can be computed as the difference in *z*-transformed hit and false alarm rates (Gutschalk *et al.*, 2008). Importantly, a second-order d' index may be determined analogously, relating the proportion of signaled commission errors (i.e., erroneous Go responses in No-Go trials) to all errors with Type II false alarms (i.e., erroneous No-Go responses in Go trials). This represents an adaptation of one of the early model-free approaches by Kunimoto *et al.* (2001) to quantify metacognitive sensitivity in an SDT framework. The ratio to which second-order d' is smaller or larger than first-order d' reflects an index of metacognitive efficiency in the manner of the meta-d' framework, although such indices are unable to fully account for the influence of Type I performance and Type I or Type II response bias on metacognitive accuracy (Evans & Azzopardi, 2007; Fleming & Lau, 2014).

The saccadic eye movement system provides a model for the brain's ability to flexibly control behavior, e.g. to automatically respond to a stimulus in one situation and to suppress this automatic response in favor of an alternative response in another situation (Munoz & Everling, 2004; Preciado & Theeuwes, 2018). Similar to the error awareness task, the antisaccade task taps the ability for behavioral inhibition in the domain of attention-to-action. Following the presentation of a peripheral cue stimulus, participants are instructed to direct their gaze in the opposite direction to its mirrored position (Figure 6B), which requires the suppression of the bottom-up automatic response to look at the stimulus (prosaccade) and the initiation of a voluntary motor command in the opposite direction (Munoz & Everling, 2004).

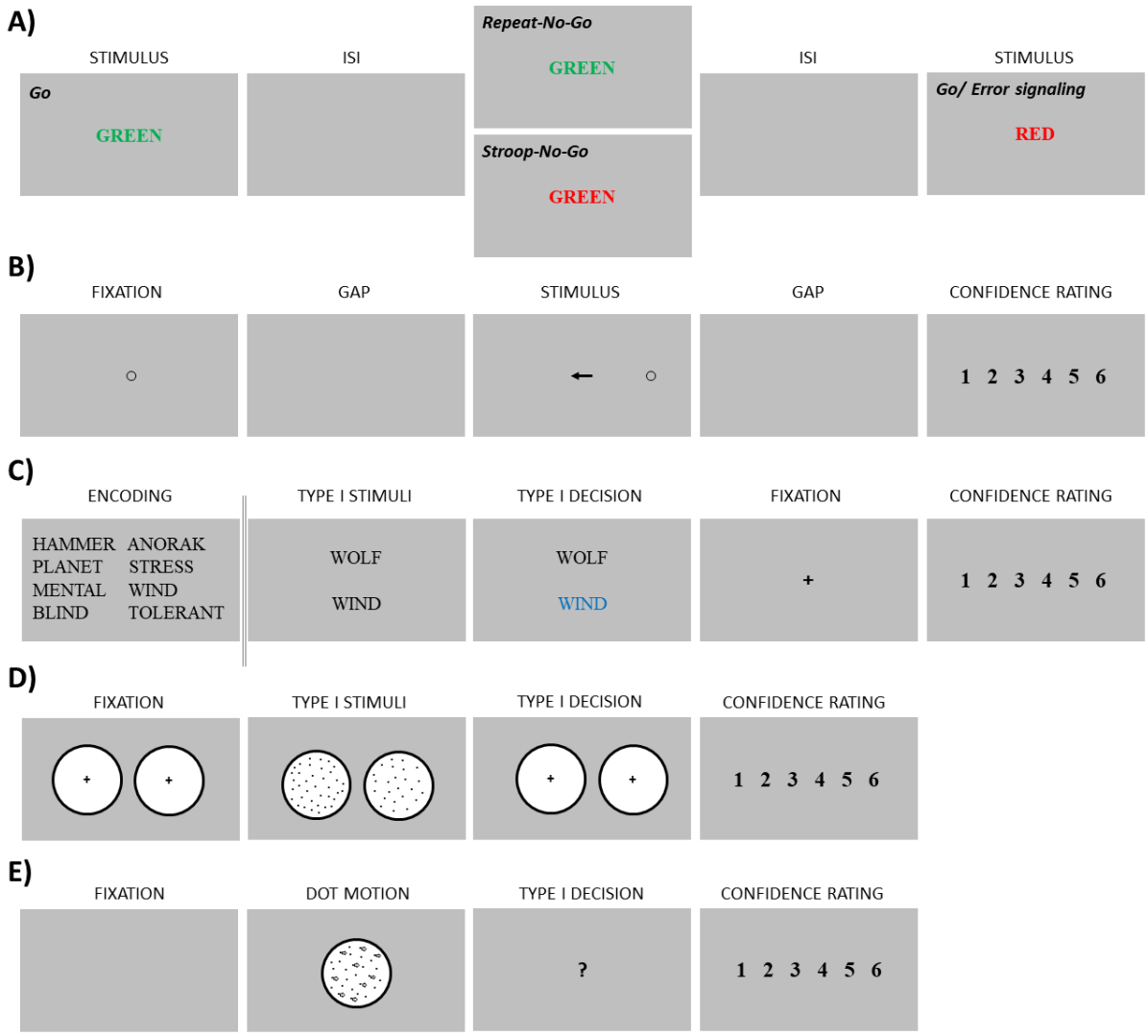

Figure 6. Prototypical trial sequences of behavioral paradigms. A) Go/No-Go error awareness task. Upon detection of a commission error, participants are instructed to indicate their error awareness by predefined button presses. ISI, interstimulus interval. B) Antisaccade task. Participants are instructed to direct their gaze in the opposite direction of an appearing cue stimulus. C) 2AFC memory task. After an encoding period, participants are instructed to designate the stimulus previously contained in the encoding list. D) Perception task with static stimuli (varying number of dots in two circles). E) Perception task with moving stimuli, in varying degrees of coherence to the left or to the right. Type I decisions are followed by metacognitive confidence ratings in Panels B-E.

Eye movements may be recorded using video-based combined pupil and corneal reflection eye trackers, which consist of a high-speed infrared light-sensitive camera and an integrated light source projecting infrared light into the eye, which is reflected off the cornea. Given a thorough pre-experiment calibration and validation procedure to map a participant's eye gaze to screen coordinates, the camera identifies the position of the eye based on the relation of the pupil's center and the corneal reflection and detects positional changes (Duchowski, 2007). Eye gaze is recorded throughout the experiment and

may thus be used to determine whether the antisaccade was conducted correctly on each individual trial. Typically, the first eye movement after cue onset is analyzed, provided that certain criteria are met, such as a minimum or maximum latency, minimum peak velocity, and maximum duration. Any eye movement towards the cue is considered a directional error and hence an erroneous response, regardless of any subsequent correction in the opposite direction (Munoz & Everling, 2004; Parton *et al.*, 2007).

In the domain of episodic memory (see Chapter I.1.4), the collection of retrospective metacognition judgments in recognition paradigms requires prior encoding of the stimuli within the experimental context. In some contexts, it may be desirable to maintain a certain delay interval (e.g. 60-90 min) between the end of the encoding and the beginning of the recognition task to allow for consolidation processes (Honey, Honey, O'Loughlin, *et al.*, 2005). Incorporating manipulations of the strength of the memory trace during encoding is not uncommon, which may exhibit selective effects on first-order retrieval and second-order metacognitive judgments (Busey *et al.*, 2000). Such manipulations of the depth of processing can be implemented e.g. by actively encoding the stimuli while performing a particular task that emphasizes either semantically rich or superficial aspects of the stimulus, resulting in deep or shallow processing (Craik & Lockhart, 1972). Stimuli from the encoding phase are then to be recognized as "old" in the retrieval phase, which also requires a set of "new" stimuli (sometimes referred to as *lures*) that were not presented to the participant in the encoding phase, but whose features (e.g. frequency, valence, vividness) should not differ from the encoded stimuli (Eisenacher & Zink, 2017; Riegel *et al.*, 2022). Recognition testing can be conducted by trial-by-trial presentations of either an old or a new stimulus (Yes/No paradigm), or by presenting both an old and a new stimulus on each trial and instructing participants to indicate whether e.g. the top or the bottom stimulus had been part of the encoding phase (2AFC paradigm, Figure 6C).

In the perceptual domain, it is customary to apply so-called staircase procedures, which adapt the difficulty of a given task to the participants' ability, thereby maintaining average performance of participants at a fixed, constant level (Macmillan & Creelman, 2005; Wright *et al.*, 2012). This can be achieved using the sequential up-and-down technique first described by Dixon and Mood (1948), whereby the stimulus intensity (i.e., the difficulty level) of a given task (e.g. the difference in dot number between two circles or the percentage of coherently moving dots, Figure 6, Panels D-E) is varied in steps of a constant size. Typically, data collection is preceded by a practice block to establish a near-threshold baseline level for each participant (Fleming *et al.*, 2016). Following a correct response, the difficulty level is increased by one step for the subsequent trial, whereas an incorrect response leads to a decrease in task difficulty (Wetherill & Levitt, 1965). Performance level may also be predetermined by requiring e.g. a certain number of consecutive correct responses to increase the difficulty level (Levitt, 1971); for instance, 2-down-1-up procedures with an expected value of ~71% correct (Weil *et al.*, 2013) or 3-down-

1-up procedures with a ~79% correct detection threshold (Kanai *et al.*, 2010) are both commonly employed.

Adaptive staircase procedures offer specific advantages for the post-experimental evaluation of metacognitive performance, as adjustment of the difficulty level achieves uniformity in task performance between participants, which helps to account for the confounding influence of Type I sensitivity on metacognitive performance (Allen *et al.*, 2017; Fleming *et al.*, 2010). However, as demonstrated by Rahnev and Fleming (2019), the mixing of high and low contrast stimuli may lead to inflated estimates of metacognitive performance across a variety of measures, including sensitivity parameters obtained from the meta-d' framework. Such staircase-related stimulus variability may be especially problematic in studies employing large step sizes between different stimulus intensities and should be taken into account in the experimental design or analyses, either by evaluating a participant's metacognitive performance at a single (e.g. the most frequent) difficulty level, or by entering stimulus variability as a covariate in analyses aiming to examine the link between metacognitive ability and a certain quantity (Rahnev & Fleming, 2019).

II.2 PSYCHOMETRIC QUESTIONNAIRE MEASUREMENT

The collection of self-reports via questionnaires or interviews represents a different methodological framework that allows to measure generalized aspects of metacognitive functioning, as described in Chapter I.1.3. Questionnaire measures are attractive to researchers for several reasons, e.g. as collected data are typically processed in a straightforward manner, requiring a simple aggregation of item scores to the corresponding scale scores. In some cases, self-ratings are supplemented by an informant rating, e.g. with the CFQ (Broadbent *et al.*, 1982), albeit suitability of this instrument for investigating metacognition is questionable, as it is mainly targeted at deficits in object-level processes. While experimental psychology lays a stronger focus on task-based measures, the following self-report questionnaires of metacognition with overall good or satisfactory psychometric criteria are common especially in applied educational or clinical contexts.

For educational purposes, metacognition is typically conceptualized rather narrowly as the ability to understand, reflect on and control one's own learning processes (Schraw & Dennison, 1994), for which offline measures are particularly well-suited due to their superior test economy compared to online measures (Veenman, 2011). This e.g. concerns the MAI (Schraw & Dennison, 1994), one of the three questionnaires employed in Study 1. Its main scales assess two major subcomponents of metacognition in the tradition of Flavell (1987), namely Knowledge about Cognition (e.g. about the self as a learner and the use of strategies) and Regulation of Cognition (i.e., processes which facilitate the control of learning).

The other two questionnaires used in Study 1 were developed with clinical applications in mind: First, the MCQ-30 (Wells & Cartwright-Hatton, 2004), which taps maladaptive expressions of self-monitoring tendencies and metacognitive beliefs and thus focuses on the specific role of metacognitive processes in the development and maintenance of psychological disorders (Nordahl *et al.*, 2022). Its five scales (Cognitive Self-Consciousness, Need to Control Thoughts, Lack of Cognitive Confidence, Positive Beliefs about Worry, Negative Beliefs about Uncontrollability/Danger) comprise themes of responsibility, punishment and superstition (Cartwright-Hatton & Wells, 1997). Second, the two scales of the BCIS (Beck *et al.*, 2004) measure willingness to question one's own perceptions and general cognitive insight in psychiatric as well as non-psychiatric samples (Fleming, Huijgen, *et al.*, 2012; Riggs *et al.*, 2012). The sum score of the Self-Certainty scale is subtracted from the Self-Reflectiveness scale to form a composite score (Martin *et al.*, 2010).

There are also several notable offline measures of metacognition which were not included in Studies 1-3. For instance, the Self-Reflection and Insight Scale (Grant *et al.*, 2002) represents a measure of insight in the context of self-regulatory and introspective processes, which distinguishes between the engagement in and need for self-reflection and a clear and confident understanding of one's own mental processes. Further structured psychometric assessments primarily used in clinical applications include the Metacognition Assessment Scale (Semerari *et al.*, 2003, 2007) in which metacognitive abilities are evaluated by an expert rating on free narratives of patients' personally relevant episodes and relationships (Bröcker *et al.*, 2017) and its self-report derivate, the Metacognition Self-Assessment Scale (Pedone *et al.*, 2017).

Finally, the Altered State of Consciousness (5D-ASC) rating scale (Dittrich, 1998) is a self-report measure not intended to tap metacognitive functioning offline, but to evaluate the induction of an altered state of consciousness in the course of a pharmacological intervention (Pokorny *et al.*, 2016), as relevant for Studies 2 and 3. Its five scales Oceanic Boundlessness, Dread of Ego Dissolution, Visionary Restructuralization, Auditory Alterations and Vigilance Reduction are conceptualized as independent key dimensions of the phenomenology characteristic for altered states of consciousness (Schmidt & Berkemeyer, 2018).

II.3 CONFIRMATORY FACTOR ANALYSIS

Confirmatory factor analysis (CFA) is a type of structural equation modelling (SEM), i.e., a method that deals with the representation of an observable or theoretical phenomenon by means of a model, in which different aspects of that phenomenon are brought into relation with each other (Kline, 2015). In particular, CFA is concerned with the measurement model part of SEM, which focuses on the

relationships between observed measures or indicators (e.g. metacognitive efficiency scores obtained in a given paradigm) and latent variables or factors, which are assumed to account for the covariations among the indicators as a common underlying cause (T. A. Brown & Moore, 2012). Associations between the indicators are therefore ascribed to a small number of (unobserved) underlying latent factors. If the observed variables are assigned to the factors *a priori*, the variation of factor interrelations within theoretically derived models allows to test the representativeness of these model assumptions for a set of empirically observed data (Mueller & Hancock, 2001). A minimum of two indicators per latent factor is required to draw meaningful conclusions about the structure of the investigated construct (Kenny *et al.*, 1998).

CFA can be used for a variety of purposes, such as psychometric evaluation, construct validation, or the detection of method effects (T. A. Brown & Moore, 2012); furthermore, Miyake *et al.* (2000) applied CFA to investigate the structure of a specific type of higher-order cognition, namely executive functions. This was achieved by the comparison of unique model specifications: In one model, a single factor was extracted from all indicators, whereas another model posited three independent factors without any shared variance, and a third model identified three separate factors that were allowed to be correlated with one another. In addition, three nested "Bi-factor models" were tested that assume the singularity of one and the unity of two factors. Accordingly, it was varied between models whether the inter-factor covariances were fixed to 0, fixed to 1.0, or freely estimated. Subsequent model tests were used to determine the adequacy of these models for the empirically observed data, eventually allowing to draw inferences about the structure of executive functions (Miyake *et al.*, 2000).

CFA is typically applied using maximum likelihood estimation procedures, for which several conditions must be met (Enders & Bandalos, 2001; Kline, 2015; Li, 2016): (a) linear latent variable interrelations; (b) no multicollinearity; (c) no violation of the normality assumption for all observed variables. The maximum likelihood estimation is used to determine those values for the factor loadings and error variances in each model specification that reproduce the empirically observed covariance matrix as closely as possible (Schermelleh-Engel *et al.*, 2003). Several indices are considered for the evaluation of the goodness-of-fit for each model, such as the χ^2-model test, which quantifies the deviation between the estimated and the empirical covariance matrix. Further indices adjusted for sample size include the Root Mean Square Error of Approximation (RMSEA), which analyzes the overall discrepancy between the hypothesized model and the population covariance matrix, the Standardized Root Mean Square Residual (SRMR), which determines the size of residual covariances within the empirical covariance matrix not accounted for by model estimates, and the Comparative Fit Index (CFI), which analyzes the discrepancy between the hypothesized model and the empirical data (Gatignon, 2010; Kline, 2015). For each of these fit indices, specific cut-off values are routinely applied to assess the goodness-of-fit of a model (Karr *et al.*, 2018).

Moreover, it is possible to conduct direct model comparisons, although it is important to note that depending on how many parameters are freely estimated or how many degrees of freedom are retained for estimation, some of the specified models are more restrictive than others, which means that they must be expected to yield a lower absolute model fit. To account for this, χ^2-difference tests can be performed to compare models at different levels of restriction; only if the test is significant, it can be assumed that the less restrictive model provides a superior explanation for the empirically observed data than the more restrictive model (Schermelleh-Engel *et al.*, 2003). In addition, Akaike (AIC) and Bayesian (BIC) information criteria indicate the adequacy of each model given its level of restriction and can be used to compare the goodness-of-fit of different models (Sen & Bradshaw, 2017).

II.4 PHARMACOLOGICAL ADMINISTRATION

In psychopharmacological experiments, the compound of interest can be administered in several ways, such as orally (Irwin & Iglewicz, 2010), intranasally (Kapetaniou *et al.*, 2021) or via intravenous, intramuscular, or subcutaneous injection routes (Jin *et al.*, 2015). Intravenous administration was found to be the overall preferable choice for ketamine (Jin *et al.*, 2015). When delivered via an intravenous access, a targeted plasma level may be rapidly achieved and maintained by a combined bolus and continuous infusion using a computerized infusion pump, injecting a precisely fixed amount of the compound of interest (e.g. ketamine) in saline solution (Berman *et al.*, 2000).

When assessing the effect of a pharmacological compound, e.g. with regard to treatment efficacy in the context of clinical psychology or psychiatry, so-called randomized controlled trials are widely considered the gold standard (Kaptchuk, 2001; Misra, 2012), as they represent the most rigorous way of determining whether a cause-effect relation exists between treatment and outcome (Sibbald & Roland, 1998). This is based in particular on the minimization of systematic error by various components, without which the overestimation of treatment effects is likely (Schulz *et al.*, 1995). *Randomization* refers to the random allocation of participants to experimental groups, which aims to eliminate both deliberate and unconscious human influence (Kaptchuk, 2001). *Double-blind* procedures require naiveté of both experimenters and participants as to whether the treatment or an inert substitute (typically a placebo) is administered. *Placebo-controlled* study designs involve comparison of drug effects with an appropriate substitute for the treatment that has no known activity expected to affect the outcome (Misra, 2012). Some researchers choose to apply active placebos, i.e., control interventions that mimic specific side effects of the experimental intervention to reduce the risk of unintentional unmasking (Jensen *et al.*, 2017), as double-blind protocols are not always feasible to maintain when treatment effects are pronounced (Sibbald & Roland, 1998; Wilsey *et al.*, 2016). It has to be considered that while

placebos are intended to have no physical effect (Gaddum, 1954), it is known that they may also exert so-called placebo effects, such as phenomenal experiences or therapeutic improvements, through a complex psychological mechanism (Beecher, 1955; Colagiuri *et al.*, 2015). Finally, pharmacological investigations may differ in whether they implement a between-subjects or a within-subjects design (Keren & Raaijmakers, 1988; Landauer, 1975). In the within-subjects design, participants undergo all experimental conditions on different occasions (e.g. placebo administration in one and drug administration in the other session), whereas they are randomly assigned to a single condition in the between-subjects design. While within-subjects designs require fewer participants and reduce statistical noise, between-subjects designs minimize expectancy and transfer effects (Keren & Raaijmakers, 1988).

II.5 FUNCTIONAL MAGNETIC RESONANCE IMAGING

Functional magnetic resonance imaging is a non-invasive technique that allows the indirect imaging of neural activity based on physiological processes in the brain (Babiloni *et al.*, 2009; Scarapicchia *et al.*, 2017). Provided sufficiently large magnetic field strengths, it is characterized by its high spatial resolution, whereas its temporal resolution is limited by the properties of the hemodynamic response (see below) and its finite signal-to-noise ratio (S.-G. Kim *et al.*, 1997; Yoo *et al.*, 2018). fMRI builds on magnetic resonance imaging (MRI), which leverages the intrinsic spin of hydrogen nuclei (protons) found throughout the human body (van Geuns *et al.*, 1999). In their default state, these protons are oriented in an almost completely random manner. The electromagnetic excitation pulse produced by the magnetic resonance (MR) scanner causes the hydrogen nuclei to align in an identical direction. Upon cessation of the excitation pulse, a relaxation of the protons to their initial state becomes observable, whereby an electromagnetic radiation is emitted that is measured as an MR signal through the use of an appropriate detector coil (Huettel *et al.*, 2014). This relaxation is dependent on distinguishable properties of the body tissue, which is integral to MRI (van Geuns *et al.*, 1999). It is furthermore important to make a distinction between longitudinal (T_1) and transverse proton relaxation (T_2/T_2^*), which constitute the basis for generating anatomical (T_1) and functional (T_2^*) brain images, respectively (G. G. Brown *et al.*, 2007). An MR acquisition requires multiple excitation pulses provided in a predetermined measurement sequence. The time interval between two excitation pulses in a sequence is referred to as repetition time (TR), and the time interval between excitation pulse and data acquisition as echo time (TE; Huettel *et al.*, 2014).

fMRI is based on the observation that regionally increased neural activity results in an increase in oxygen metabolism and thus in increased cerebral blood flow (D. G. Norris, 2006; Watts *et al.*, 2018). Oxygen is

transported by the iron-containing (and hence magnetic) protein complex hemoglobin in erythrocytes (Sharma & Premachandra, 1991). Oxygenated and deoxygenated hemoglobin differ in their local magnetic properties, which is at the core of the blood-oxygen-level-dependent (BOLD) contrast, the preeminent method of fMRI (Lindquist *et al.*, 2009; Logothetis, 2002). Unlike oxygenated hemoglobin, deoxygenated hemoglobin causes microscopic magnetic field inhomogeneities in T_2^*-weighted images due to a decrease in T_2^* relaxation times (Guensch *et al.*, 2021; Markett, 2016; Uludağ *et al.*, 2009). An increase in brain activity due to a particular event should result in increased local energy consumption and thereby, oxygenation differences in the corresponding region. To counteract deoxygenation in this region, a reactive local oversupply of oxygenated hemoglobin is delivered (Fox & Raichle, 1986), which is known as the *hemodynamic response* (Friston *et al.*, 1998). A brief initial dip (Xiaoping Hu & Yacoub, 2012) is followed by a steep increase in oxygenation, with the hemodynamic response peaking about five seconds later and then continuing to decline steadily over several seconds until falling below baseline (i.e., the post-stimulus undershoot); the BOLD signal returns to baseline level only about twenty seconds after the onset of the corresponding event (Markett, 2016; Menon *et al.*, 1995). The entirety of this sequence is referred to as the *hemodynamic response function* (Lindquist *et al.*, 2009; see Figure 7).

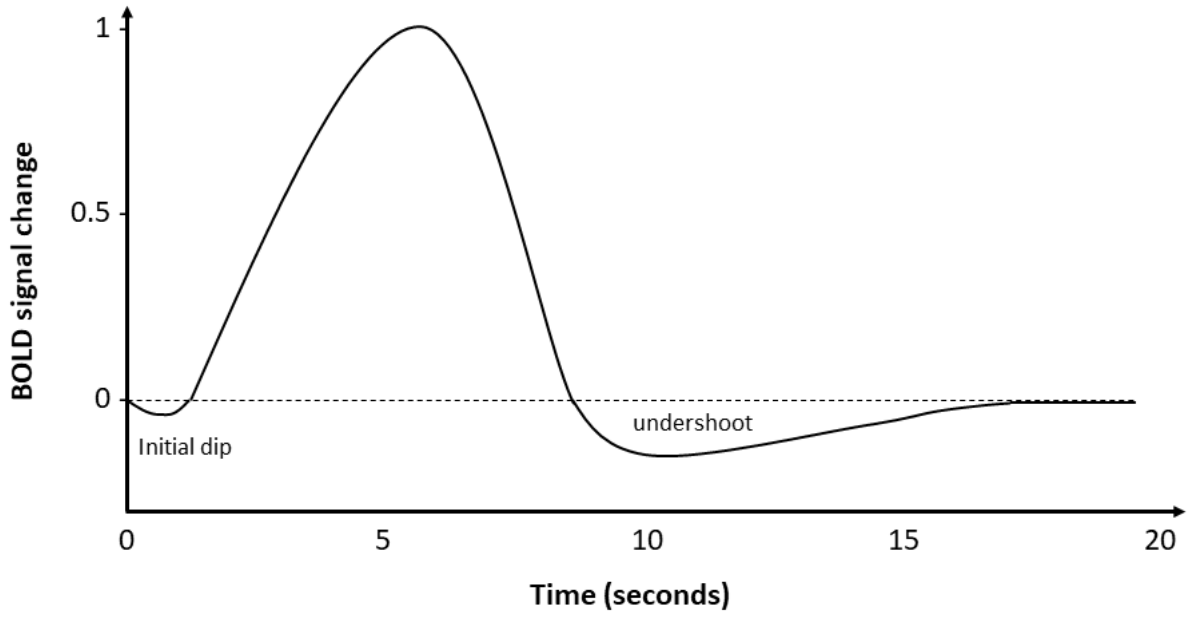

Figure 7. BOLD hemodynamic response to a single short-duration stimulus (schematic representation). In response to a brief local deoxygenation that may arise as a result of initial oxygen uptake (initial dip), there is a steep increase in local oxygenation level until the positive main BOLD response peaks after approximately five seconds. The return to baseline is usually preceded by a post-stimulus undershoot of varying duration (Barth & Poser, 2011).

Before fMRI-related hypothesis testing can be carried out, a series of preprocessing steps has to be applied to the acquired BOLD data to correct for deviations across participants along spatial and, in some cases, temporal dimensions (Huettel *et al.*, 2014). Among a variety of other possible steps, these typically include head-motion correction during realignment, co-registration of anatomical and functional brain images, spatial normalization into a standardized stereotaxic space and spatial smoothing, which altogether substantially alters the acquired data, but ensures the meaningfulness of subsequent first (participant-wise) and second (group-wise) level analyses (Markett, 2016). In a standard fMRI data analysis, experimental conditions are contrasted, meaning that the activation levels evoked by two or more independent variables (or two or more levels of the same independent variable) are compared. This follows the logic of subtraction, as two or more quantities are assumed to differ only in one property, the independent variable (Huettel *et al.*, 2014). Statistically significant differences in whole-brain level intensity are thereby evaluated as meaningful differences in neural activation due to the effect of the independent variable.

Moreover, it may in many cases be desirable to understand the effects of one's independent variable on the complex interplay of brain regions within the underlying neurocircuitry. After all, the human brain would be improperly characterized as a structure consisting of purely modular, functionally specialized parts (Sporns *et al.*, 2005), as a brain area's functioning critically depends on its interactions with other brain regions (Fox & Friston, 2012; D. V. Smith *et al.*, 2016). The measurement of brain region co-activation – i.e., functional connectivity – can therefore inform the understanding of the neurofunctional effects of one's independent variable on a considerably deeper level (Büchel, 2004; Stevens, 2016). Functional connectivity patterns can be assessed either in the absence (i.e., during the "resting-state") or in the presence of an active task (Stevens, 2016). An example for the latter is psychophysiological interaction analysis, a measure of effective functional connectivity that reveals regional changes in the relationship between activations within different brain areas, dependent on the interaction between a physiological (e.g. activity in a given seed region) and a psychological (e.g. a specified context or task) factor (Friston *et al.*, 1997; O'Reilly *et al.*, 2012).

III. STUDY SYNOPSES

Table 2. Overview of studies included in the present dissertation.

Publications in the current thesis	
Study 1	**Lehmann, M.**, Hagen, J., & Ettinger, U. (2022). Unity and diversity of metacognition. *Journal of Experimental Psychology: General, 151*(10), 2396–2417. doi: 10.1037/xge0001197
Study 2	**Lehmann, M.**, Neumann, C., Wasserthal, S., Schultz, J., Delis, A., Trautner, P., Hurlemann, R., & Ettinger, U. (2021). Effects of ketamine on brain function during metacognition of episodic memory. *Neuroscience of Consciousness, 7*(1), niaa028. doi: 10.1093/nc/niaa028
Study 3	**Lehmann, M.**, Neumann, C., Wasserthal, S., Delis, A., Schultz, J., Hurlemann, R., & Ettinger, U. (2022). Ketamine increases fronto-posterior functional connectivity during meta-perceptual confidence ratings. *Behavioural Brain Research, 430*, 113925. doi: 10.1016/j.bbr.2022.113925

Further relevant publications
Lehmann, M., & Ettinger, U. (under review). Metacognitive monitoring in schizotypy: Systematic literature review and new empirical data.
Wasserthal, S., **Lehmann, M.**, Neumann, C., Delis, A., Philipsen, A., Hurlemann, R., Ettinger, U., & Schultz, J. (to be submitted). Effects of NMDA-receptor blockade by ketamine on mentalizing and its neural correlates in humans – a randomized control trial.
Lehmann, M., Plieger, T., Reuter, M., & Ettinger, U. (under review). Insights into the molecular genetic basis of individual differences in metacognition.

III.1 STUDY 1: THE STRUCTURE OF INDIVIDUAL DIFFERENCES IN METACOGNITION

Lehmann, M., Hagen, J., & Ettinger, U. (2022). Unity and diversity of metacognition. *Journal of Experimental Psychology: General, 151*(10)*,* 2396–2417. doi: 10.1037/xge0001197

As outlined in detail in Chapter I.1.4, a variety of factors contributes as to why pivotal questions about the structural organization of the ability for metacognition remain largely unresolved. In essence, this pertains to the employment of underpowered samples, insufficiently specific instruments to measure metacognition, inconsistent use of metacognition metrics, and the inherent limitations of correlational analysis methods that cannot adequately account for the task impurity problem. Taking these factors into account, this study complemented the correlational approach by a latent variable approach. For the latter, confirmatory factor analyses were conducted, which allowed to investigate the question of the structure of metacognition at the level of latent variables (representing functional domains), whose interrelationships were defined in several models, the representativeness of which could be assessed with respect to the empirical data (Miyake *et al.*, 2000). First, the so-labeled "specificity model" assumed an entirely modular structure, which was achieved by fixing the covariances between the three latent variables to 0. Conversely, the "G-factor model" implied the extraction of a single metacognitive resource across domains, with all covariances between the latent variables fixed to 1.0. Models allowing the coexistence of generality and specificity comprised the "Three-factor model" and three nested Bi-factor models. In the Three-factor model, all inter-domain covariances were freely estimated, whereas in the Bi-factor models, one of the three inter-domain covariances was fixed to 1.0 (thereby equating two domains), with the other two covariances freely estimated and then equated.

A total of 155 healthy individuals underwent a series of tasks, two metacognition tasks for each experimental domain: The attention-to-action-domain comprised *a)* an antisaccade task, which required participants to direct their eye gaze in the opposite direction of a given cue stimulus before providing a confidence rating on a discrete six-level scale in having reacted accordingly, and *b)* a Go/No-Go error awareness task with two different No-Go conditions and context-dependent signaling of error awareness. In the domain of visual perception, *a)* a classical random-dots motion perception task, with a number of dots either moving to the left or to the right in varying degrees of coherence, was followed by *b)* a magnitude comparison/ set discrimination task with static stimuli, with either the left or the right of two circles containing more dots; performance was titrated by an adaptive staircase (see Chapter II.1.2) in both tasks and confidence ratings on a discrete six-level scale were collected after each response. Finally, the domain of episodic memory comprised two structurally analogue retrieval tasks with *a)* word and *b)* face stimuli that should be classified either as old (i.e., encoded ~90 minutes prior

to retrieval) or as new stimuli, with trial-by-trial ratings on the same confidence rating scale. Furthermore, participants provided self-report ratings on three different questionnaires (MCQ-30, MAI and BCIS), which were conceptualized as measures of "metacognition in everyday life", i.e., of the latent, multi-level metacognitive assumptions about oneself assumed to exist across tasks and situations in the model by Toglia and Kirk (2000; see Chapter I.1.3).

Confirmatory factor analyses revealed an outcome structure that was overall more consistent with a domain-general than a domain-specific account, as they yielded a mostly good fit for the G-factor model, whereas invalid model estimates were obtained for the specificity model, which yielded a clearly inferior fit compared to all other models. As the G-factor model constituted the most restrictive model, having the smallest amount of freely estimated parameters, a model decision in favor of the G-factor model was considered; however, the G-factor model's CFI was slightly below the lower boundary of an acceptable fit. Good or acceptable fits were obtained for the models providing a combination of generality and specificity, in particular for the Bi-factor model which equated attention-to-action and memory, suggesting a common metacognitive resource for these two domains and a substantial degree of functional specialization in metacognitive processes underlying visual perceptual judgments. Given the absence of clear-cut differences in CFA model fits according to χ^2-difference tests and AIC/BIC, no particular model was endorsed; instead, the result pattern was discussed descriptively with a particular focus on the Bi-factor model with an isolated perceptual domain. This was corroborated by the results from correlational analysis, in which significant inter-domain correlations were found between the antisaccade task and both episodic memory tasks. Given the non-significance of correlations for the majority of task pairings, a substantial degree of contribution of idiosyncratic domain-specific signals to the structure of metacognition needs to be considered, which is less compatible with a "G-factor", but naturally accommodated in a "unity and diversity" account assuming the coexistence of generality and specificity. Metacognitive bias, on the other hand, was found to be characterized by a high degree of generality across tasks. Taken together, the results suggest that a composite factor arguably best characterized as "metacognition for cognition" (comprising behavioral inhibition in the domain of attention-to-action as well as episodic memory retrieval) may be distinguished from a possibly independent metacognitive resource subserving visual perceptual judgments (Figure 8).

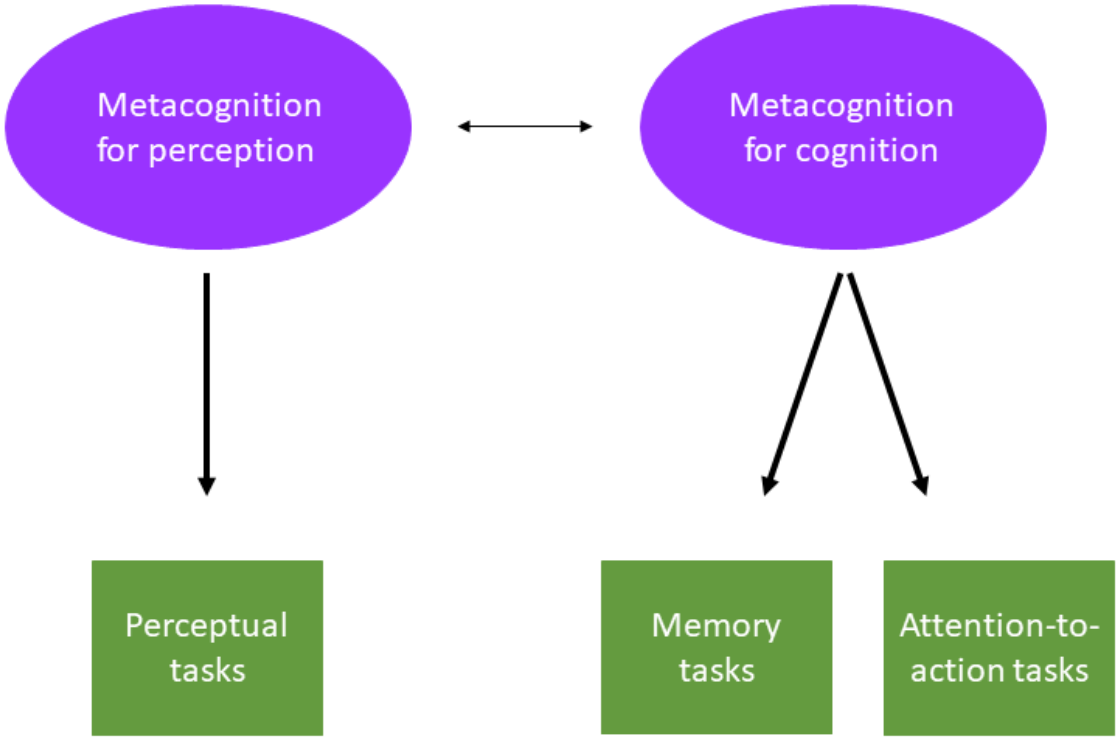

Figure 8. Key finding of Study 1: Unity and diversity in the capacity for metacognitive monitoring. In contrast to the binary assumption of generality versus specificity in metacognitive processes (Figure 1), as sketched out by Rouault, McWilliams, *et al.* (2018), it was found that perceptual judgments may be monitored and controlled by a dedicated metacognitive subsystem, whereas a substantial degree of commonality can be assumed in the metacognitive processes underlying judgments in two more "cognitive" domains, i.e., episodic memory and behavioral inhibition in the domain of attention-to-action.

No reliable association emerged between online metacognitive performance in the laboratory and offline self-report measures of metacognition. There was, however, a moderate association with metacognitive bias, evident particularly in the memory domain. Overall, this underlines the paradox that providing realistic insight into one's own monitoring processes already requires a high degree of metacognitive ability, suggesting that self-reports may only represent a valuable asset in research on participants with advanced metacognitive skills (Grossner *et al.*, 2021; Lysaker, Buck, *et al.*, 2011).

III.2 STUDY 2: PHARMACOLOGICAL MODULATION OF METAMEMORY

Lehmann, M., Neumann, C., Wasserthal, S., Schultz, J., Delis, A., Trautner, P., Hurlemann, R., & Ettinger, U. (2021). Effects of ketamine on brain function during metacognition of episodic memory. *Neuroscience of Consciousness, 7*(1), niaa028. doi: 10.1093/nc/niaa028

Whereas the dose-dependent anesthetic effect of the NMDAR antagonist ketamine has long been established, researchers have only begun to piece together a thorough account of the wide range of cognitive and perceptual alterations induced by the pharmacological compound at subanesthetic levels (Vlisides *et al.*, 2018). This study was the first to investigate the effects of an intravenous ketamine challenge – and thereby the involvement of the glutamatergic system – on the precision of metamnestic evaluations while acquiring fMRI data. For a number of reasons outlined in Chapter I.4, the integrity of metacognitive operations was assumed to be perturbed as a consequence of ketamine administration.

Implementing a randomized, double-blind, placebo-controlled between-subjects design, 53 healthy volunteers were administered either a psychotomimetic dose of ketamine or a placebo saline solution. Metacognitive reports about retrieval judgments were collected via trial-by-trial confidence ratings on a discrete six-level scale in two separate study phases, either in the presence or the absence of drug infusion: In Study Phase I, a first word list was encoded prior to drug administration and MRI assessment, and subsequently retrieved during drug infusion and MRI data collection. Conversely, Study Phase II consisted of the encoding of a second word list as participants were yet undergoing drug infusion and placed in the MRI scanner, for which retrieval performance was tested after completion of the MRI scanning and termination of the infusion. Thereby, ketamine effects on encoding, retrieval and metacognitive performance could be evaluated independently of each other.

It was found that while acute ketamine administration left retrieval performance and response times unaffected, the ketamine-induced altered state of consciousness was characterized by a deterioration of absolute metacognitive sensitivity (see Figure 9A) and a tendency for overconfidence, i.e., larger metacognitive bias. Importantly, however, the performance-corrected index of relative metacognitive sensitivity or efficiency was not significantly different between the groups; this was surprising given the absence of ketamine effects on retrieval performance, which in principle should not reflect negatively on metacognition. Exploratory application of a hierarchical Bayesian analysis (Figure 9, Panels B-C) on group-level values of metacognitive efficiency revealed a bimodal distribution for the two groups; however, the credible interval of this group difference narrowly overlapped with zero, so inference of a significant group difference in metacognitive efficiency cannot be undertaken with sufficient certainty (Fleming, 2017). This could be interpreted as a result of non-significant group-heterogeneity in retrieval

performance, but also of insufficient statistical power (Cohen, 1992). Ultimately, the pharmacological challenge was associated with moderate impacts on metacognitive evaluations, with both a tendency for reduced accuracy and overconfidence under ketamine.

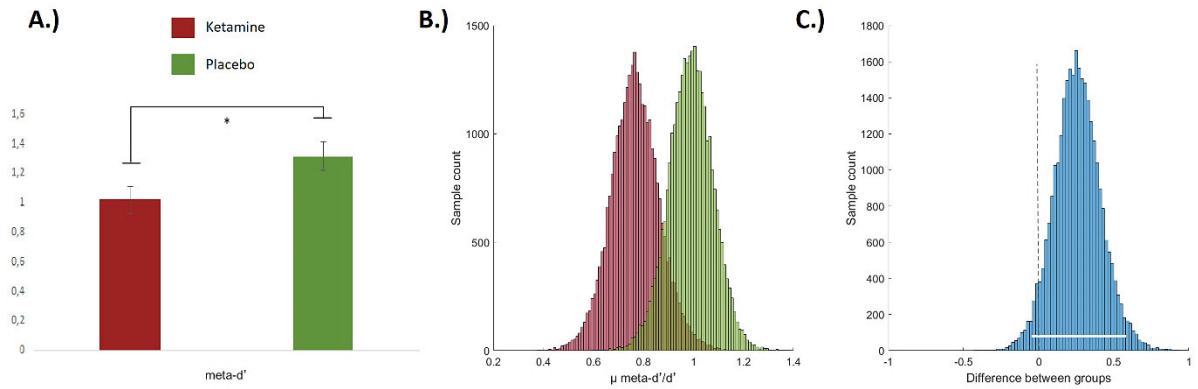

Figure 9. Key finding of Study 2: Pharmacological modulation of metamemory processes. Panel A: Significant between-group difference in metacognitive sensitivity in Study Phase I. Error bars indicate standard errors; *$p < .05$. Panel B: Hierarchical Bayesian estimation of metacognitive efficiency ($\mu_{meta\text{-}d'/d'}$) in Study Phase I, resulting in a bimodal distribution for group-level estimates. Panel C: Difference in group posteriors for hierarchical Bayesian estimation (in log units), with the credible interval narrowly overlapping with zero, as indicated by the white horizontal line. Red bar/ histogram: Ketamine, Green bar/ histogram: Placebo.

Secondary study aims encompassed identifying the effects of varying levels of processing during encoding on retrieval and metacognitive performance as well as investigating group differences on the 5D-ASC (Dittrich, 1998). Significant group differences on 5D-ASC ratings confirmed that ketamine elicited an altered state of consciousness, and a significant difference between deeply and shallowly encoded words was found consistently across behavioral outcome measures and response times, with the exception of metacognitive efficiency in Study Phase I. Finally, there was no significant difference between groups in Study Phase II, neither at the behavioral nor at the brain functional level, suggesting that the described effects were restricted to the altered state of consciousness during acute ketamine administration.

With respect to group differences in BOLD activation, ketamine was associated with higher neural activity during second-order confidence ratings in several posterior cortical clusters, anchored in superior and inferior parietal lobule as well as occipital structures (calcarine and lingual gyrus). The specificity of this effect was limited, as it was observed aggregating over both metacognitive confidence ratings and a matched control condition. Nonetheless, this up-regulation of posterior visuospatial

cortical areas was indicative of an association with the "hot zone" of the NCCs (Koch *et al.*, 2016), and activations in these areas track phenomenal qualities and the overall reliability of the sensory input to the metacognitive process rather than the genuine metacognitive evaluation itself (Bang & Fleming, 2018), which is thought to be a function of a frontoparietal network, as previously established (Vaccaro & Fleming, 2018). Consequently, early aspects of the metacognitive process could be impacted by ketamine, as the metacognitive system is required to raise additional efforts to make sense of a distorted input signal. A different, but not mutually exclusive explanation would regard the up-regulation of posterior cortical brain areas during metacognition as signaling alterations in conscious experiences, which would lead to trial-to-trial variability in the placement of confidence criteria.

III.3 STUDY 3: PHARMACOLOGICAL MODULATION OF METAPERCEPTION

Lehmann, M., Neumann, C., Wasserthal, S., Delis, A., Schultz, J., Hurlemann, R., & Ettinger, U. (2022). Ketamine increases fronto-posterior functional connectivity during meta-perceptual confidence ratings. *Behavioural Brain Research, 430*, 113925. doi: 10.1016/j.bbr.2022.113925

With respect to dissociations between metacognitive processes in relation to perceptual judgments and episodic memory retrieval both at the level of functional properties and neurobiological underpinnings, a subsequent aim was to assess the involvement of the glutamatergic system in the formation of metaperceptual judgments. In this context, an important methodological difference between the domains has to be considered: In the perceptual domain, it is customary to adapt the difficulty of the respective task to each individual's performance by means of a staircase procedure, which entails both advantages and disadvantages for the computation of metacognitive ability (see Chapter II.1.2). Indeed, such an adaptive staircase procedure was applied in this study employing a randomized, double-blind, placebo-controlled between-subjects design, in which participants completed a perceptual decision-making task with trial-by-trial ratings on a six-level metacognitive confidence scale during fMRI, while being administered either the identical psychotomimetic dose of ketamine as in Study 2 or placebo saline solution. After exclusion of dropouts and participants with failed titration of task performance at a targeted level, data from 45 participants were included in final data analysis.

In close resemblance to the modulation of metamemory by ketamine in Study 2, a significant drug effect on absolute metacognitive sensitivity (meta-d') was found at the behavioral level, whereas the group difference on metacognitive efficiency was narrowly non-significant. There was again no clear evidence of confounding of metacognitive efficiency parameters by any ketamine effects on Type I performance and/or reaction times; likewise, newly added parameters related to the adaptive staircase procedure

(i.e., the absolute perceptual threshold at the beginning and during the experiment, as well as the overall variability of difficulty levels) were not significantly different between groups. Hence, the divergence between ketamine effects on absolute (meta-d') and relative (meta-d'/d') metacognitive sensitivity could yet again be due to non-significant group-heterogeneity in Type I performance as well as a potential lack of statistical power, which would be associated with the extensive exclusion of participants due to unsuccessful calibration of task difficulty. However, the overall informativeness of absolute metacognitive sensitivity indices should be increased due to application of the staircase method (Fleming & Lau, 2014). Finally, there was no significant association of ketamine with metacognitive bias in this study; descriptively, it became salient that ketamine was associated with inordinately higher levels of confidence in a subset of participants, but with markedly lower levels of confidence in others. This could be related to potentially trait-like individual differences which might be amplified by ketamine application.

At the brain functional level, there were no significant group differences in mean BOLD activation, contrary to expectation. However, exploratory investigations of regional changes in activity of different brain areas in relation to a specific task (here: genuine metacognitive introspection versus a control condition involving only the motor component) revealed increased functional connectivity between core frontal areas of metacognition and posterior (i.e., occipital, temporal, and posterior frontal) structures (Figure 10). This could reflect the deployment of information from structures involved e.g. in the recollection of visual memories, somatosensory processes or the integration of internal action-feedback loops. Such an explanation would be consistent with the inference that metacognition is domain-generally affected by acute ketamine administration, but that the fundamental functionality of metacognitive operations is only moderately compromised, as the metacognitive system incorporates and integrates additional information, e.g. in the form of re-representations of object-level features.

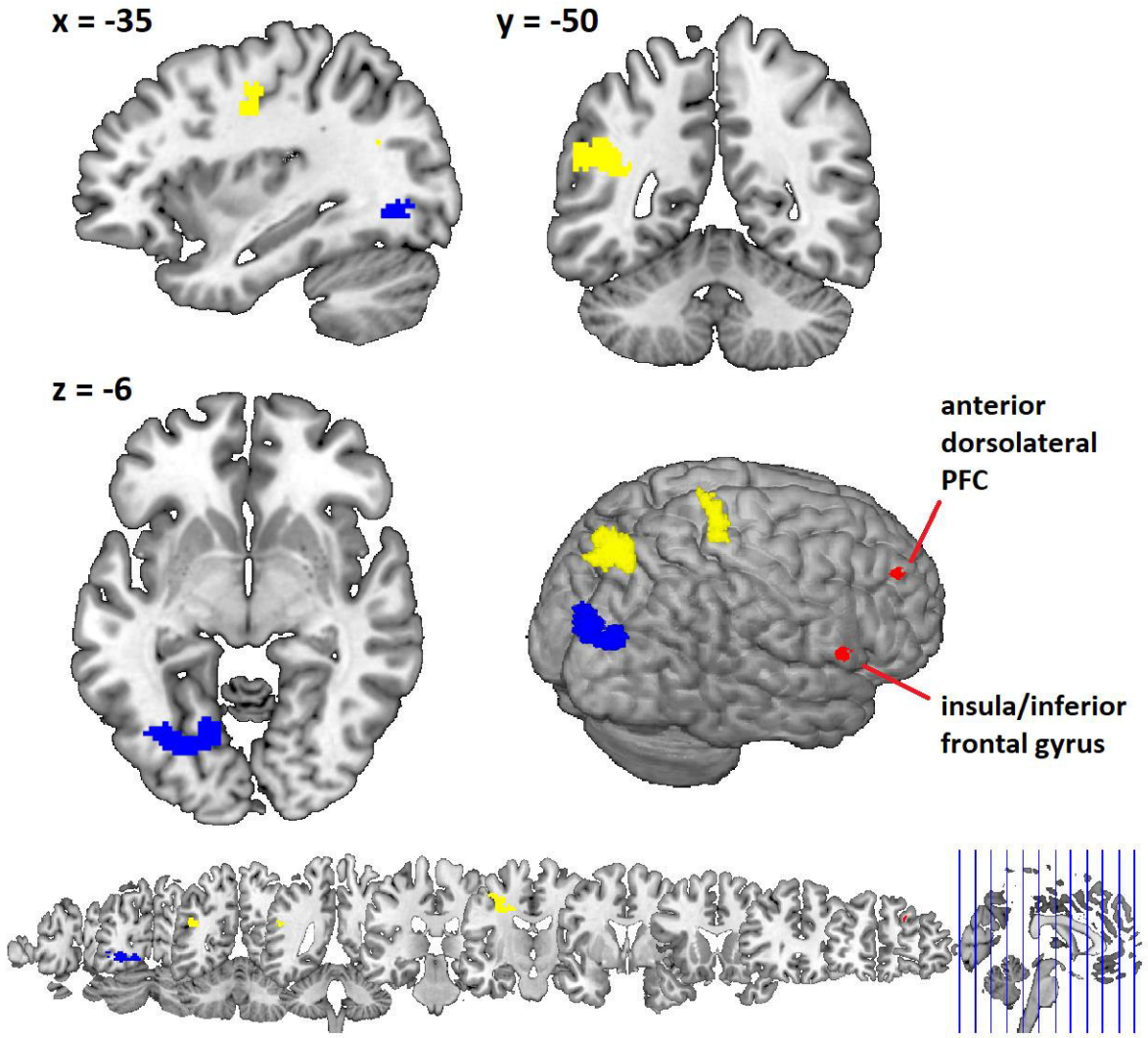

Figure 10. Key finding of Study 3: Increased frontoposterior task-specific connectivity during metaperceptual confidence ratings. Blue and yellow colors indicate areas with significantly ($P < .05$, corrected for family-wise error) higher context-dependent connectivity under ketamine compared to placebo with the seed voxel in anterior dorsolateral prefrontal cortex (PFC, yellow) and with the seed voxel in right insula/inferior frontal gyrus (blue) during trials involving genuine metacognitive introspection versus a matched control condition. Positions of seed voxels are highlighted in red. Republished from "Ketamine increases fronto-posterior functional connectivity during meta-perceptual confidence ratings" by M. Lehmann, C. Neumann, S. Wasserthal, A. Delis, J. Schultz, R. Hurlemann, and U. Ettinger, 2022, *Behavioural Brain Research, 430*, p. 7; copyright (2022), with permission from Elsevier.

IV. DISCUSSION

The three studies included in the present thesis aimed to advance the scientific understanding of the architecture of metacognition, both regarding the organization of its functional properties (**Study 1**) and with respect to its biological substrates at the brain functional and neurotransmitter levels (**Studies 2 and 3**).

IV.1 INTEGRATION

The integrity and functionality of the metacognitive system is fundamental to a wide range of areas and research questions, not the least of which is our understanding of mental health and psychopathology (Seow *et al.*, 2021). Nonetheless, the literature review in the introductory chapters identified a number of open questions with regard to the structure of metacognition across situations and demands, as well as with regard to the neuropharmacological basis of metacognitive processes. The findings of **Study 1** suggest that neither a fully unitary nor a fully modular account of metacognition appear sufficiently adequate, as substantial commonalities and differences in metacognitive ability estimates were found across task domains. Instead, a rather balanced contribution of unified resources and domain-specific signals may be inferred, as the metacognitive system feeds off both task-independent and task-specific components and processes. More specifically, convergent evidence from correlational and latent variable approaches corroborated the notion that perceptual judgments may be monitored by a dedicated metacognitive subsystem, whereas the metacognitive processes underlying judgments in the domains of attention-to-action and episodic memory appear to draw – at least to a substantial degree – upon a unitary resource. This leads to the question as to what is distinctive about metaperception and which mechanism underlies the clustering of more "cognitive" domains such as memory and attention-to-action, which will be discussed in the following.

In general terms, metacognition is understood as a process that integrates sensory evidence and information about one's interactions with the external world (Wokke *et al.*, 2020); within this conceptualization, however, it is reasonable to postulate the existence of two broad, distinct categories of metacognitive mechanisms: Those that monitor primarily externally-generated (e.g. perceptual sensory information) versus those that monitor primarily internally-generated (e.g. memory or motor commands) signals; note, however, that action monitoring as a whole can be considered as a hybrid function integrating both externally-generated and internally-generated information (Arbuzova *et al.*, 2022). This is also reflected at the level of anatomical connectivity, with separate coding schemes within the PFC for external and internal information (Passingham *et al.*, 2010). Conversely, recent evidence

suggests that metacognitive domains may not be empirically aligned along a dimension of externally- versus internally-generated information source, a conclusion mostly based on the lack of correlation between memory and motor tasks (Arbuzova *et al.*, 2022). This null finding and its discrepancy with the results obtained in Study 1 should not be given too much weight, however, as the study, at this point only available as preprint, employed a possibly insufficient sample size of only forty participants.

Consistent with the notion of generality and specificity coexisting within the structural organization of metacognition, Morales *et al.* (2018) provided evidence for the coexistence of domain-general and domain-specific signals in the human brain. More precisely, they differentiated generic lower-level confidence-related signals in the frontal and posterior midline (i.e., frontoparietal and cingulo-opercular regions) from content-rich metacognitive representations in aPFC tracking higher-order contextual information with a high degree of specificity (Bang & Fleming, 2018). In keeping with this, Baird *et al.* (2013) suggested the existence of unique neural networks anchored in medial and lateral regions of the aPFC for metaperceptual and metamemory judgments based on analyses of resting-state connectivity. Other evidence, however, suggests the frontopolar aPFC to contribute essentially domain-specific codes to a markedly greater extent for the computation of perceptual metacognition than what is necessary for metamnestic processes. For instance, Fleming *et al.* (2014) described specific deteriorations in metaperception for patients with lesions to the aPFC, whereas the accuracy of metamemory judgments was found to be unimpaired.

With regard to the mechanisms behind this potentially distinctive involvement of frontopolar cortical regions in metaperception, one might postulate that certain coding schemes could have evolved which are of specific relevance to aspects of perceptual awareness in general and perceptual metacognition in particular. As metaperception is inherently linked to assessing the quality of ongoing experiences, perceptual confidence – unlike memory-related confidence – appears to be driven by the strength of a signal (Koizumi *et al.*, 2015; Peters *et al.*, 2017; Samaha *et al.*, 2019), with distinct detection-specific activation profiles within the frontopolar cortex (Mazor *et al.*, 2020). According to a higher-order state space approach, the higher-order (metacognitive) inference level feeds off these detection-like signals as low-dimensional perceptual awareness states (Fleming, 2020). In this vein, it is eminently plausible to assume a distinct circuit underlying the metacognitive evaluation of more “cognitive” information, where in the case of confidence ratings following memory or attention-to-action judgments, the signal being evaluated is primarily an internally generated one. The placement of a wider range of metacognitive domains into the established knowledge landscape on the brain functional basis of metacognition proves intricate due to the traditional focus on mnestic and perceptual processes. In light of the empirical evidence outlined above, however, it can altogether be assumed that the behavioral dissociation between "metacognition for cognition" and "metacognition for perception" is also mirrored at certain levels of the underlying neurofunctional architecture.

Meanwhile, a predominantly domain-general account seems adequate for metacognitive bias, with substantial intercorrelations across most tasks. An overall similar pattern of results was recently described by Xiao Hu *et al.* (2022), who identified a high degree of generality in the mechanisms underlying the confidence rating process itself and a high degree of specificity for metacognitive accuracy estimates between perceptual and memory tasks. In sum, Study 1 corroborates and expands the existing research landscape on the unity and diversity of task-based metacognitive processes.

Whereas a substantial degree of shared variance between point estimates of metacognition was revealed within the task-based approach, no reliable evidence was found for correlations of metacognitive ability estimates across method-groups. As indicated in Chapter I.1.4, the identification of any such association is considerably complicated by the fact that online and offline approaches share few instrumental characteristics (Rouault, McWilliams, *et al.*, 2018). It follows from Toglia and Kirk's (2000) model that inferring online measures of metacognition from offline measures and vice versa involves a higher level of conceptual abstraction than the comparison of online measures collected in different tasks or domains. Moreover, correlations between online measures could also be associated with factors unrelated to genuine metacognition that may be activated within a given task. It is noteworthy, however, that an association between online metacognitive bias and offline self-report monitoring tendencies was observable to some extent.

Consequently, it is necessary to point to the well-known drawbacks of questionnaire measurements, e.g. the desire for positive self-representation, which may be particularly relevant in the context of self-report assessment of metacognition: After all, the accurate evaluation of one's own monitoring processes across situations itself presupposes an advanced degree of metacognitive insight, which is only likely to apply to a subset of the sample. This aligns with previously expressed concerns that the employment of self-report tools for measuring metacognition is based on overly simplistic assumptions (Grossner *et al.*, 2021). On the other hand, the potentially insufficient reliability of behavioral measures might contribute to this null finding (Dang *et al.*, 2020), which may, after all, be more strongly influenced by state-related fluctuations and/or distorted metacognition metrics than self-report measures.

Importantly, a remarkably homogeneous pattern of findings regarding behavioral effects of ketamine emerged across **Studies 2 and 3**: In both studies, the induction of a substantially altered state of consciousness was evidenced by significant group differences on all scales of the 5D-ASC questionnaire, with the most pronounced effects on Vigilance Reduction (i.e., feelings of drowsiness) and Oceanic Boundlessness (i.e., pleasant aspects of the experience). This ensured and validated the phenomenological effects of the pharmacological compound, as any ketamine effects on metacognition would be expected specifically within the context of an altered state of consciousness (Carhart-Harris *et al.*, 2014). Neither study observed significant ketamine effects on task performance nor a significant

deceleration in reaction times following ketamine administration, although it should be noted that a descriptively higher Type I sensitivity in the placebo group was consistently observed across experiments. The main finding of significant group differences in absolute metacognitive sensitivity (meta-d') and non-significant group differences in relative metacognitive sensitivity (meta-d'/d') was likewise achieved in both studies and will be subject to a more detailed discussion in the following. Finally, the only noticeable difference in ketamine effects on behavioral outcome measures between the two studies emerged with respect to metacognitive bias: Whereas Study 2, in line with the observation of overconfidence in incorrect responses in individuals with schizophrenia (Balzan, 2016), found a significantly increased performance-corrected metacognitive bias in the ketamine group, no significant group difference emerged in Study 3. Descriptively, however, the values of the ketamine group were found to be clustered around the margins on either side of the distribution of bias scores. In line with Study 1, which suggested a task-independent, trait-like character for metacognitive bias, one might therefore argue that a priori trait-like Type II response tendencies were transiently biased towards under- or overconfidence by ketamine. Pending availability of additional data on the association of ketamine with general confidence level, such considerations remain of a speculative nature.

With respect to metacognitive performance, both studies yielded evidence for a significant effect of ketamine on absolute metacognitive sensitivity (meta-d'), whereas frequentist and Bayesian approaches failed to reveal significant ketamine effects on performance-corrected relative metacognitive sensitivity (meta-d'/d'). As discussed previously, post-hoc disentanglement of the two opposing explanations (*i*) effect does truly not exist, because non-significant group-heterogeneity in Type I sensitivity inflated parameter estimates of absolute metacognitive sensitivity, and (*ii*) effect exists, but remained insignificant due to a lack of statistical power, is not possible to a satisfactory extent. In extensive, multi-method experimental studies such as the ones in question, recruitment of even larger sample sizes is not always economically justifiable. It follows, however, that any statement on the involvement of glutamate in metacognitive processes must be made with great caution.

Considering the application of a staircase procedure in Study 3, which – in spite of methodological concerns (Rahnev & Fleming, 2019; see Chapter II.1.2) – offers a straightforward possibility to eliminate performance-related differences in metacognitive ability estimates, inferences from Study 3 could potentially be drawn with greater confidence, given the higher relative informativeness of ketamine effects on absolute metacognitive sensitivity. Despite the absence of significant group differences in metacognitive efficiency, Study 3 therefore suggests glutamate to represent a building block in the neuropharmacological basis of metaperceptual processes. With respect to the observations of Lofwall *et al.* (2006) and Carter, Kleykamp, *et al.* (2013), who described no effect of intramuscular ketamine on memory-related metacognition, the degree of involvement of glutamate in metamemory is perhaps not quite as clear in comparison. In any case, further research on the neuropharmacological basis of

metacognitive processes is warranted. Taking into account the aforementioned limitations, however, the observed metacognitive deficits during acute ketamine administration could be regarded as further evidence for the adequacy of the glutamate model of psychosis (Javitt, 2012; Javitt *et al.*, 2012). Going forward, pharmacological application of uncompetitive NMDAR antagonists such as ketamine could hence increasingly be used for expanding knowledge on neurobiological and psycho-behavioral aspects of schizophrenia spectrum disorders as well as for advancing the development of therapeutic applications.

In sum, the behavioral findings from Studies 2 and 3 suggest a domain-general role of the glutamatergic system in metacognition, while the moderate effect sizes are highly suggestive that additional neurotransmitter systems underlie the overall integrity and functionality of metacognitive operations. Nevertheless, the pattern of findings from Studies 2 and 3 indicates that glutamate might represent a common pharmacological substrate for metacognitive operations, possibly as an enabling factor rather than as an actual core substrate of confidence formation, which is supported by the convergent observation of moderate behavioral impacts as well as posterior neural correlates of these effects.

At the brain functional level, activation differences between groups with respect to the metacognitive confidence rating element of the trial sequence were found in posterior rather than frontal structures in both studies. This might initially be surprising, since metacognition is thought to be the function of a frontoparietal network (Vaccaro & Fleming, 2018) considered to comprise the core substrates of confidence formation (Bang & Fleming, 2018). Posterior regions, on the other hand, are associated with the hot zone of content-specific neural correlates of phenomenal and sensory information (Koch *et al.*, 2016). Importantly, posterior brain areas have been implicated in metacognition, especially in metamemory (Elman *et al.*, 2012; McCurdy *et al.*, 2013), which may account for the unspecific up-regulation of posterior regions during second-order confidence ratings in Study 2 that was not similarly observed with perceptual stimuli in Study 3. Instead, Study 3 yielded evidence for increased context-dependent functional connectivity between the medial frontal and prefrontal core areas of metacognition with contralateral posterior brain regions such as the occipital or medio-temporal cortex. Along these lines, it is established that an interplay of frontal and posterior brain regions is characteristic of metacognition, in the sense of a hierarchical loop between more (pre-)frontally encoded meta-level processes feeding off more posteriorly encoded object-level processes – e.g. visual or sensorimotor information (Fleming & Dolan, 2012). The finding of increased frontoposterior functional connectivity may hence be indicative of the representation of object-level information precision maintained from regions involved in primary stimulus categorization which may then be re-represented for use in metacognitive report (Fleming, Huijgen, *et al.*, 2012). Such re-representations of maintained object-level features might increasingly be invoked under ketamine, as part of a compensatory mechanism to maximize the amount of information provided to the metacognitive process upon detection of a

perturbed input signal. Given the moderate effect sizes of behavioral ketamine effects on metacognition, this might also explain why the fundamental functionality of metacognitive processes appears to be only moderately affected. While Studies 2 and 3 thereby indicate consistent cross-domain impacts of acute ketamine administration on metacognition, these are more likely to pertain to a common early basis or signal source available to both mnestic and perceptual processes, and possibly the influence of ketamine-induced trial-by-trial fluctuations in conscious awareness. Importantly, these alterations are nonetheless specific to the Type II process, as the BOLD results of both studies indicate that the metacognitive system incorporates and integrates additional information by means of a compensatory looping mechanism re-representing object-level features. These findings may thereby provide insight into how the pharmacologically challenged metacognitive system operates to generate a sufficiently accurate outcome. It is therefore cautiously concluded that Studies 2 and 3 were able to reveal a low-level common pharmacological substrate for mnestic and perceptual metacognition.

Overall, the glutamatergic system appears to constitute an enabling factor for the integrity and functionality of metacognitive processes. Challenging it markedly disrupts the metacognitive system, which can, however, largely be compensated for by the provision of re-represented object-level information. Importantly, it is not suggested that an involvement of the glutamatergic system constitutes an exhaustive account of the neurobiological basis of metacognition at the neurotransmitter level. Highly differentiated higher-order functions such as metacognition likely require the complex interplay of various neurotransmitters, which might include e.g. the dopaminergic (Joensson *et al.*, 2015) or noradrenergic (Hauser *et al.*, 2017) systems. As such a unity and diversity structure has previously been illustrated for executive functions (Miyake *et al.*, 2000), the combination of both domain-general and domain-specific features could be deeply embedded into the structural organization of many higher-order cognitions. Such a postulate seems reasonable to ensure a flexible balance between the respective advantages of generality and specificity outlined in Chapter I.1.4, and may ultimately be reflected in many of the established patterns of individual differences across various life contexts and domains. The findings from Studies 1 to 3 are hence overall suggestive of unity and diversity of metacognition at the level of functional properties and neurobiological (in particular, neuropharmacological) underpinnings.

IV.2 LIMITATIONS

Due to the interplay of several factors, the stand-alone interpretative value of the studies included in this dissertation is limited. Although Study 1 extended its scope of analysis beyond the classically implemented domains of visual perception and episodic memory, the addition of other task domains would have been desirable: For instance, domains such as motor function, emotion, or even semantic and working memory could have been included to paint as holistic a picture as possible of the structure of metacognitive processes. Likewise, the use of three or more indicators (tasks) per domain would have been optimal (Kenny *et al.*, 1998); however, this could not be implemented for reasons of test economy.

Generally, shortcomings regarding the study sample can be cited for all studies: For Study 1, this pertains mostly to the unbalanced gender distribution with markedly more female participants, as well as to the restriction to a young adult sample (18-35 years) to reduce variance in first-order task performance. This may have raised the difficulty, however, to identify e.g. an association between online and offline measures, since age might exert a moderating influence (Harty *et al.*, 2013). Likewise, only little is known about the developmental path of metacognition across the life span, with some indication of a shift from domain-specificity to domain-generality over the course of childhood (Geurten *et al.*, 2018). Meanwhile, questions regarding sample size must be raised over Study 2 and perhaps even more so over Study 3, in which a technical error in staircase calibration resulted in the exclusion from analysis of about a third of the originally tested participants. Given sample size limitations, the implication of ketamine effects on metacognition may only be made tentatively, and as pointed out previously, it is ultimately possible that non-significant group-heterogeneity in Type I sensitivity may have inflated the group differences in metacognitive performance in both studies.

A related issue concerns the employment of a between-subjects design in Studies 2 and 3, as such is always associated with the risk that despite the random allocation of participants, a priori intrinsic differences between groups may distort the outcome. Likewise, the straightforward inference of the role of the glutamatergic system for the integrity of metacognitive processes based on ketamine application cannot be undertaken with conclusive certainty, since uncompetitive NMDAR antagonists only partly affect effective physiological glutamatergic neurotransmission (Lipton, 2004), which at the same time likely contributes to the good safety record of ketamine in neuroscientific research (Wolff & Winstock, 2006).

Finally, the instrumental characteristics (e.g. at the level of stimulus parameters) were not optimally balanced to represent a “pure” index for a relationship of interest (across tasks, domains and method-groups) that is not also influenced by method variance. This issue applies to all three studies, partly owing to scientific conventions, such as the fact that staircase procedures are common and readily

applicable only in the perceptual domain. However, this represents a potential experimental confound which e.g. might contribute to the finding of a relatively isolated perceptual domain in Study 1, as the latent factor of metaperceptual indicators could arguably be modeled more homogenously than other factors. For Studies 2 and 3, confidence in the postulate of a common pharmacological substrate of metacognitive processes is necessarily limited, as it results from the post-hoc comparison of different studies with different samples and – to some extent – different instrumental task characteristics. Nonetheless, the finding of similar effects of the same pharmacological intervention across studies could equally be cited as an argument for an increased generalizability of the observation.

Consideration of these shortcomings gives way to a nuanced view on the studies included in the present dissertation, as they merely contribute puzzle pieces to a much larger mosaic of the architecture of metacognition, from which impulses for follow-up studies can be extrapolated.

IV.3 FUTURE RESEARCH AVENUES

Despite the growing interest in systematic investigations of metacognition, research on individual differences in the ability for metacognitive monitoring still has plenty of ground to cover. Most importantly, robust links between metacognition measured in the laboratory and relevant outcome parameters in everyday life have yet to be established: While the G-factor of intelligence, for instance, is known to be associated with academic achievement, job performance, and health-related factors (Gottfredson & Deary, 2004; Mayes *et al.*, 2009; Watkins *et al.*, 2007), such knowledge with regard to metacognition is still in its infancy. In light of this, the identification and development of appropriate proxies for assessing metacognition in real-world settings (e.g. domestic, recreational and/or occupational activities) represents a promising avenue. Similarly, for the task-based approach to measuring metacognition, it would be purposeful to integrate aspects with a strong and discernible relevance to everyday life, which fundamentally alludes to the need for incorporation of additional task domains in the experimental study of metacognition to obtain as holistic a picture of its structure and taxonomy as possible.

Likewise, additional research is warranted to foster understanding of the biological underpinnings of metacognition at the neurotransmitter level. As opposed to many previous research efforts, future studies should make sure to employ state-of-the-art approaches to quantifying metacognitive performance. Beyond the approach presented in this dissertation, which infers the involvement of specific neurotransmitter systems by the degree to which metacognitive processes are corrupted by a pharmacological intervention, a no less promising undertaking could be to focus on the potential pharmacological enhancement of metacognitive abilities. The possibility of such enhancement was

demonstrated by the study of Hauser *et al.* (2017; see Chapter I.2.2), despite the fact that according to the *Theory of conscious states* by Carhart-Harris *et al.* (2014), secondary consciousness is fundamentally optimized towards upholding abstracted, metacognitive states, and the external application of psychoactive (and in particular, psychedelic or dissociative) drugs is generally associated with a shift away from secondary consciousness. However, pharmacological enhancement would rather be expected as a result of administration of psychostimulants such as methylphenidate, which interacts indirectly with dopamine and noradrenaline and has previously been associated with cognitive enhancement in a variety of domains (Devilbiss & Berridge, 2008; Marco *et al.*, 2011; McDonald *et al.*, 2017) and even an increased ability for conscious error detection (Hester *et al.*, 2012). Stimulation of the vagus nerve, which exerts a modulatory effect on neurotransmitter systems (Villani *et al.*, 2019), might represent yet another approach for enhancement of metacognitive accuracy by external applications. Similarly, while evidence regarding the malleability of metacognitive ability and the transfer of training effects across domains is still inconsistent (Carpenter *et al.*, 2019; Haddara & Rahnev, 2022; Rouy *et al.*, 2022), future studies could include assessing how potential training effects are mediated at the neurotransmitter level.

Beyond brain functional and neuropharmacological approaches, consideration of the molecular genetic basis of metacognitive processes is elemental to understanding the biological substructures of metacognition. To this end, it is arguably constructive to undertake research efforts to identify genotype variations that may account for individual differences in metacognition. Finally, the wider research program in which the studies included in the present thesis are embedded also encompasses investigations on the integrity of metacognitive monitoring in schizotypy, a subclinical endophenotype of schizophrenia harboring risk for acute psychosis (Barrantes-Vidal *et al.*, 2015; Thomas *et al.*, 2021), and on the effects of a subanesthetic dose of ketamine on highly schizotypal individuals.

IV.4 CONCLUSION

The studies included in this dissertation contribute towards a fundamental account of the structure and taxonomy of metacognitive operations across different domains and requirements, as well as towards the emergent understanding of the biological underpinnings underlying the capacity for metacognitive monitoring at the neurotransmitter level. Taken together, a combination of generalized (unity) and functionally specified (diversity) components appears to be adequate for both: Study 1 suggested the coexistence of a fairly unified metacognitive subsystem for evaluating "cognitive" (i.e., memory and attention-to-action) judgments and a modular subsystem for metaperceptual reports, whereas conceptual and methodological constraints forestall the linking of estimates of metacognitive ability across method-groups. On the other hand, Studies 2 and 3 shed light on a potentially shared substrate at the neurotransmitter level for metamemory and metaperception, as challenges to the glutamatergic system appear to disrupt a common early confidence signal which may partially be compensated for by a process re-representing object-level information to the metacognitive confidence rating process. With regard to the moderate effect sizes, however, the contribution of additional neurotransmitter systems to the integrity of metacognitive processes certainly needs to be assumed, consistent with a unity and diversity account of the pharmacological substrates of metacognition. Overall, it appears only adequate to assume a construct as complex as metacognition to be governed by a complex interplay of neurobiological underpinnings and partially overlapping subsystems. At this stage, the architecture of metacognition remains but an imperfect model, and only future research may allow us to unravel its full edifice.

V. BIBLIOGRAPHY

Abdallah, C. G., De Feyter, H. M., Averill, L. A., Jiang, L., Averill, C. L., Chowdhury, G. M. I., Purohit, P., de Graaf, R. A., Esterlis, I., Juchem, C., Pittman, B. P., Krystal, J. H., Rothman, D. L., Sanacora, G., & Mason, G. F. (2018). The effects of ketamine on prefrontal glutamate neurotransmission in healthy and depressed subjects. *Neuropsychopharmacology*, *43*(10), 2154–2160. https://doi.org/10.1038/s41386-018-0136-3

Adam, K. C. S., Doss, M. K., Pabon, E., Vogel, E. K., & de Wit, H. (2020). Δ9-Tetrahydrocannabinol (THC) impairs visual working memory performance: a randomized crossover trial. *Neuropsychopharmacology*, *45*(11), 1807–1816. https://doi.org/10.1038/s41386-020-0690-3

Adell, A., Jiménez-Sánchez, L., López-Gil, X., & Romón, T. (2012). Is the Acute NMDA Receptor Hypofunction a Valid Model of Schizophrenia? *Schizophrenia Bulletin*, *38*(1), 9–14. https://doi.org/10.1093/schbul/sbr133

Ais, J., Zylberberg, A., Barttfeld, P., & Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition*, *146*, 377–386. https://doi.org/10.1016/j.cognition.2015.10.006

Akinfiresoye, L., & Tizabi, Y. (2013). Antidepressant effects of AMPA and ketamine combination: role of hippocampal BDNF, synapsin, and mTOR. *Psychopharmacology*, *230*(2), 291–298. https://doi.org/10.1007/s00213-013-3153-2

Akturk, A. O., & Sahin, I. (2011). Literature Review on Metacognition and its Measurement. *Procedia - Social and Behavioral Sciences*, *15*, 3731–3736. https://doi.org/10.1016/j.sbspro.2011.04.364

Aleksandrova, L. R., Phillips, A. G., & Wang, Y. T. (2017). Antidepressant effects of ketamine and the roles of AMPA glutamate receptors and other mechanisms beyond NMDA receptor antagonism. *Journal of Psychiatry & Neuroscience*, *42*(4), 222–229. https://doi.org/10.1503/jpn.160175

Alfimova, M. V., Golimbet, V., Lezheiko, T., & Korovaitseva, G. (2017). A Potential Role of the 5-HTTLPR Polymorphism in Self-Reported Executive Functioning. *Spanish Journal of Psychology*, *20*(2017), 1–9. https://doi.org/10.1017/sjp.2017.6

Allen, M., Frank, D., Schwarzkopf, D. S., Fardo, F., Winston, J. S., Hauser, T. U., & Rees, G. (2016). Unexpected arousal modulates the influence of sensory noise on confidence. *ELife*, *5*(e18103), 1–17. https://doi.org/10.7554/eLife.18103

Allen, M., Glen, J. C., Müllensiefen, D., Schwarzkopf, D. S., Fardo, F., Frank, D., Callaghan, M. F., & Rees, G. (2017). Metacognitive ability correlates with hippocampal and prefrontal microstructure. *NeuroImage*, *149*, 415–423. https://doi.org/10.1016/j.neuroimage.2017.02.008

Andreou, C., Moritz, S., Veith, K., Veckenstedt, R., & Naber, D. (2014). Dopaminergic Modulation of Probabilistic Reasoning and Overconfidence in Errors: A Double-Blind Study. *Schizophrenia Bulletin*, *40*(3), 558–565. https://doi.org/10.1093/schbul/sbt064

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *The Neuroscientist*, *18*(3), 251–270. https://doi.org/10.1177/1073858411403316

Anis, N. A., Berry, S. C., Burton, N. R., & Lodge, D. (1983). The dissociative anaesthetics, ketamine and phencyclidine, selectively reduce excitation of central mammalian neurones by N-methyl-aspartate. *British Journal of Pharmacology*, *79*(2), 565–575. https://doi.org/10.1111/j.1476-5381.1983.tb11031.x

Arbuzova, P., Maurer, L. K., & Filevich, E. (2022). Metacognitive domains are not aligned along a

dimension of internal-external information source. *BioRxiv*, 2022.05.03.490468. https://doi.org/10.1101/2022.05.03.490468

Arbuzova, P., Peters, C., Röd, L., Koß, C., Maurer, H., Maurer, L. K., Müller, H., Verrel, J., & Filevich, E. (2021). Measuring metacognition of direct and indirect parameters of voluntary movement. *Journal of Experimental Psychology: General*, *150*(11), 2208–2229. https://doi.org/10.1037/xge0000892

Artigas, F., Nutt, D. J., & Shelton, R. (2002). Mechanism of action of antidepressants. *Psychopharmacology Bulletin*, *36 Suppl 2*, 123–132.

Asztély, F., & Gustafsson, B. (1996). Ionotropic glutamate receptors. *Molecular Neurobiology*, *12*(1), 1. https://doi.org/10.1007/BF02740744

Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research*, *150*, 45–53. https://doi.org/10.1016/S0079-6123(05)50004-9

Babiloni, C., Pizzella, V., Gratta, C. Del, Ferretti, A., & Romani, G. L. (2009). Fundamentals of Electroencefalography, Magnetoencefalography, and Functional Magnetic Resonance Imaging. *International Review of Neurobiology*, *86*, 67–80. https://doi.org/10.1016/S0074-7742(09)86005-4

Bacon, E., Danion, J. M., Kauffmann-Muller, F., Schelstraete, M. A., Bruant, A., Sellal, F., & Grange, D. (1998). Confidence level and feeling of knowing for episodic and semantic memory: An investigation of lorazepam effects on metamemory. *Psychopharmacology*, *138*(3–4), 318–325. https://doi.org/10.1007/s002130050677

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, *10*(9), 659–669. https://doi.org/10.1038/nrn2667

Baer, C., Ghetti, S., & Odic, D. (2021). Perceptual and Memory Metacognition in Children. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*. https://escholarship.org/uc/item/30g4w4sc

Bahrami, B., Olsen, K., Latham, P., Roepstorff, A., Rees, G., & Frith, C. (2010). Optimally Interacting Minds. *Science*, *329*(5995), 1081–1085. https://doi.org/10.1126/science.1185718

Baird, B., Cieslak, M., Smallwood, J., Grafton, S. T., & Schooler, J. W. (2015). Regional white matter variation associated with domain-specific metacognitive accuracy. *Journal of Cognitive Neuroscience*, *27*(3), 440–452. https://doi.org/10.1162/jocn_a_00741

Baird, B., Mota-Rolim, S. A., & Dresler, M. (2019). The cognitive neuroscience of lucid dreaming. *Neuroscience and Biobehavioral Reviews*, *100*, 305–323. https://doi.org/10.1016/j.neubiorev.2019.03.008

Baird, B., Mrazek, M. D., Phillips, D. T., & Schooler, J. W. (2014). Domain-specific enhancement of metacognitive ability following meditation training. *Journal of Experimental Psychology: General*, *143*(5), 1972–1979. https://doi.org/10.1037/a0036882

Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *The Journal of Neuroscience*, *33*(42), 16657–16665. https://doi.org/10.1523/jneurosci.0786-13.2013

Baker, L., & Cerro, L. (2000). Assessing Metacognition in Children and Adults. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 99–145). Buros Institute of Mental Measurements.

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

Bang, D., & Fleming, S. M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, *115*(23), 6082–6087. https://doi.org/10.1073/pnas.1800795115

Barrantes-Vidal, N., Grant, P., & Kwapil, T. R. (2015). The Role of Schizotypy in the Study of the Etiology of Schizophrenia Spectrum Disorders. *Schizophrenia Bulletin*, *41*(suppl_2), S408–S416. https://doi.org/10.1093/schbul/sbu191

Barrett, A. B., Dienes, Z., & Seth, A. K. (2013). Measures of metacognition on signal-detection theoretic models. *Psychological Methods*, *18*(4), 535. https://doi.org/10.1037/a0033268

Barth, M., & Poser, B. A. (2011). Advances in High-Field BOLD fMRI. *Materials*, *4*(11). https://doi.org/10.3390/ma4111941

Bastin, J., Deman, P., David, O., Gueguen, M., Benis, D., Minotti, L., Hoffman, D., Combrisson, E., Kujala, J., Perrone-Bertolotti, M., Kahane, P., Lachaux, J.-P., & Jerbi, K. (2017). Direct Recordings from Human Anterior Insula Reveal its Leading Role within the Error-Monitoring Network. *Cerebral Cortex*, *27*(2), 1545–1557. https://doi.org/10.1093/cercor/bhv352

Beck, A. T., Baruch, E., Balter, J. M., Steer, R. A., & Warman, D. M. (2004). A new instrument for measuring insight: The Beck Cognitive Insight Scale. *Schizophrenia Research*, *68*(2–3), 319–329. https://doi.org/10.1016/S0920-9964(03)00189-0

Beecher, H. K. (1955). The Powerful Placebo. *Journal of the American Medical Association*, *159*(17), 1602–1606. https://doi.org/10.1001/jama.1955.02960340022006

Berman, R. M., Cappiello, A., Anand, A., Oren, D. A., Heninger, G. R., Charney, D. S., & Krystal, J. H. (2000). Antidepressant effects of ketamine in depressed patients. *Biological Psychiatry*, *47*(4), 351–354. https://doi.org/10.1016/S0006-3223(99)00230-9

Besiroglu, L., Çetinkaya, N., Selvi, Y., & Atli, A. (2011). Effects of selective serotonin reuptake inhibitors on thought-action fusion, metacognitions, and thought suppression in obsessive-compulsive disorder. *Comprehensive Psychiatry*, *52*(5), 556–561. https://doi.org/10.1016/j.comppsych.2010.10.003

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1341–1354. https://doi.org/10.1037/0278-7393.32.6.1341

Bickel, S., & Javitt, D. C. (2009). Neurophysiological and neurochemical animal models of schizophrenia: Focus on glutamate. *Behavioural Brain Research*, *204*(2), 352–362. https://doi.org/10.1016/j.bbr.2009.05.005

Bittar, T. P., & Labonté, B. (2021). Functional Contribution of the Medial Prefrontal Circuitry in Major Depressive Disorder and Stress-Induced Depressive-Like Behaviors. *Frontiers in Behavioral Neuroscience*, *15*. https://doi.org/10.3389/fnbeh.2021.699592

Boldt, A., De Gardelle, V., & Yeung, N. (2017). The impact of evidence reliability on sensitivity and bias in decision confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1520. https://doi.org/10.1037/xhp0000404

Boldt, A., & Gilbert, S. J. (2022). Partially Overlapping Neural Correlates of Metacognitive Monitoring and Metacognitive Control. *The Journal of Neuroscience*, *42*(17), 3622–3635. https://doi.org/10.1523/jneurosci.1326-21.2022

Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *The Journal of Neuroscience*, *35*(8), 3478–3484. https://doi.org/10.1523/jneurosci.0797-14.2015

Bowdle, A. T., Radant, A. D., Cowley, D. S., Kharasch, E. D., Strassman, R. J., & Roy-Byrne, P. P. (1998). Psychedelic Effects of Ketamine in Healthy Volunteers : Relationship to Steady-state Plasma Concentrations. *Anesthesiology*, *88*(1), 82–88. https://doi.org/10.1097/00000542-199801000-00015

Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology*, *103*(4), 751–757. https://doi.org/10.1037/h0037190

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, *21*(1), 1–16. https://doi.org/10.1111/j.2044-8260.1982.tb01421.x

Bröcker, A.-L., Bayer, S., Stuke, F., Giemsa, P., Heinz, A., Bermpohl, F., Lysaker, P. H., & Montag, C. (2017). The Metacognition Assessment Scale (MAS-A): Results of a pilot study applying a German translation to individuals with schizophrenia spectrum disorders. *Psychology and Psychotherapy: Theory, Research and Practice*, *90*(3), 401–418. https://doi.org/10.1111/papt.12122

Brogaard, B. (2011). Are there unconscious perceptual processes? *Consciousness and Cognition*, *20*(2), 449–463. https://doi.org/10.1016/j.concog.2010.10.002

Brown, C., Fezoui, M., Selig, W. M., Schwartz, C. E., & Ellis, J. L. (2004). Antitussive activity of sigma-1 receptor agonists in the guinea-pig. *British Journal of Pharmacology*, *141*(2), 233–240. https://doi.org/10.1038/sj.bjp.0705605

Brown, G. G., Perthen, J. E., Liu, T. T., & Buxton, R. B. (2007). A Primer on Functional Magnetic Resonance Imaging. *Neuropsychology Review*, *17*(2), 107–125. https://doi.org/10.1007/s11065-007-9028-8

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361–379). Guilford press.

Büchel, C. (2004). Perspectives on the estimation of effective connectivity from neuroimaging data. *Neuroinformatics*, *2*(2), 169–173. https://doi.org/10.1385/NI:2:2:169

Burgess, P. (1997). Theory and Methodology in Executive Function Research. In P. Rabbitt (Ed.), *Theory and Methodology of Frontal and Executive Function.* (pp. 81–116). Psychology Press.

Burns, J. M., & Boyer, E. W. (2013). Antitussives and substance abuse. *Substance Abuse and Rehabilitation*, *4*, 75–82. https://doi.org/10.2147/sar.s36761

Busey, T. A., & Arici, A. (2009). On the Role of Individual Items in Recognition Memory and Metacognition: Challenges for Signal Detection Theory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(5), 1123–1136. https://doi.org/10.1037/a0016646

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin and Review*, *7*(1), 26–48. https://doi.org/10.3758/BF03210724

Canli, T., & Lesch, K.-P. (2007). Long story short: the serotonin transporter in emotion regulation and social cognition. *Nature Neuroscience*, *10*(9), 1103–1109. https://doi.org/10.1038/nn1964

Carhart-Harris, R. L., Leech, R., Hellyer, P. J., Shanahan, M., Feilding, A., Tagliazucchi, E., Chialvo, D. R., & Nutt, D. (2014). The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in Human Neuroscience*, *8*(20), 1–22. https://doi.org/10.3389/fnhum.2014.00020

Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, *148*(1), 51–64. https://doi.org/10.1037/xge0000505

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, *32*(2), 121–138. https://doi.org/10.1017/S0140525X09000545

Carter, L. P., Kleykamp, B. A., Griffiths, R. R., & Mintzer, M. Z. (2013). Cognitive effects of intramuscular ketamine and oral triazolam in healthy volunteers. *Psychopharmacology*, *226*(1), 53–63. https://doi.org/10.1007/s00213-012-2883-x

Carter, L. P., Reissig, C. J., Johnson, M. W., Klinedinst, M. A., Griffiths, R. R., & Mintzer, M. Z. (2013). Acute cognitive effects of high doses of dextromethorphan relative to triazolam in humans. *Drug and Alcohol Dependence*, *128*(3), 206–213. https://doi.org/10.1016/j.drugalcdep.2012.08.025

Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about worry and intrusions: The Meta-Cognitions Questionnaire and its correlates. *Journal of Anxiety Disorders*, *11*(3), 279–296. https://doi.org/10.1016/s0887-6185(97)00011-x

Charles, L., Gaillard, R., Amado, I., Krebs, M.-O., Bendjemaa, N., & Dehaene, S. (2017). Conscious and unconscious performance monitoring: Evidence from patients with schizophrenia. *NeuroImage*, *144*, 153–163. https://doi.org/10.1016/j.neuroimage.2016.09.056

Chen, H.-S. V., & Lipton, S. A. (2006). The chemical biology of clinically tolerated NMDA receptor antagonists. *Journal of Neurochemistry*, *97*(6), 1611–1626. https://doi.org/10.1111/j.1471-4159.2006.03991.x

Chen, Y.-C. I., Choi, J.-K., Andersen, S. L., Rosen, B. R., & Jenkins, B. G. (2005). Mapping dopamine D2/D3 receptor function using pharmacological magnetic resonance imaging. *Psychopharmacology*, *180*(4), 705–715. https://doi.org/10.1007/s00213-004-2034-0

Chua, E. F., Schacter, D. L., Rand-Giovannetti, E., & Sperling, R. A. (2006). Understanding metamemory: Neural correlates of the cognitive process and subjective level of confidence in recognition memory. *NeuroImage*, *29*(4), 1150–1160. https://doi.org/10.1016/j.neuroimage.2005.09.058

Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural correlates of metamemory: A comparison of feeling-of-knowing and retrospective confidence judgments. *Journal of Cognitive Neuroscience*, *21*(9), 1751–1765. https://doi.org/10.1162/jocn.2009.21123

Clark, A., Friston, K., & Wilkinson, S. (2019). Bayesing qualia: Consciousness as inference, not raw datum. *Journal of Consciousness Studies*, *26*(9–10), 19–33.

Clos, M., Bunzeck, N., & Sommer, T. (2019a). Dopamine is a double-edged sword: dopaminergic modulation enhances memory retrieval performance but impairs metacognition. *Neuropsychopharmacology*, *44*(3), 555–563. https://doi.org/10.1038/s41386-018-0246-y

Clos, M., Bunzeck, N., & Sommer, T. (2019b). Dopamine Enhances Item Novelty Detection via Hippocampal and Associative Recall via Left Lateral Prefrontal Cortex Mechanisms. *The Journal of Neuroscience*, *39*(40), 7920–7933. https://doi.org/10.1523/jneurosci.0495-19.2019

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, *1*(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Colagiuri, B., Schenk, L. A., Kessler, M. D., Dorsey, S. G., & Colloca, L. (2015). The placebo effect: From concepts to genes. *Neuroscience*, *307*, 171–190. https://doi.org/10.1016/j.neuroscience.2015.08.017

Collingridge, G. L. (1987). The role of NMDA receptors in learning and memory. *Nature*, *330*(6149), 604–605. https://doi.org/10.1038/330604a0

Collingridge, G. L., Volianskis, A., Bannister, N., France, G., Hanna, L., Mercier, M., Tidball, P., Fang, G., Irvine, M. W., Costa, B. M., Monaghan, D. T., Bortolotto, Z. A., Molnár, E., Lodge, D., & Jane, D. E.

(2013). The NMDA receptor as a target for cognitive enhancement. *Neuropharmacology*, *64*, 13–26. https://doi.org/10.1016/j.neuropharm.2012.06.051

Condy, C., Wattiez, N., Rivaud-Péchoux, S., & Gaymard, B. (2005). Ketamine-induced distractibility: An oculomotor study in monkeys. *Biological Psychiatry*, *57*(4), 366–372. https://doi.org/10.1016/j.biopsych.2004.10.036

Conn, P. J., & Pin, J. P. (1997). Pharmacology and functions of metabotropic glutamate receptors. *Annual Review of Pharmacology and Toxicology*, *37*, 205–237. https://doi.org/10.1146/annurev.pharmtox.37.1.205

Corssen, G., & Domino, E. F. (1966). Dissociative Anesthesia: Further Pharmacologic Studies and First Clinical Experience with the Phencyclidine Derivative Cl-581. *Anesthesia and Analgesia*, *45*(1), 29–40.

Coull, J. T., Morgan, H., Cambridge, V. C., Moore, J. W., Giorlando, F., Adapa, R., Corlett, P. R., & Fletcher, P. C. (2011). Ketamine perturbs perception of the flow of time in healthy volunteers. *Psychopharmacology*, *218*(3), 543–556. https://doi.org/10.1007/s00213-011-2346-9

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Crick, F., & Koch, C. (1990). Toward a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275.

Cummings, K. A., & Popescu, G. K. (2015). Glycine-dependent activation of NMDA receptors. *Journal of General Physiology*, *145*(6), 513–527. https://doi.org/10.1085/jgp.201411302

D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Balteau, E., Luxen, A., Maquet, P., & Salmon, E. (2007). Distinct Regions of the Medial Prefrontal Cortex Are Associated with Self-referential Processing and Perspective Taking. *Journal of Cognitive Neuroscience*, *19*(6), 935–944. https://doi.org/10.1162/jocn.2007.19.6.935

Dang, J., King, K. M., & Inzlicht, M. (2020). Why Are Self-Report and Behavioral Measures Weakly Correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

David, A. S., Bedford, N., Wiffen, B., & Gilleen, J. (2012). Failures of metacognition and lack of insight in neuropsychiatric disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1379–1390. https://doi.org/10.1098/rstb.2012.0002

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, *17*(4), 335–350. https://doi.org/10.1080/09548980601004024

de Gardelle, V., & Mamassian, P. (2014). Does Confidence Use a Common Currency Across Two Visual Tasks? *Psychological Science*, *25*(6), 1286–1288. https://doi.org/10.1177/0956797614528956

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*, *132*(9), 2531–2540. https://doi.org/10.1093/brain/awp111

Desender, K., Buc Calderon, C., Van Opstal, F., & Van den Bussche, E. (2017). Avoiding the conflict: Metacognitive awareness drives the selection of low-demand contexts. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(7), 1397–1410. https://doi.org/10.1037/xhp0000391

Desender, K., Donner, T. H., & Verguts, T. (2021). Dynamic expressions of confidence within an

evidence accumulation framework. *Cognition*, *207*, 104522. https://doi.org/10.1016/j.cognition.2020.104522

Desender, K., Vermeylen, L., & Verguts, T. (2022). Dynamic influences on static measures of metacognition. *Nature Communications*, *13*(1), 4208. https://doi.org/10.1038/s41467-022-31727-0

Devilbiss, D. M., & Berridge, C. W. (2008). Cognition-Enhancing Doses of Methylphenidate Preferentially Increase Prefrontal Cortex Neuronal Responsiveness. *Biological Psychiatry*, *64*(7), 626–635. https://doi.org/10.1016/j.biopsych.2008.04.037

Di Gregorio, F., Maier, M. E., & Steinhauser, M. (2020). Are errors detected before they occur? Early error sensations revealed by metacognitive judgments on the timing of error awareness. *Consciousness and Cognition*, *77*, 102857. https://doi.org/10.1016/j.concog.2019.102857

Dienes, Z., & Seth, A. (2010). Gambling on the unconscious: A comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Consciousness and Cognition*, *19*(2), 674–681. https://doi.org/10.1016/j.concog.2009.09.009

Dimaggio, G., Maillard, P., MacBeth, A., & Kramer, U. (2019). Effects of Therapeutic Alliance and Metacognition on Outcome in a Brief Psychological Treatment for Borderline Personality Disorder. *Psychiatry*, *82*(2), 143–157. https://doi.org/10.1080/00332747.2019.1610295

Dingledine, R., Borges, K., Bowie, D., & Traynelis, S. F. (1999). The glutamate receptor ion channels. *Pharmacological Reviews*, *51*(1), 7–61.

Dittrich, A. (1998). The standardized psychometric assessment of altered states of consciousness (ASCs) in humans. *Pharmacopsychiatry*, *31*(SUPPL. 2), 80–84. https://doi.org/10.1055/s-2007-979351

Dixon, W. J., & Mood, A. M. (1948). A Method for Obtaining and Analyzing Sensitivity Data. *Journal of the American Statistical Association*, *43*(241), 109–126. https://doi.org/10.1080/01621459.1948.10483254

Doss, S., Wandinger, K.-P., Hyman, B. T., Panzer, J. A., Synofzik, M., Dickerson, B., Mollenhauer, B., Scherzer, C. R., Ivinson, A. J., Finke, C., Schöls, L., Müller vom Hagen, J., Trenkwalder, C., Jahn, H., Höltje, M., Biswal, B. B., Harms, L., Ruprecht, K., Buchert, R., … Prüss, H. (2014). High prevalence of NMDA receptor IgA/IgM antibodies in different dementia types. *Annals of Clinical and Translational Neurology*, *1*(10), 822–832. https://doi.org/10.1002/acn3.120

Double, K. S., & Birney, D. P. (2019). Reactivity to Measures of Metacognition. *Frontiers in Psychology*, *10*(2755). https://doi.org/10.3389/fpsyg.2019.02755

Drueke, B., Gauggel, S., Weise, L., Forkmann, T., & Mainz, V. (2022). Metacognitive judgements and abilities in patients with affective disorders. *Current Psychology*. https://doi.org/10.1007/s12144-022-02838-0

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). Springer. https://doi.org/10.1007/978-1-84628-609-4

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

Efklides, A., & Vlachopoulos, S. P. (2012). Measurement of Metacognitive Knowledge of Self, Task, and Strategies in Mathematics. *European Journal of Psychological Assessment*, *28*(3), 227–239. https://doi.org/10.1027/1015-5759/a000145

Eichner, C., & Berna, F. (2016). Acceptance and Efficacy of Metacognitive Training (MCT) on Positive Symptoms and Delusions in Patients With Schizophrenia: A Meta-analysis Taking Into Account Important Moderators. *Schizophrenia Bulletin*, *42*(4), 952–962.

https://doi.org/10.1093/schbul/sbv225

Eisenacher, S., Rausch, F., Ainser, F., Mier, D., Veckenstedt, R., Schirmbeck, F., Lewien, A., Englisch, S., Andreou, C., Moritz, S., Meyer-Lindenberg, A., Kirsch, P., & Zink, M. (2015). Investigation of metamemory functioning in the at-risk mental state for psychosis. *Psychological Medicine*, *45*(15), 3329–3340. https://doi.org/10.1017/S0033291715001373

Eisenacher, S., & Zink, M. (2017). The Importance of Metamemory Functioning to the Pathogenesis of Psychosis. *Frontiers in Psychology*, *8*(304), 1–15. https://doi.org/10.3389/fpsyg.2017.00304

Elman, J. A., Klostermann, E. C., Marian, D. E., Verstaen, A., & Shimamura, A. P. (2012). Neural correlates of metacognitive monitoring during episodic and semantic retrieval. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 599–609. https://doi.org/10.3758/s13415-012-0096-8

Enders, C. K., & Bandalos, D. L. (2001). The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Evans, S., & Azzopardi, P. (2007). Evaluation of a "bias-free" measure of awareness. *Spatial Vision*, *20*(1–2), 61–77. https://doi.org/10.1163/156856807779369742

Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *The Journal of Neuroscience*, *38*(2), 263–277. https://doi.org/10.1523/jneurosci.0322-17.2017

Faßbender, K., Bey, K., Lippold, J. V, Aslan, B., Hurlemann, R., & Ettinger, U. (2021). GABAergic modulation of performance in response inhibition and interference control tasks. *Journal of Psychopharmacology*, *35*(12), 1496–1509. https://doi.org/10.1177/02698811211032440

Fitzgerald, L. M., Arvaneh, M., & Dockree, P. M. (2017). Domain-specific and domain-general processes underlying metacognitive judgments. *Consciousness and Cognition*, *49*, 264–277. https://doi.org/10.1016/j.concog.2017.01.011

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906.

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 21–29). Lawrence Erlbaum Associates.

Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex*, *16*(11), 1623–1630. https://doi.org/10.1093/cercor/bhj097

Fleming, S. M. (2017). HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *2017*(1), 1–14. https://doi.org/10.1093/nc/nix007

Fleming, S. M. (2020). Awareness as inference in a higher-order state space. *Neuroscience of Consciousness*, *2020*(1). https://doi.org/10.1093/nc/niz020

Fleming, S. M. (2021). *Know Thyself - The science of self-awareness* (1st ed.). Basic Books.

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general bayesian framework for metacognitive computation. *Psychological Review*, *124*(1), 91–114. https://doi.org/10.1037/rev0000045

Fleming, S. M., & Dolan, R. J. (2010). Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Consciousness and Cognition*, *19*(1), 352–363. https://doi.org/10.1016/j.concog.2009.11.002

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

Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1280–1286. https://doi.org/10.1098/rstb.2012.0021

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience*, *32*(18), 6117–6125. https://doi.org/10.1523/jneurosci.6489-11.2012

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*(443), 1–9. https://doi.org/10.3389/fnhum.2014.00443

Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J.-C. (2016). Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness*, *2016*(1), 1–12. https://doi.org/10.1093/nc/niw018

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain*, *137*(10), 2811–2822. https://doi.org/10.1093/brain/awu221

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541–1543. https://doi.org/10.1126/science.1191883

Fletcher, L., & Carruthers, P. (2012). Metacognition and reasoning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1366–1378. https://doi.org/10.1098/rstb.2011.0413

Ford, C. P. (2014). The role of D2-autoreceptors in regulating dopamine neuron activity and transmission. *Neuroscience*, *282*, 13–22. https://doi.org/10.1016/j.neuroscience.2014.01.025

Foster, A. C., & Fagg, G. E. (1987). Taking apart NMDA receptors. *Nature*, *329*(6138), 395–396. https://doi.org/10.1038/329395a0

Fox, P. T., & Friston, K. J. (2012). Distributed processing; distributed functions? *NeuroImage*, *61*(2), 407–426. https://doi.org/10.1016/j.neuroimage.2011.12.051

Fox, P. T., & Raichle, M. E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proceedings of the National Academy of Sciences*, *83*(4), 1140–1144. https://doi.org/10.1073/pnas.83.4.1140

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, *6*(3), 218–229. https://doi.org/10.1006/nimg.1997.0291

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-Related fMRI: Characterizing Differential Responses. *NeuroImage*, *7*(1), 30–40. https://doi.org/10.1006/nimg.1997.0306

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2213–2223.

https://doi.org/10.1098/rstb.2012.0123

Frith, C. D., & Dolan, R. (1996). The role of the prefrontal cortex in higher cognitive functions. *Cognitive Brain Research*, *5*(1), 175–181. https://doi.org/10.1016/S0926-6410(96)00054-7

Frohlich, J., & Van Horn, J. D. (2014). Reviewing the ketamine model for schizophrenia. *Journal of Psychopharmacology*, *28*(4), 287–302. https://doi.org/10.1177/0269881113512909

Fujikawa, D. G. (2015). The Role of Excitotoxic Programmed Necrosis in Acute Brain Injury. *Computational and Structural Biotechnology Journal*, *13*, 212–221. https://doi.org/10.1016/j.csbj.2015.03.004

Gaddum, J. H. (1954). Clinical pharmacology. *Proceedings of the Royal Society of Medicine*, *47*(3), 195–204. https://pubmed.ncbi.nlm.nih.gov/13155508

Gajdos, T., Fleming, S. M., Saez Garcia, M., Weindel, G., & Davranche, K. (2019). Revealing subthreshold motor contributions to perceptual confidence. *Neuroscience of Consciousness*, *2019*(1), 1–8. https://doi.org/10.1093/nc/niz001

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin and Review*, *10*(4), 843–876. https://doi.org/10.3758/BF03196546

Garris, P. A., Budygin, E. A., Phillips, P. E. M., Venton, B. J., Robinson, D. L., Bergstrom, B. P., Rebec, G. V, & Wightman, R. M. (2003). A role for presynaptic mechanisms in the actions of nomifensine and haloperidol. *Neuroscience*, *118*(3), 819–829. https://doi.org/10.1016/S0306-4522(03)00005-8

Gatignon, H. (2010). Confirmatory factor analysis. In *Statistical analysis of management data* (pp. 59–122). Springer. https://doi.org/10.1007/978-1-4419-1270-1_4

Geurten, M., Meulemans, T., & Lemaire, P. (2018). From domain-specific to domain-general? The developmental path of metacognition for strategy selection. *Cognitive Development*, *48*, 62–81. https://doi.org/10.1016/j.cogdev.2018.08.002

Gherman, S., & Philiastides, M. G. (2018). Human VMPFC encodes early signatures of confidence in perceptual decisions. *ELife*, *7*, e38293. https://doi.org/10.7554/eLife.38293

Gitlin, J., Chamadia, S., Locascio, J. J., Ethridge, B. R., Pedemonte, J. C., Hahm, E. Y., Ibala, R., Mekonnen, J., Colon, K. M., Qu, J., & Akeju, O. (2020). Dissociative and Analgesic Properties of Ketamine Are Independent. *Anesthesiology*, *133*(5), 1021–1028. https://doi.org/10.1097/ALN.0000000000003529

Goodman, L. A., & Kruskal, W. H. (1954). Measures of Association for Cross Classifications. *Journal of the American Statistical Association*, *49*(268), 732–764. https://doi.org/10.2307/2281536

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence Predicts Health and Longevity, but Why? *Current Directions in Psychological Science*, *13*(1), 1–4. https://doi.org/10.1111/j.0963-7214.2004.01301001.x

Grant, A. M., Franklin, J., & Langford, P. (2002). The self-reflection and insight scale: A new measure of private self-consciousness. *Social Behavior and Personality: An International Journal*, *30*(8), 821–835. https://doi.org/10.2224/sbp.2002.30.8.821

Green, D. G., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley & Sons, Inc.

Grigsby, J., & Schneiders, J. L. (1991). Neuroscience, Modularity and Personality Theory: Conceptual Foundations of a Model of Complex Human Functioning. *Psychiatry*, *54*(1), 21. https://doi.org/10.1080/00332747.1991.11024528

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 and Biobehavioral Reviews*, *55*, 88–97. https://doi.org/10.1016/j.neubiorev.2015.04.006

Grossner, E. C., Bernier, R. A., Brenner, E. K., Chiou, K. S., & Hillary, F. G. (2018). Prefrontal gray matter volume predicts metacognitive accuracy following traumatic brain injury. *Neuropsychology*, *32*(4), 484–494. https://doi.org/10.1037/neu0000446

Grossner, E. C., Brenner, E., Bernier, R., Rabinowitz, A., & Hillary, F. (2021). Metacognition in Individuals Aging with Traumatic Brain Injury. *PsyArXiv*. https://doi.org/10.31234/osf.io/967zj

Guensch, D. P., Michel, M. C., Huettenmoser, S. P., Jung, B., Gulac, P., Segiser, A., Longnus, S. L., & Fischer, K. (2021). The blood oxygen level dependent (BOLD) effect of in-vitro myoglobin and hemoglobin. *Scientific Reports*, *11*(1), 11464. https://doi.org/10.1038/s41598-021-90908-x

Guggenmos, M. (2021). Measuring metacognitive performance: type 1 performance dependence and test-retest reliability. *Neuroscience of Consciousness*, *2021*(1), 1–14. https://doi.org/10.1093/nc/niab040

Guilford, J. P., & Perry, N. C. (1951). Estimation of other coefficients of correlation from the phi coefficient. *Psychometrika*, *16*(3), 335–346. https://doi.org/10.1007/BF02310556

Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural Correlates of Auditory Perceptual Awareness under Informational Masking. *PLoS Biology*, *6*(6), e138. https://doi.org/10.1371/journal. pbio.0060138

Haddara, N., & Rahnev, D. (2022). The Impact of Feedback on Perceptual Decision-Making and Metacognition: Reduction in Bias but No Change in Sensitivity. *Psychological Science*, *33*(2), 259–275. https://doi.org/10.1177/09567976211032887

Hamm, J. A., Renard, S. B., Fogley, R. L., Leonhardt, B. L., Dimaggio, G., Buck, K. D., & Lysaker, P. H. (2012). Metacognition and Social Cognition in Schizophrenia: Stability and Relationship to Concurrent and Prospective Symptom Assessments. *Journal of Clinical Psychology*, *68*(12), 1303–1312. https://doi.org/10.1002/jclp.21906

Hardingham, G. E., & Bading, H. (2010). Synaptic versus extrasynaptic NMDA receptor signalling: implications for neurodegenerative disorders. *Nature Reviews Neuroscience*, *11*(10), 682–696. https://doi.org/10.1038/nrn2911

Harty, S., O'Connell, R. G., Hester, R., & Robertson, I. H. (2013). Older adults have diminished awareness of errors in the laboratory and daily life. *Psychology and Aging*, *28*(4), 1032–1041. https://doi.org/10.1037/a0033567

Hauser, T. U., Allen, M., Purg, N., Moutoussis, M., Rees, G., & Dolan, R. J. (2017). Noradrenaline blockade specifically enhances metacognitive performance. *ELife*, *6*, 1–13. https://doi.org/10.7554/eLife.24901

Hertzog, C., & Robinson, A. E. (2005). Metacognition and Intelligence. In *Handbook of understanding and measuring intelligence* (pp. 101–123). Sage Publications, Inc. https://doi.org/10.4135/9781452233529.n7

Hester, R., Foxe, J. J., Molholm, S., Shpaner, M., & Garavan, H. (2005). Neural mechanisms involved in error processing: A comparison of errors made with and without awareness. *NeuroImage*, *27*(3), 602–608. https://doi.org/10.1016/j.neuroimage.2005.04.035

Hester, R., Sanjay Nandam, L., O'Connell, R. G., Wagner, J., Strudwick, M., Nathan, P. J., Mattingley, J. B., & Bellgrove, M. A. (2012). Neurochemical enhancement of conscious error awareness. *The Journal of Neuroscience*, *32*(8), 2619–2627. https://doi.org/10.1523/jneurosci.4052-11.2012

Heyes, C., Bang, D., Shea, N., Frith, C. D., & Fleming, S. M. (2020). Knowing Ourselves Together: The Cultural Origins of Metacognition. *Trends in Cognitive Sciences*, *24*(5), 349–362. https://doi.org/10.1016/j.tics.2020.02.007

Higham, P. A., Perfect, T. J., & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 57–80. https://doi.org/10.1037/a0013865

Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You'd Better Think Twice: Post-Decision Perceptual Confidence. *NeuroImage*, *99*, 323–331. https://doi.org/10.1016/j.neuroimage.2014.05.049

Honey, G. D., Honey, R. A. E., O'Loughlin, C., Sharar, S. R., Kumaran, D., Suckling, J., Menon, D. K., Sleator, C., Bullmore, E. T., & Fletcher, P. C. (2005). Ketamine disrupts frontal and hippocampal contribution to encoding and retrieval of episodic memory: An fMRI study. *Cerebral Cortex*, *15*(6), 749–759. https://doi.org/10.1093/cercor/bhh176

Honey, G. D., Honey, R. A. E., Sharar, S. R., Turner, D. C., Pomarol-Clotet, E., Kumaran, D., Simons, J. S., Hu, X., Rugg, M. D., Bullmore, E. T., & Fletcher, P. C. (2005). Impairment of specific episodic memory processes by sub-psychotic doses of ketamine: The effects of levels of processing at encoding and of the subsequent retrieval task. *Psychopharmacology*, *181*(3), 445–457. https://doi.org/10.1007/s00213-005-0001-z

Hu, Xiao, Yang, C., & Luo, L. (2022). Are the contributions of processing experience and prior beliefs to confidence ratings domain-general or domain-specific? *Journal of Experimental Psychology: General*, 1–17. https://doi.org/10.1037/xge0001257

Hu, Xiaoping, & Yacoub, E. (2012). The story of the initial dip in fMRI. *NeuroImage*, *62*(2), 1103–1108. https://doi.org/10.1016/j.neuroimage.2012.03.005

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). *Functional magnetic resonance imaging* (3rd ed.). Sinauer Associates.

Humphreys, G. F., & Lambon Ralph, M. A. (2017). Mapping Domain-Selective and Counterpointed Domain-General Higher Cognitive Functions in the Lateral Parietal Cortex: Evidence from fMRI Comparisons of Difficulty-Varying Semantic Versus Visuo-Spatial Tasks, and Functional Connectivity Analyses. *Cerebral Cortex*, *27*(8), 4199–4212. https://doi.org/10.1093/cercor/bhx107

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: A critical review. *Psychophysiology*, *43*(3), 302–313. https://doi.org/10.1111/j.1469-8986.2006.00403.x

Ionescu, D. F., Felicione, J. M., Gosai, A., Cusin, C., Shin, P., Shapero, B. G., & Deckersbach, T. (2018). Ketamine-Associated Brain Changes: A Review of the Neuroimaging Literature. *Harvard Review of Psychiatry*, *26*(6), 320–339. https://doi.org/10.1097/HRP.0000000000000179

Irak, M., & Tosun, A. (2008). Exploring the role of metacognition in obsessive–compulsive and anxiety symptoms. *Journal of Anxiety Disorders*, *22*(8), 1316–1325. https://doi.org/10.1016/j.janxdis.2008.01.012

Irwin, S. A., & Iglewicz, A. (2010). Oral Ketamine for the Rapid Treatment of Depression and Anxiety in Patients Receiving Hospice Care. *Journal of Palliative Medicine*, *13*(7), 903–908. https://doi.org/10.1089/jpm.2010.9808

Izaute, M., & Bacon, E. (2006). Effects of the amnesic drug lorazepam on complete and partial information retrieval and monitoring accuracy. *Psychopharmacology*, *188*(4), 472–481. https://doi.org/10.1007/s00213-006-0492-2

Jang, Y., Lee, H., Kim, Y., & Min, K. (2020). The Relationship between Metacognitive Ability and Metacognitive Accuracy. *Metacognition and Learning*, *15*(3), 411–434. https://doi.org/10.1007/s11409-020-09232-w

Javitt, D. C. (2007). Glutamate and Schizophrenia: Phencyclidine, N-Methyl-d-Aspartate Receptors, and Dopamine–Glutamate Interactions. *International Review of Neurobiology*, *78*, 69–108. https://doi.org/10.1016/S0074-7742(06)78003-5

Javitt, D. C. (2012). Twenty-five Years of Glutamate in Schizophrenia: Are We There Yet? *Schizophrenia Bulletin*, *38*(5), 911–913. https://doi.org/10.1093/schbul/sbs100

Javitt, D. C., Zukin, S. R., Heresco-Levy, U., & Umbricht, D. (2012). Has an angel shown the way? Etiological and therapeutic implications of the PCP/NMDA model of schizophrenia. *Schizophrenia Bulletin*, *38*(5), 958–966. https://doi.org/10.1093/schbul/sbs069

Jenkins, A. C., Macrae, C. N., & Mitchell, J. P. (2008). Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proceedings of the National Academy of Sciences*, *105*(11), 4507–4512. https://doi.org/10.1073/pnas.0708785105

Jensen, J. S., Bielefeldt, A. Ø., & Hróbjartsson, A. (2017). Active placebo control groups of pharmacological interventions were rarely used but merited serious consideration: a methodological overview. *Journal of Clinical Epidemiology*, *87*, 35–46. https://doi.org/10.1016/j.jclinepi.2017.03.001

Jin, J. F., Zhu, L. L., Chen, M., Xu, H. M., Wang, H. F., Feng, X. Q., Zhu, X. P., & Zhou, Q. (2015). The optimal choice of medication administration route regarding intravenous, intramuscular, and subcutaneous injection. *Patient Preference and Adherence*, *9*, 923–942. https://doi.org/10.2147/PPA.S87271

Joensson, M., Thomsen, K. R., Andersen, L. M., Gross, J., Mouridsen, K., Sandberg, K., Østergaard, L., & Lou, H. C. (2015). Making sense: Dopamine activates conscious self-monitoring through medial prefrontal cortex. *Human Brain Mapping*, *36*(5), 1866–1877. https://doi.org/10.1002/hbm.22742

Johnson, J. W., & Kotermanski, S. E. (2006). Mechanism of action of memantine. *Current Opinion in Pharmacology*, *6*(1), 61–67. https://doi.org/10.1016/j.coph.2005.09.007

Jones, J. L., Mateus, C. F., Malcolm, R. J., Brady, K. T., & Back, S. E. (2018). Efficacy of Ketamine in the Treatment of Substance Use Disorders: A Systematic Review. *Frontiers in Psychiatry*, *9*(277). https://doi.org/10.3389/fpsyt.2018.00277

Jourdain, P., Bergersen, L. H., Bhaukaurally, K., Bezzi, P., Santello, M., Domercq, M., Matute, C., Tonello, F., Gundersen, V., & Volterra, A. (2007). Glutamate exocytosis from astrocytes controls synaptic strength. *Nature Neuroscience*, *10*(3), 331–339. https://doi.org/10.1038/nn1849

Kaas, A. L., van Mier, H., & Goebel, R. (2007). The Neural Correlates of Human Working Memory for Haptically Explored Object Orientations. *Cerebral Cortex*, *17*(7), 1637–1649. https://doi.org/10.1093/cercor/bhl074

Kanai, R., Walsh, V., & Tseng, C. (2010). Subjective discriminability of invisibility: A framework for distinguishing perceptual and attentional failures of awareness. *Consciousness and Cognition*, *19*(4), 1045–1057. https://doi.org/10.1016/j.concog.2010.06.003

Kantrowitz, J. T., & Javitt, D. C. (2010). N-methyl-d-aspartate (NMDA) receptor dysfunction or dysregulation: the final common pathway on the road to schizophrenia? *Brain Research Bulletin*, *83*(3–4), 108–121. https://doi.org/10.1016/j.brainresbull.2010.04.006

Kapetaniou, G. E., Reinhard, M. A., Christian, P., Jobst, A., Tobler, P. N., Padberg, F., & Soutschek, A. (2021). The role of oxytocin in delay of gratification and flexibility in non-social decision making.

ELife, *10*, e61844. https://doi.org/10.7554/eLife.61844

Kaptchuk, T. J. (2001). The double-blind, randomized, placebo-controlled trial: Gold standard or golden calf? *Journal of Clinical Epidemiology*, *54*(6), 541–549. https://doi.org/10.1016/S0895-4356(00)00347-4

Karmiloff-Smith, A. (1994). Beyond Modularity: A Developmental Perspective on Cognitive Science. *International Journal of Language & Communication Disorders*, *29*(1), 95–105. https://doi.org/10.3109/13682829409041485

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kelemen, W. L., Frost, P. J., & Weaver, C. A. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition*, *28*(1), 92–107. https://doi.org/10.3758/BF03211579

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). McGraw-Hill.

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. https://doi.org/10.1038/nature07200

Keren, G. B., & Raaijmakers, J. G. W. (1988). On between-subjects versus within-subjects comparisons in testing utility theory. *Organizational Behavior and Human Decision Processes*, *41*(2), 233–247. https://doi.org/10.1016/0749-5978(88)90028-3

Kim, S.-G., Richter, W., & Uğurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, *37*(4), 631–636. https://doi.org/10.1002/mrm.1910370427

Kim, S., Senju, A., Sodian, B., Paulus, M., Itakura, S., Okuno, A., Ueno, M., & Proust, J. (2021). Memory Monitoring and Control in Japanese and German Preschoolers. *Memory & Cognition*. https://doi.org/10.3758/s13421-021-01263-1

Kircher, T. T. J., Koch, K., Stottmeister, F., & Durst, V. (2007). Metacognition and reflexivity in patients with schizophrenia. *Psychopathology*, *40*(4), 254–260. https://doi.org/10.1159/000101730

Klein, T., Ullsperger, M., & Danielmeier, C. (2013). Error awareness and the insula: links to neurological and psychiatric diseases. *Frontiers in Human Neuroscience*, *7*(14). https://doi.org/10.3389/fnhum.2013.00014

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford press.

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22

Koike, H., Iijima, M., & Chaki, S. (2011). Involvement of AMPA receptor in both the rapid and sustained antidepressant-like effects of ketamine in animal models of depression. *Behavioural Brain Research*, *224*(1), 107–111. https://doi.org/10.1016/j.bbr.2011.05.035

Koizumi, A., Maniscalco, B., & Lau, H. (2015). Does perceptual confidence facilitate cognitive control?

Attention, Perception, & Psychophysics, *77*(4), 1295–1306. https://doi.org/10.3758/s13414-015-0843-3

Koriat, A. (2012). When are two heads better than one and why? *Science*, *336*(6079), 20–22. https://doi.org/10.1126/science.1216549

Köther, U., Veckenstedt, R., Vitzthum, F., Roesch-Ely, D., Pfueller, U., Scheu, F., & Moritz, S. (2012). "Don't give me that look" - Overconfidence in false mental state perception in schizophrenia. *Psychiatry Research*, *196*(1), 1–8. https://doi.org/10.1016/j.psychres.2012.03.004

Krall, J., Balle, T., Krogsgaard-Larsen, N., Sørensen, T. E., Krogsgaard-Larsen, P., Kristiansen, U., & Frølund, B. (2015). GABAA Receptor Partial Agonists and Antagonists: Structure, Binding Mode, and Pharmacology. In U. B. T.-A. in P. Rudolph (Ed.), *Diversity and Functions of GABA Receptors: A Tribute to Hanns Möhler, Part A* (Vol. 72, pp. 201–227). Academic Press. https://doi.org/10.1016/bs.apha.2014.10.003

Krystal, J. H., Karper, L. P., Seibyl, J. P., Freeman, G. K., Delaney, R., Bremner, J. D., Heninger, G. R., Bowers, M. B., & Charney, D. S. (1994). Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Archives of General Psychiatry*, *51*(3), 199–214. https://doi.org/10.1001/archpsyc.1994.03950030035004

Krystal, J. H., Sanacora, G., & Duman, R. S. (2013). Rapid-Acting Glutamatergic Antidepressants: The Path to Ketamine and Beyond. *Biological Psychiatry*, *73*(12), 1133–1141. https://doi.org/10.1016/j.biopsych.2013.03.026

Kuchling, F., Fields, C., & Levin, M. (2022). Metacognition as a Consequence of Competing Evolutionary Time Scales. *Entropy*, *24*(5). https://doi.org/10.3390/e24050601

Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and Accuracy of Near-Threshold Discrimination Responses. *Consciousness and Cognition*, *10*(3), 294–340. https://doi.org/10.1006/ccog.2000.0494

Lahti, A. C., Weiler, M. A., Tamara, M., Parwani, A., & Tamminga, C. A. (2001). Effects of Ketamine in Normal and Schizophrenic Volunteers. *Neuropsychopharmacology*, *25*(4), 455–467. https://doi.org/10.1016/S0893-133X(01)00243-3

Landauer, A. A. (1975). Crossover effects in experimental investigations with human subjects. *Pharmacology Biochemistry and Behavior*, *3*(6), 1141–1145. https://doi.org/10.1016/0091-3057(75)90031-3

Lapate, R. C., Samaha, J., Rokers, B., Postle, B. R., & Davidson, R. J. (2020). Perceptual metacognition of human faces is causally supported by function of the lateral prefrontal cortex. *Communications Biology*, *3*(1), 360. https://doi.org/10.1038/s42003-020-1049-3

Lau, A., & Tymianski, M. (2010). Glutamate receptors, neurotoxicity and neurodegeneration. *Pflügers Archiv - European Journal of Physiology*, *460*(2), 525–542. https://doi.org/10.1007/s00424-010-0809-1

Lau, H. C., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, *15*(8), 365–373. https://doi.org/10.1016/j.tics.2011.05.009

Laube, B., Hirai, H., Sturgess, M., Betz, H., & Kuhse, J. (1997). Molecular Determinants of Agonist Discrimination by NMDA Receptor Subunits: Analysis of the Glutamate Binding Site on the NR2B Subunit. *Neuron*, *18*(3), 493–503. https://doi.org/10.1016/S0896-6273(00)81249-0

LeDoux, J. E., Michel, M., & Lau, H. (2020). A little history goes a long way toward understanding why we study consciousness the way we do today. *Proceedings of the National Academy of Sciences*,

117(13), 6976–6984. https://doi.org/10.1073/pnas.1921623117

Lee, A. L. F., Ruby, E., Giles, N., & Lau, H. (2018). Cross-domain association in metacognitive efficiency depends on first-order task types. *Frontiers in Psychology*, *9*(2464), 1–10. https://doi.org/10.3389/fpsyg.2018.02464

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lee, U., Ku, S., Noh, G., Baek, S., Choi, B., & Mashour, G. A. (2013). Disruption of frontal-parietal communication by ketamine, propofol, and sevoflurane. *Anesthesiology*, *118*(6), 1264–1275. https://doi.org/10.1097/ALN.0b013e31829103f5

Lee, Y., Syeda, K., Maruschak, N. A., Cha, D. S., Mansur, R. B., Wium-Andersen, I. K., Woldeyohannes, H. O., Rosenblat, J. D., & McIntyre, R. S. (2016). A New Perspective on the Anti-Suicide Effects With Ketamine Treatment: A Procognitive Effect. *Journal of Clinical Psychopharmacology*, *36*(1). https://doi.org/10.1097/JCP.0000000000000441

Lehmann, M., Hagen, J., & Ettinger, U. (2022). Unity and Diversity of Metacognition. *Journal of Experimental Psychology: General*, *151*(10), 2396–2417. https://doi.org/10.1037/xge0001197

Lehmann, M., Neumann, C., Wasserthal, S., Delis, A., Schultz, J., Hurlemann, R., & Ettinger, U. (2022). Ketamine increases fronto-posterior functional connectivity during meta-perceptual confidence ratings. *Behavioural Brain Research*, *430*, 113925. https://doi.org/10.1016/j.bbr.2022.113925

Lehmann, M., Neumann, C., Wasserthal, S., Schultz, J., Delis, A., Trautner, P., Hurlemann, R., & Ettinger, U. (2021). Effects of ketamine on brain function during metacognition of episodic memory. *Neuroscience of Consciousness*, *2021*(1). https://doi.org/10.1093/nc/niaa028

Lei, W., Chen, J., Yang, C., Guo, Y., Feng, P., Feng, T., & Li, H. (2020). Metacognition-related regions modulate the reactivity effect of confidence ratings on perceptual decision-making. *Neuropsychologia*, *144*, 107502. https://doi.org/10.1016/j.neuropsychologia.2020.107502

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477.

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lindquist, M. A., Meng Loh, J., Atlas, L. Y., & Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. *NeuroImage*, *45*(1, Supplement 1), S187–S198. https://doi.org/10.1016/j.neuroimage.2008.10.065

Linnaeus, C. (1735). *Systema naturae; sive, Regna tria naturae: systematice proposita per classes, ordines, genera & species*. Haak.

Lipton, S. A. (2004). Paradigm shift in NMDA receptor antagonist drug development: Molecular mechanism of uncompetitive inhibition by memantine in the treatment of Alzheimer's disease and other neurologic disorders. *Journal of Alzheimer's Disease*, *6*, S61–S74. https://doi.org/10.3233/JAD-2004-6S610

Lodge, D., & Mercier, M. S. (2015). Ketamine and phencyclidine: the good, the bad and the unexpected. *British Journal of Pharmacology*, *172*(17), 4254–4276. https://doi.org/10.1111/bph.13222

Lofwall, M. R., Griffiths, R. R., & Mintzer, M. Z. (2006). Cognitive and subjective acute dose effects of intramuscular ketamine in healthy adults. *Experimental and Clinical Psychopharmacology*, *14*(4), 439–449. https://doi.org/10.1037/1064-1297.14.4.439

Logothetis, N. K. (2002). The neural basis of the blood–oxygen–level–dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1424), 1003–1037. https://doi.org/10.1098/rstb.2002.1114

Lopez-Morinigo, J. D., Ajnakina, O., Martínez, A. S. E., Escobedo-Aedo, P. J., Ruiz-Ruano, V. G., Sánchez-Alonso, S., Mata-Iturralde, L., Muñoz-Lorenzo, L., Ochoa, S., Baca-Garciá, E., & David, A. S. (2020). Can metacognitive interventions improve insight in schizophrenia spectrum disorders? A systematic review and meta-analysis. *Psychological Medicine*, *50*(14), 2289–2301. https://doi.org/10.1017/S0033291720003384

Lou, H. C., Skewes, J. C., Thomsen, K. R., Overgaard, M., Lau, H. C., Mouridsen, K., & Roepstorff, A. (2011). Dopaminergic stimulation enhances confidence and accuracy in seeing rapidly presented words. *Journal of Vision*, *11*(2), 1–6. https://doi.org/10.1167/11.2.1

Lydiard, R. B. (2003). The role of GABA in anxiety disorders. *Journal of Clinical Psychiatry*, *64*, 21–27.

Lysaker, P. H., Buck, K. D., Carcione, A., Procacci, M., Salvatore, G., Nicolò, G., & Dimaggio, G. (2011). Addressing metacognitive capacity for self reflection in the psychotherapy for schizophrenia: A conceptual model of the key tasks and processes. *Psychology and Psychotherapy: Theory, Research and Practice*, *84*(1), 58–69. https://doi.org/10.1348/147608310X520436

Lysaker, P. H., Carcione, A., Dimaggio, G., Johannesen, J. K., Nicolò, G., Procacci, M., & Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia: associations with neurocognition, symptoms, insight and quality of life. *Acta Psychiatrica Scandinavica*, *112*(1), 64–71. https://doi.org/10.1111/j.1600-0447.2005.00514.x

Lysaker, P. H., & Dimaggio, G. (2014). Metacognitive capacities for reflection in schizophrenia: Implications for developing treatments. *Schizophrenia Bulletin*, *40*(3), 487–491. https://doi.org/10.1093/schbul/sbu038

Lysaker, P. H., Dimaggio, G., Buck, K. D., Callaway, S. S., Salvatore, G., Carcione, A., Nicolò, G., & Stanghellini, G. (2011). Poor insight in schizophrenia: links between different forms of metacognition with awareness of symptoms, treatment need, and consequences of illness. *Comprehensive Psychiatry*, *52*(3), 253–260. https://doi.org/10.1016/j.comppsych.2010.07.007

Lysaker, P. H., Dimaggio, G., Buck, K. D., Carcione, A., & Nicolò, G. (2007). Metacognition within narratives of schizophrenia: Associations with multiple domains of neurocognition. *Schizophrenia Research*, *93*(1), 278–287. https://doi.org/10.1016/j.schres.2007.02.016

Lysaker, P. H., George, S., Chaudoin–Patzoldt, K. A., Pec, O., Bob, P., Leonhardt, B. L., Vohs, J. L., James, A. V, Wickett, A., Buck, K. D., & Dimaggio, G. (2017). Contrasting metacognitive, social cognitive and alexithymia profiles in adults with borderline personality disorder, schizophrenia and substance use disorder. *Psychiatry Research*, *257*, 393–399. https://doi.org/10.1016/j.psychres.2017.08.001

Lysaker, P. H., Kukla, M., Dubreucq, J., Gumley, A., McLeod, H., Vohs, J. L., Buck, K. D., Minor, K. S., Luther, L., Leonhardt, B. L., Belanger, E. A., Popolo, R., & Dimaggio, G. (2015). Metacognitive deficits predict future levels of negative symptoms in schizophrenia controlling for neurocognition, affect recognition, and self-expectation of goal attainment. *Schizophrenia Research*, *168*(1–2), 267–272. https://doi.org/10.1016/j.schres.2015.06.015

Ma, L., Skoblenick, K., Johnston, K., & Everling, S. (2018). Ketamine alters lateral prefrontal oscillations in a rule-based working memory task. *The Journal of Neuroscience*, *38*(10), 2482–2494. https://doi.org/10.1523/jneurosci.2659-17.2018

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory - A user's guide*. Lawrence Erlbaum Associates.

Maki, R. H., & Swett, S. (1987). Metamemory for narrative text. *Memory & Cognition*, *15*(1), 72–83. https://doi.org/10.3758/BF03197713

Malhotra, A. K., Pinals, D. A., Weingartner, H., Sirocco, K., Missar, C. D., Pickar, D., & Breier, A. (1996). NMDA receptor function and human cognition: The effects of ketamine in healthy volunteers. *Neuropsychopharmacology*, *14*(5), 301–307. https://doi.org/10.1016/0893-133X(95)00137-3

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: meta-d′, response-specific meta-d′, and the unequal variance SDT model. In S. M. Fleming & C. D. Frith (Eds.), *The Cognitive Neuroscience of Metacognition* (pp. 25–66). Springer. https://doi.org/10.1007/978-3-642-45190-4_3

Marco, E. M., Adriani, W., Ruocco, L. A., Canese, R., Sadile, A. G., & Laviola, G. (2011). Neurobehavioral adaptations to methylphenidate: The issue of early adolescent exposure. *Neuroscience and Biobehavioral Reviews*, *35*(8), 1722–1739. https://doi.org/10.1016/j.neubiorev.2011.02.011

Markett, S. (2016). Functional magnetic resonance imaging (fMRI). In M. Reuter & C. Montag (Eds.), *Neuroeconomics: Studies in neuroscience, psychology and behavioral economics* (pp. 375–397). Springer.

Martín-Luengo, B., Zinchenko, O., Dolgoarshinnaia, A., & Leminen, A. (2021). Retrospective confidence judgments: Meta-analysis of functional magnetic resonance imaging studies. *Human Brain Mapping*, *42*(10), 3005–3022. https://doi.org/10.1002/hbm.25397

Martin, J. M., Warman, D. M., & Lysaker, P. H. (2010). Cognitive insight in non-psychiatric individuals and individuals with psychosis: an examination using the Beck Cognitive Insight Scale. *Schizophrenia Research*, *121*(1–3), 39–45. https://doi.org/10.1016/j.schres.2010.03.028

Massé, M., & Lecomte, T. (2015). Metacognitive profiles in individuals with a first episode of psychosis and their relation to social functioning and perceived social support. *Schizophrenia Research*, *166*(1), 60–64. https://doi.org/10.1016/j.schres.2015.05.020

Massin-Krauss, M., Bacon, E., & Danion, J. M. (2002). Effects of the benzodiazepine lorazepam on monitoring and control processes in semantic memory. *Consciousness and Cognition*, *11*(1), 123–137. https://doi.org/10.1006/ccog.2001.0538

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 509–527. https://doi.org/10.1037/a0014876

Mayer, M. L., Westbrook, G. L., & Guthrie, P. B. (1984). Voltage-dependent block by Mg2+ of NMDA responses in spinal cord neurones. *Nature*, *309*(5965), 261–263. https://doi.org/10.1038/309261a0

Mayes, S. D., Calhoun, S. L., Bixler, E. O., & Zimmerman, D. N. (2009). IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences*, *19*(2), 238–241. https://doi.org/10.1016/j.lindif.2008.09.001

Mazancieux, A., Dinze, C., Souchay, C., & Moulin, C. J. A. (2020). Metacognitive domain specificity in feeling-of-knowing but not retrospective confidence. *Neuroscience of Consciousness*, *2020*(1), 1–11. https://doi.org/10.1093/nc/niaa001

Mazancieux, A., Fleming, S., Souchay, C., & Moulin, C. (2020). Is there a G factor for metacognition?

Correlations in retrospective metacognitive sensitivity across tasks. *Journal of Experimental Psychology: General*, *149*. https://doi.org/10.1037/xge0000746

Mazor, M., Friston, K. J., & Fleming, S. M. (2020). Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. *ELife*, *9*, 1–34. https://doi.org/10.7554/eLife.53900

McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *The Journal of Neuroscience*, *33*(5), 1897–1906. https://doi.org/10.1523/jneurosci.1890-12.2013

McDonald, B. C., Flashman, L. A., Arciniegas, D. B., Ferguson, R. J., Xing, L., Harezlak, J., Sprehn, G. C., Hammond, F. M., Maerlender, A. C., Kruck, C. L., Gillock, K. L., Frey, K., Wall, R. N., Saykin, A. J., & McAllister, T. W. (2017). Methylphenidate and Memory and Attention Adaptation Training for Persistent Cognitive Symptoms after Traumatic Brain Injury: A Randomized, Placebo-Controlled Trial. *Neuropsychopharmacology*, *42*(9), 1766–1775. https://doi.org/10.1038/npp.2016.261

Meldrum, B. S. (2000). Glutamate as a Neurotransmitter in the Brain: Review of Physiology and Pathology. *The Journal of Nutrition*, *130*(4), 1007S-1015S. https://doi.org/10.1093/jn/130.4.1007S

Menon, R. S., Ogawa, S., Hu, X., Strupp, J. P., Anderson, P., & Uğurbil, K. (1995). BOLD Based Functional MRI at 4 Tesla Includes a Capillary Bed Contribution: Echo-Planar Imaging Correlates with Previous Optical Imaging Using Intrinsic Signals. *Magnetic Resonance in Medicine*, *33*(3), 453–459. https://doi.org/10.1002/mrm.1910330323

Mesbahi-Vasey, S., Veras, L., Yonkunas, M., Johnson, J. W., & Kurnikova, M. G. (2017). All atom NMDA receptor transmembrane domain model development and simulations in lipid bilayers and water. *PLoS ONE*, *12*(6), e0177686–e0177686. https://doi.org/10.1371/journal.pone.0177686

Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(2), 288–294. https://doi.org/10.1037/0278-7393.12.2.288

Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. The MIT Press. https://doi.org/10.7551/mitpress/4561.001.0001

Metcalfe, J., & Son, L. K. (2012). Anoetic, noetic, and autonoetic metacognition. In M. J. Beran, J. L. Brandl, J. Perner, & J. Proust (Eds.), *The foundations of metacognition* (pp. 289–301). Oxford University Press.

Meuwese, J. D. I., Van Loon, A. M., Scholte, H. S., Lirk, P. B., Vulink, N. C. C., Hollmann, M. W., & Lamme, V. A. F. (2013). NMDA receptor antagonist ketamine impairs feature integration in visual perception. *PLoS ONE*, *8*(11), 1–12. https://doi.org/10.1371/journal.pone.0079326

Miele, D. B., Wager, T. D., Mitchell, J. P., & Metcalfe, J. (2011). Dissociating Neural Correlates of Action Monitoring and Metacognition of Agency. *Journal of Cognitive Neuroscience*, *23*(11), 3620–3636. https://doi.org/10.1162/jocn_a_00052

Mintzer, M. Z., & Griffiths, R. R. (2003). Lorazepam and scopolamine: A single-dose comparison of effects on human memory and attentional processes. *Experimental and Clinical Psychopharmacology*, *11*(1), 56–72. https://doi.org/10.1037/1064-1297.11.1.56

Mintzer, M. Z., Kleykamp, B. A., & Griffiths, R. R. (2010). Dose effects of triazolam and scopolamine on metamemory. *Experimental and Clinical Psychopharmacology*, *18*(1), 17–31. https://doi.org/10.1037/a0018061

Misra, S. (2012). Randomized double blind placebo control studies, the “gold standard” in intervention

based studies. *Indian Journal of Sexually Transmitted Diseases*, *33*(2), 131–134. https://doi.org/10.4103/0253-7184.102130

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

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

Miyamoto, K., Trudel, N., Kamermans, K., Lim, M. C., Lazari, A., Verhagen, L., Wittmann, M. K., & Rushworth, M. F. S. (2021). Identification and disruption of a neural mechanism for accumulating prospective metacognitive information prior to decision-making. *Neuron*, *109*(8), 1396–1408. https://doi.org/10.1016/j.neuron.2021.02.024

Moeller, S. J., Fleming, S. M., Gan, G., Zilverstand, A., Malaker, P., d'Oleire Uquillas, F., Schneider, K. E., Preston-Campbell, R. N., Parvaz, M. A., Maloney, T., Alia-Klein, N., & Goldstein, R. Z. (2016). Metacognitive impairment in active cocaine use disorder is associated with individual differences in brain structure. *European Neuropsychopharmacology*, *26*(4), 653–662. https://doi.org/10.1016/j.euroneuro.2016.02.009

Molenberghs, P., Trautwein, F.-M., Böckler, A., Singer, T., & Kanske, P. (2016). Neural correlates of metacognitive ability and of feeling confident: a large-scale fMRI study. *Social Cognitive and Affective Neuroscience*, *11*(12), 1942–1951. https://doi.org/10.1093/scan/nsw093

Monteggia, L. M., & Zarate, C. (2015). Antidepressant actions of ketamine: from molecular mechanisms to clinical practice. *Current Opinion in Neurobiology*, *30*, 139–143. https://doi.org/10.1016/j.conb.2014.12.004

Moore, J. W., Turner, D. C., Corlett, P. R., Arana, F. S., Morgan, H. L., Absalom, A. R., Adapa, R., de Wit, S., Everitt, J. C., Gardner, J. M., Pigott, J. S., Haggard, P., & Fletcher, P. C. (2011). Ketamine administration in healthy volunteers reproduces aberrant agency experiences associated with schizophrenia. *Cognitive Neuropsychiatry*, *16*(4), 364–381. https://doi.org/10.1080/13546805.2010.546074

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *The Journal of Neuroscience*, *38*(14), 3534–3546. https://doi.org/10.1523/jneurosci.2360-17.2018

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive Psychology*, *78*, 99–147. https://doi.org/10.1016/j.cogpsych.2015.01.002

Moreira, C. M., Rollwage, M., Kaduk, K., Wilke, M., & Kagan, I. (2018). Post-decision wagering after perceptual judgments reveals bi-directional certainty readouts. *Cognition*, *176*, 40–52. https://doi.org/10.1016/j.cognition.2018.02.026

Morgan, C. J. A., Mofeez, A., Brandner, B., Bromley, L., & Curran, H. V. (2004). Acute Effects of Ketamine on Memory Systems and Psychotic Symptoms in Healthy Volunteers. *Neuropsychopharmacology*, *29*(1), 208–218. https://doi.org/10.1038/sj.npp.1300342

Moritz, S., Andreou, C., Schneider, B. C., Wittekind, C. E., Menon, M., Balzan, R. P., & Woodward, T. S. (2014). Sowing the seeds of doubt: A narrative review on metacognitive training in schizophrenia. *Clinical Psychology Review*, *34*(4), 358–366. https://doi.org/10.1016/j.cpr.2014.04.004

Moritz, S., Gläscher, J., Sommer, T., Büchel, C., & Braus, D. F. (2006). Neural correlates of memory confidence. *NeuroImage*, *33*(4), 1188–1193. https://doi.org/10.1016/j.neuroimage.2006.08.003

Moritz, S., Lysaker, P. H., Hofmann, S. G., & Hautzinger, M. (2018). Going meta on metacognitive interventions. *Expert Review of Neurotherapeutics*, *18*(10), 739–741. https://doi.org/10.1080/14737175.2018.1520636

Moritz, S., & Woodward, T. S. (2006). The contribution of metamemory deficits to schizophrenia. *Journal of Abnormal Psychology*, *115*(1), 15–25. https://doi.org/10.1037/0021-843X.15.1.15

Moritz, S., & Woodward, T. S. (2007). Metacognitive training in schizophrenia: from basic research to knowledge translation and intervention. *Current Opinion in Psychiatry*, *20*(6), 619–625. https://doi.org/10.1097/YCO.0b013e3282f0b8ed

Moritz, S., Woodward, T. S., & Ruff, C. C. (2003). Source monitoring and memory confidence in schizophrenia. *Psychological Medicine*, *33*(1), 131–139. https://doi.org/10.1017/S0033291702006852

Moritz, S., Woodward, T. S., Whitman, J. C., & Cuttler, C. (2005). Confidence in errors as a possible basis for delusions in schizophrenia. *Journal of Nervous and Mental Disease*, *193*(1), 9–16. https://doi.org/10.1097/01.nmd.0000149213.10692.00

Moses-Payne, M. E., Rollwage, M., Fleming, S. M., & Roiser, J. P. (2019). Postdecision Evidence Integration and Depressive Symptoms. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00639

Moss, G. P. (1996). Basic terminology of stereochemistry (IUPAC Recommendations 1996). *Pure and Applied Chemistry*, *68*(12), 2193–2222. https://doi.org/10.1351/pac199668122193

Mueller, R. O., & Hancock, G. R. (2001). Factor Analysis and Latent Structure, Confirmatory. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 5239–5244). Elsevier Ltd. https://doi.org/10.1016/b0-08-043076-7/00426-5

Muetzelfeldt, L., Kamboj, S. K., Rees, H., Taylor, J., Morgan, C. J. A., & Curran, H. V. (2008). Journey through the K-hole: Phenomenological aspects of ketamine use. *Drug and Alcohol Dependence*, *95*(3), 219–229. https://doi.org/10.1016/j.drugalcdep.2008.01.024

Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*(3), 218–228. https://doi.org/10.1038/nrn1345

Murrough, J. W., Iosifescu, D. V, Chang, L. C., Al Jurdi, R. K., Green, C. E., Perez, A. M., Iqbal, S., Pillemer, S., Foulkes, A., & Shah, A. (2013). Antidepressant efficacy of ketamine in treatment-resistant major depression: a two-site randomized controlled trial. *American Journal of Psychiatry*, *170*(10), 1134–1142. https://doi.org/10.1176/appi.ajp.2013.13030392

Narendran, R., Frankle, W. G., Keefe, R., Gil, R., Martinez, D., Slifstein, M., Kegeles, L. S., Talbot, P. S., Huang, Y., Hwang, D.-R., Khenissi, L., Cooper, T. B., Laruelle, M., & Abi-Dargham, A. (2005). Altered Prefrontal Dopaminergic Function in Chronic Recreational Ketamine Users. *American Journal of Psychiatry*, *162*(12), 2352–2359. https://doi.org/10.1176/appi.ajp.162.12.2352

Nedergaard, M., Takano, T., & Hansen, A. J. (2002). Beyond the role of glutamate as a neurotransmitter. *Nature Reviews Neuroscience*, *3*(9), 748–755. https://doi.org/10.1038/nrn916

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*(1), 109–133. https://doi.org/10.1037/0033-2909.95.1.109

Nelson, T. O. (1986). ROC curves and measures of discrimination accuracy: A reply to Swets. *Psychological Bulletin*, *100*(1), 128–132. https://doi.org/10.1037/0033-2909.100.1.128

Nelson, T. O., & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 26, pp. 125–173). Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Neubig, R. R., Spedding, M., Kenakin, T., & Christopoulos, A. (2003). International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on Terms and Symbols in Quantitative Pharmacology. *Pharmacological Reviews*, *55*(4), 597–606. https://doi.org/10.1124/pr.55.4.4

Nichols, S., & Stich, S. (2003). How to read your own mind: A cognitive theory of self-consciousness. *Consciousness: New Philosophical Essays*, 157–200.

Nieuwenhuis, S., Richard Ridderinkhof, K., Blom, J., Band, G. P. H., & 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. https://doi.org/10.1017/S0048577201001111

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Nolan, M. F., Dudman, J. T., Dodson, P. D., & Santoro, B. (2007). HCN1 Channels Control Resting and Active Integrative Properties of Stellate Cells from Layer II of the Entorhinal Cortex. *The Journal of Neuroscience*, *27*(46), 12440–12451. https://doi.org/10.1523/jneurosci.2358-07.2007

Nordahl, H., Anyan, F., Hjemdal, O., & Wells, A. (2022). The network structure of dysfunctional metacognition: Analysis of the MCQ-30. *Acta Psychologica*, *227*, 103622. https://doi.org/10.1016/j.actpsy.2022.103622

Norman, E. (2020). Why Metacognition Is Not Always Helpful. *Frontiers in Psychology*, *11*, 1537. https://doi.org/10.3389/fpsyg.2020.01537

Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in psychology. *Review of General Psychology*, *23*(4), 403–424. https://doi.org/10.1177/1089268019883821

Norris, C. M., Blalock, E. M., Thibault, O., Brewer, L. D., Clodfelter, G. V, Porter, N. M., & Landfield, P. W. (2006). Electrophysiological Mechanisms of Delayed Excitotoxicity: Positive Feedback Loop Between NMDA Receptor Current and Depolarization-Mediated Glutamate Release. *Journal of Neurophysiology*, *96*(5), 2488–2500. https://doi.org/10.1152/jn.00593.2005

Norris, D. G. (2006). Principles of magnetic resonance assessment of brain function. *Journal of Magnetic Resonance Imaging*, *23*(6), 794–807. https://doi.org/10.1002/jmri.20587

Nutt, D. (2014). A brave new world for psychology? *Psychologist*, *27*(9), 658–660.

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604–609. https://doi.org/10.1093/scan/nss055

Ogden, K. K., & Traynelis, S. F. (2011). New advances in NMDA receptor pharmacology. *Trends in Pharmacological Sciences*, *32*(12), 726–733. https://doi.org/10.1016/j.tips.2011.08.003

Olivares, D., K. Deshpande, V., Shi, Y., K. Lahiri, D., H. Greig, N., T. Rogers, J., & Huang, X. (2012). N-Methyl D-Aspartate (NMDA) Receptor Antagonists and Memantine Treatment for Alzheimer's Disease, Vascular Dementia and Parkinson's Disease. *Current Alzheimer Research*, *9*(6), 746–758. https://doi.org/10.2174/156720512801322564

Oranje, B., van Berckel, B. N. M., Kemner, C., van Ree, J. M., Kahn, R. S., & Verbaten, M. N. (2000). The Effects of a Sub-Anaesthetic Dose of Ketamine on Human Selective Attention.

Neuropsychopharmacology, *22*(3), 293–302. https://doi.org/10.1016/S0893-133X(99)00118-9

Orser, B. A., Pennefather, P. S., & MacDonald, J. F. (1997). Multiple Mechanisms of Ketamine Blockade of N-methyl-D-aspartate Receptors. *Anesthesiology*, *86*(4), 903–917. https://doi.org/10.1097/00000542-199704000-00021

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin*, *136*(1), 65–86. https://doi.org/10.1037/a0017815

Overgaard, M., Rote, J., Mouridsen, K., & Ramsøy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition*, *15*(4), 700–708. https://doi.org/10.1016/j.concog.2006.04.002

Øye, I., Paulsen, O., & Maurset, A. (1992). Effects of ketamine on sensory perception: Evidence for a role of N- methyl-D-aspartate receptors. *Journal of Pharmacology and Experimental Therapeutics*, *260*(3), 1209–1213.

Pallavicini, C., Vilas, M. G., Villarreal, M., Zamberlan, F., Muthukumaraswamy, S., Nutt, D., Carhart-Harris, R., & Tagliazucchi, E. (2019). Spectral signatures of serotonergic psychedelics and glutamatergic dissociatives. *NeuroImage*, *200*, 281–291. https://doi.org/10.1016/j.neuroimage.2019.06.053

Palmer, E. C., David, A. S., & Fleming, S. M. (2014). Effects of age on metacognitive efficiency. *Consciousness and Cognition*, *28*(1), 151–160. https://doi.org/10.1016/j.concog.2014.06.007

Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology Review*, *15*(3), 105–130. https://doi.org/10.1007/s11065-005-7091-6

Pannu, J. K., Kaszniak, A. W., & Rapcsak, S. Z. (2005). Metamemory for faces following frontal lobe damage. *Journal of the International Neuropsychological Society*, *11*(6), 668–676. https://doi.org/10.1017/S1355617705050873

Park, C. I., Kim, H. W., Jeon, S., Hwang, E. H., Kang, J. I., & Kim, S. J. (2020). Metacognitive beliefs predict early response to pharmacological treatment in patients with obsessive–compulsive disorder. *Psychopharmacology*, *237*(11), 3489–3496. https://doi.org/10.1007/s00213-020-05630-9

Parton, A., Nachev, P., Hodgson, T. L., Mort, D., Thomas, D., Ordidge, R., Morgan, P. S., Jackson, S., Rees, G., & Husain, M. (2007). Role of the human supplementary eye field in the control of saccadic eye movements. *Neuropsychologia*, *45*(5), 997–1008. https://doi.org/10.1016/j.neuropsychologia.2006.09.007

Pasquali, A., Timmermans, B., & Cleeremans, A. (2010). Know thyself: Metacognitive networks and measures of consciousness. *Cognition*, *117*(2), 182–190. https://doi.org/10.1016/j.cognition.2010.08.010

Passie, T., Adams, H.-A., Logemann, F., Brandt, S. D., Wiese, B., & Karst, M. (2021). Comparative effects of (S)-ketamine and racemic (R/S)-ketamine on psychopathology, state of consciousness and neurocognitive performance in healthy volunteers. *European Neuropsychopharmacology*, *44*, 92–104. https://doi.org/10.1016/j.euroneuro.2021.01.005

Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2010). Medial frontal cortex: from self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences*, *14*(1), 16–21. https://doi.org/10.1016/j.tics.2009.11.001

Paul, E. J., Smith, J. D., Valentin, V. V., Turner, B. O., Barbey, A. K., & Ashby, F. G. (2015). Neural networks underlying the metacognitive uncertainty response. *Cortex*, *71*, 306–322. https://doi.org/10.1016/j.cortex.2015.07.028

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the Mind's Eye: The Metacognition of Visual Imagery. *Psychological Science*, *22*(12), 1535–1542. https://doi.org/10.1177/0956797611417134

Pedone, R., Semerari, A., Riccardi, I., Procacci, M., Nicolò, G., & Carcione, A. (2017). Development of a self-report measure of metacognition: The Metacognition Self-Assessment Scale (MSAS). Instrument description and factor structure. *Clinical Neuropsychiatry: Journal of Treatment Evaluation*, *14*(3), 185–194.

Peltoniemi, M. A., Hagelberg, N. M., Olkkola, K. T., & Saari, T. I. (2016). Ketamine: A Review of Clinical Pharmacokinetics and Pharmacodynamics in Anesthesia and Pain Therapy. *Clinical Pharmacokinetics*, *55*(9), 1059–1077. https://doi.org/10.1007/s40262-016-0383-6

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, *10*(2), 257–261. https://doi.org/10.1038/nn1840

Peters, M. A. K., Thesen, T., Ko, Y. D., Maniscalco, B., Carlson, C., Davidson, M., Doyle, W., Kuzniecky, R., Devinsky, O., Halgren, E., & Lau, H. (2017). Perceptual confidence neglects decision-incongruent evidence in the brain. *Nature Human Behaviour*, *1*(7), 139. https://doi.org/10.1038/s41562-017-0139

Platel, J.-C., Lacar, B., & Bordey, A. (2007). GABA and glutamate signaling: homeostatic control of adult forebrain neurogenesis. *Journal of Molecular Histology*, *38*(4), 303–311. https://doi.org/10.1007/s10735-007-9103-8

Pokorny, T., Preller, K. H., Kraehenmann, R., & Vollenweider, F. X. (2016). Modulatory effect of the 5-HT1A agonist buspirone and the mixed non-hallucinogenic 5-HT1A/2A agonist ergotamine on psilocybin-induced psychedelic experience. *European Neuropsychopharmacology*, *26*(4), 756–766. https://doi.org/10.1016/j.euroneuro.2016.01.005

Preciado, D., & Theeuwes, J. (2018). To look or not to look? Reward, selection history, and oculomotor guidance. *Journal of Neurophysiology*, *120*(4), 1740–1752. https://doi.org/10.1152/jn.00275.2018

Pronin, E. (2008). How we see ourselves and how we see others. *Science*, *320*(5880), 1177–1180. https://doi.org/10.1126/science.1154199

Rahnev, D. (2017). Top-Down Control of Perceptual Decision Making by the Prefrontal Cortex. *Current Directions in Psychological Science*, *26*(5), 464–469. https://doi.org/10.1177/0963721417709807

Rahnev, D., & Fleming, S. M. (2019). How experimental procedures influence estimates of metacognitive ability. *Neuroscience of Consciousness*, *2019*(1). https://doi.org/10.1093/nc/niz009

Rausch, M., Müller, H. J., & Zehetleitner, M. (2015). Metacognitive sensitivity of subjective reports of decisional confidence and visual experience. *Consciousness and Cognition*, *35*, 192–205. https://doi.org/10.1016/j.concog.2015.02.011

Rausch, M., & Zehetleitner, M. (2022). Evaluating false positive rates of standard and hierarchical measures of metacognitive accuracy. *PsyArXiv*. https://doi.org/10.31234/osf.io/593av

Reyes, G., Silva, J. R., Jaramillo, K., Rehbein, L., & Sackur, J. (2015). Self-Knowledge Dim-Out: Stress Impairs Metacognitive Accuracy. *PLoS ONE*, *10*(8), e0132320. https://doi.org/10.1371/journal.pone.0132320

Reyes, G., Vivanco-Carlevari, A., Medina, F., Manosalva, C., de Gardelle, V., Sackur, J., & Silva, J. R. (2020). Hydrocortisone decreases metacognitive efficiency independent of perceived stress. *Scientific Reports*, *10*(1), 14100. https://doi.org/10.1038/s41598-020-71061-3

Reynolds, R. E., & Wade, S. E. (1986). Thinking about Thinking about Thinking: Reflections on Metacognition. *Harvard Educational Review*, *56*, 307–318. https://doi.org/10.17763/haer.56.3.a2v45g440pg45173

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140. https://doi.org/10.1016/j.bandc.2004.09.016

Riegel, M., Wierzba, M., Wypych, M., Ritchey, M., Jednoróg, K., Grabowska, A., Vuilleumier, P., & Marchewka, A. (2022). Distinct medial-temporal lobe mechanisms of encoding and amygdala-mediated memory reinstatement for disgust and fear. *NeuroImage*, *251*, 118889. https://doi.org/10.1016/j.neuroimage.2022.118889

Riggs, S. E., Grant, P. M., Perivoliotis, D., & Beck, A. T. (2012). Assessment of cognitive insight: a qualitative review. *Schizophrenia Bulletin*, *38*(2), 338–350. https://doi.org/10.1093/schbul/sbq085

Robinson, D. M., & Keating, G. M. (2006). Memantine - A Review of its Use in Alzheimer's Disease. *Drugs*, *66*(11), 1515–1534. https://doi.org/10.2165/00003495-200666110-00015

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

Rollwage, M., Dolan, R. J., & Fleming, S. M. (2018). Metacognitive Failure as a Feature of Those Holding Radical Beliefs. *Current Biology*, *28*(24), 4014–4021. https://doi.org/10.1016/j.cub.2018.10.053

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-09075-3

Rouault, M., Lebreton, M., & Pessiglione, M. (2022). A shared brain system forming confidence judgment across cognitive domains. *Cerebral Cortex*, bhac146. https://doi.org/10.1093/cercor/bhac146

Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human Metacognition Across Domains: Insights from Individual Differences and Neuroimaging. *Personality Neuroscience*, *1*(e17). https://doi.org/10.1017/pen.2018.16

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. https://doi.org/10.1016/j.biopsych.2017.12.017

Rouy, M., de Gardelle, V., Reyes, G., Sackur, J., Vergnaud, J. C., Filevich, E., & Faivre, N. (2022). Metacognitive improvement: Disentangling adaptive training from experimental confounds. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001185

Rouy, M., Saliou, P., Nalborczyk, L., Pereira, M., Roux, P., & Faivre, N. (2021). Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. *Neuroscience and Biobehavioral Reviews*, *126*, 329–337. https://doi.org/10.1016/j.neubiorev.2021.03.017

Ruby, E., Giles, N., & Lau, H. (2017). Finding domain-general metacognitive mechanisms requires using appropriate tasks. *BioRxiv*, 211805. https://doi.org/10.1101/211805

Rummel, J., & Meiser, T. (2013). The role of metacognition in prospective memory: Anticipated task demands influence attention allocation strategies. *Consciousness and Cognition*, *22*(3), 931–943.

https://doi.org/10.1016/j.concog.2013.06.006

Sadeghi, S., Ekhtiari, H., Bahrami, B., & Ahmadabadi, M. N. (2017). Metacognitive Deficiency in a Perceptual but Not a Memory Task in Methadone Maintenance Patients. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-06707-w

Samaha, J., & Postle, B. R. (2017). Correlated individual differences suggest a common mechanism underlying metacognition in visual perception and visual short-term memory. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1867), 1–10. https://doi.org/10.1098/rspb.2017.2035

Samaha, J., Switzky, M., & Postle, B. R. (2019). Confidence boosts serial dependence in orientation estimation. *Journal of Vision*, *19:4*(25), 1–13. https://doi.org/10.1167/19.4.25

Sanchez-Prieto, J., Budd, D. C., Herrero, I., Vázquez, E., & Nicholls, D. G. (1996). Presynaptic receptors and the control of glutamate exocytosis. *Trends in Neurosciences*, *19*(6), 235–239. https://doi.org/10.1016/0166-2236(96)10031-X

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069–1078. https://doi.org/10.1016/j.concog.2009.12.013

Saraç, S., & Karakelle, S. (2012). On-line and off-line assessment of metacognition. *International Electronic Journal of Elementary Education*, *4*(2), 301–315. https://www.iejee.com/index.php/IEJEE/article/view/201

Sauvé, G., Lavigne, K. M., Pochiet, G., Brodeur, M. B., & Lepage, M. (2020). Efficacy of psychological interventions targeting cognitive biases in schizophrenia: A systematic review and meta-analysis. *Clinical Psychology Review*, *78*(June 2020), 101854. https://doi.org/10.1016/j.cpr.2020.101854

Scarapicchia, V., Brown, C., Mayo, C., & Gawryluk, J. R. (2017). Functional Magnetic Resonance Imaging and Functional Near-Infrared Spectroscopy: Insights from Combined Recording Studies. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00419

Schauwecker, P. E. (2010). Neuroprotection by glutamate receptor antagonists against seizure-induced excitotoxic cell death in the aging brain. *Experimental Neurology*, *224*(1), 207–218. https://doi.org/10.1016/j.expneurol.2010.03.013

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9

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

Schilling, L., Moritz, S., Kriston, L., Krieger, M., & Nagel, M. (2018). Efficacy of metacognitive training for patients with borderline personality disorder: Preliminary results. *Psychiatry Research*, *262*, 459–464. https://doi.org/10.1016/j.psychres.2017.09.024

Schmechtig, A., Lees, J., Perkins, A., Altavilla, A., Craig, K. J., Dawson, G. R., William Deakin, J. F., Dourish, C. T., Evans, L. H., Koychev, I., Weaver, K., Smallman, R., Walters, J., Wilkinson, L. S., Morris, R., Williams, S. C. R., & Ettinger, U. (2013). The effects of ketamine and risperidone on eye movement control in healthy volunteers. *Translational Psychiatry*, *3*(12), e334-10. https://doi.org/10.1038/tp.2013.109

Schmidt, T. T., & Berkemeyer, H. (2018). The altered states database: Psychometric data of altered states of consciousness. *Frontiers in Psychology*, *9*(1028), 1–5. https://doi.org/10.3389/fpsyg.2018.01028

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326. https://doi.org/10.1016/j.tics.2011.05.006

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical Evidence of Bias: Dimensions of Methodological Quality Associated With Estimates of Treatment Effects in Controlled Trials. *Journal of the American Medical Association*, *273*(5), 408–412. https://doi.org/10.1001/jama.1995.03520290060030

Schwartz, J. H. (2002). Neurotransmitters. In V. S. Ramachandran (Ed.), *Encyclopedia of the Human Brain: Neuroscience and Biobehavioral Psychology* (pp. 601–611). Academic Press. https://doi.org/10.1016/B0-12-227210-2/00251-X

Schwartz, T., Sachdeva, S., & Stahl, S. (2012). Glutamate Neurocircuitry: Theoretical Underpinnings in Schizophrenia. *Frontiers in Pharmacology*, *3*, 195. https://doi.org/10.3389/fphar.2012.00195

Scott, B., & Levy, M. (2013). Metacognition: Examining the Components of a Fuzzy Concept. *Educational Research EJournal*, *2*, 120–131. https://doi.org/10.5838/erej.2013.22.04

Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., & Alleva, G. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychology & Psychotherapy*, *10*(4), 238–261. https://doi.org/10.1002/cpp.362

Semerari, A., Carcione, A., Dimaggio, G., Nicolò, G., & Procacci, M. (2007). Understanding minds: Different functions and different disorders? The contribution of psychotherapy research. *Psychotherapy Research*, *17*(1), 106–119. https://doi.org/10.1080/10503300500536953

Sen, S., & Bradshaw, L. (2017). Comparison of Relative Fit Indices for Diagnostic Model Selection. *Applied Psychological Measurement*, *41*(6), 422–438. https://doi.org/10.1177/0146621617695521

Seow, T. X. F., Rouault, M., Gillan, C. M., & Fleming, S. M. (2021). How Local and Global Metacognition Shape Mental Health. *Biological Psychiatry*, *90*(7), 436–446. https://doi.org/10.1016/j.biopsych.2021.05.013

Sharma, R., & Premachandra, B. R. (1991). Membrane-bound hemoglobin as a marker of oxidative injury in adult and neonatal red blood cells. *Biochemical Medicine and Metabolic Biology*, *46*(1), 33–44. https://doi.org/10.1016/0885-4505(91)90048-P

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. https://doi.org/10.1016/j.tics.2014.01.006

Shea, N., & Frith, C. D. (2019). The Global Workspace Needs Metacognition. *Trends in Cognitive Sciences*, *23*(7), 560–571. https://doi.org/10.1016/j.tics.2019.04.007

Shekhar, M., & Rahnev, D. (2018). Distinguishing the Roles of Dorsolateral and Anterior PFC in Visual Metacognition. *The Journal of Neuroscience*, *38*(22), 5078–5087. https://doi.org/10.1523/jneurosci.3484-17.2018

Shen, K., Kalwarowsky, S., Clarence, W., Brunamonti, E., & Paré, M. (2010). Beneficial effects of the NMDA antagonist ketamine on decision processes in visual search. *The Journal of Neuroscience*,

30(29), 9947–9953. https://doi.org/10.1523/jneurosci.6317-09.2010

Shushruth, S. (2013). Exploring the Neural Basis of Consciousness through Anesthesia. *The Journal of Neuroscience*, *33*(5), 1757–1758. https://doi.org/10.1523/jneurosci.5215-12.2013

Sibbald, B., & Roland, M. (1998). Understanding controlled trials. Why are randomised controlled trials important? *British Medical Journal*, *316*(7126), 201. https://doi.org/10.1136/bmj.316.7126.201

Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I Was So Sure! Metacognitive Judgments Are Less Accurate Given Prospectively than Retrospectively. *Frontiers in Psychology*, *7*, 218. https://doi.org/10.3389/fpsyg.2016.00218

Sinner, B., & Graf, B. M. (2008). Ketamine. In J. Schüttler & H. Schwilden (Eds.), *Modern Anesthetics* (pp. 313–333). Springer. https://doi.org/10.1007/978-3-540-74806-9_15

Skoblenick, K., & Everling, S. (2014). N-methyl-D-aspartate Receptor Antagonist Ketamine Impairs Action-monitoring Activity in the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *26*(3), 577–592. https://doi.org/10.1162/jocn_a_00519

Sleigh, J., Harvey, M., Voss, L., & Denny, B. (2014). Ketamine – More mechanisms of action than just NMDA blockade. *Trends in Anaesthesia and Critical Care*, *4*(2), 76–81. https://doi.org/10.1016/j.tacc.2014.03.002

Smith, J. D., Beran, M. J., Couchman, J. J., Coutinho, M. V. C., & Boomer, J. B. (2009). Animal metacognition: Problems and prospects. *Comparative Cognition & Behavior Reviews*, *4*, 40–53. https://doi.org/10.3819/ccbr.2009.40004

Smith, D. V., Gseir, M., Speer, M. E., & Delgado, M. R. (2016). Toward a cumulative science of functional integration: A meta-analysis of psychophysiological interactions. *Human Brain Mapping*, *37*(8), 2904–2917. https://doi.org/10.1002/hbm.23216

Solanto, M. V, Marks, D. J., Wasserstein, J., Mitchell, K., Abikoff, H., Alvir, J. M. J., & Kofman, M. D. (2010). Efficacy of meta-cognitive therapy for adult ADHD. *The American Journal of Psychiatry*, *167*(8), 958–968. https://doi.org/10.1176/appi.ajp.2009.09081123

Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., & Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Consciousness and Cognition*, *20*(4), 1787–1792. https://doi.org/10.1016/j.concog.2010.12.011

Song, J. H. H., Loyal, S., & Lond, B. (2021). Metacognitive Awareness Scale, Domain Specific (MCAS-DS): Assessing Metacognitive Awareness During Raven's Progressive Matrices. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.607577

Souchay, C., Guillery-Girard, B., Pauly-Takacs, K., Wojcik, D., & Eustache, F. (2013). Subjective Experience of Episodic Memory and Metacognition: A Neurodevelopmental Approach. *Frontiers in Behavioral Neuroscience*, *7*. https://doi.org/10.3389/fnbeh.2013.00212

Spada, M. M., Caselli, G., Nikčević, A. V., & Wells, A. (2014). Metacognition in addictive behaviors. *Addictive Behaviors*, *44*, 9–15. https://doi.org/10.1016/j.addbeh.2014.08.002

Spada, M. M., & Wells, A. (2008). Metacognitive beliefs about alcohol use: Development and validation of two self-report scales. *Addictive Behaviors*, *33*(4), 515–527. https://doi.org/10.1016/j.addbeh.2007.10.011

Spada, M. M., & Wells, A. (2009). A metacognitive model of problem drinking. *Clinical Psychology & Psychotherapy*, *16*(5), 383–393. https://doi.org/10.1002/cpp.620

Spada, M. M., Zandvoort, M., & Wells, A. (2007). Metacognitions in Problem Drinkers. *Cognitive Therapy and Research*, *31*(5), 709–716. https://doi.org/10.1007/s10608-006-9066-1

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and self-regulated learning constructs. *Educational Research and Evaluation*, *10*(2), 117–139. https://doi.org/10.1076/edre.10.2.117.27905

Sporns, O., Tononi, G., & Kötter, R. (2005). The Human Connectome: A Structural Description of the Human Brain. *PLoS Computational Biology*, *1*(4), e42. https://doi.org/10.1371/journal.pcbi.0010042

Stahl, S. M. (2013). Mechanism of action of dextromethorphan/quinidine: comparison with ketamine. *CNS Spectrums*, *18*(5), 225–227. https://doi.org/10.1017/S109285291300062X

Steffens, M., Becker, B., Neumann, C., Kasparbauer, A. M., Meyhöfer, I., Weber, B., Mehta, M. A., Hurlemann, R., & Ettinger, U. (2016). Effects of ketamine on brain function during smooth pursuit eye movements. *Human Brain Mapping*, *37*(11), 4047–4060. https://doi.org/10.1002/hbm.23294

Stevens, M. C. (2016). The contributions of resting state and task-based functional connectivity studies to our understanding of adolescent brain network maturation. *Neuroscience and Biobehavioral Reviews*, *70*, 13–32. https://doi.org/10.1016/j.neubiorev.2016.07.027

Stone, J. M., Abel, K. M., Allin, M. P. G., van Haren, N., Matsumoto, K., McGuire, P. K., & Fu, C. H. Y. (2011). Ketamine-Induced Disruption of Verbal Self-Monitoring Linked to Superior Temporal Activation. *Pharmacopsychiatry*, *44*(1), 33–38. https://doi.org/10.1055/s-0030-1267942

Swinney, D. C. (2004). Biochemical mechanisms of drug action: what does it take for success? *Nature Reviews Drug Discovery*, *3*(9), 801–808. https://doi.org/10.1038/nrd1500

Szczepanik, J. E., Brycz, H., Kleka, P., Fanslau, A., Zarate, C. A., & Nugent, A. C. (2020). Metacognition and emotion – How accurate perception of own biases relates to positive feelings and hedonic capacity. *Consciousness and Cognition*, *82*, 102936. https://doi.org/10.1016/j.concog.2020.102936

Thomas, E. H. X., Steffens, M., Harms, C., Rossell, S. L., Gurvich, C., & Ettinger, U. (2021). Schizotypy, neuroticism, and saccadic eye movements: New data and meta-analysis. *Psychophysiology*, *58*(1), e13706. https://doi.org/10.1111/psyp.13706

Timmermans, B., Schilbach, L., Pasquali, A., & Cleeremans, A. (2012). Higher order thoughts in action: consciousness as an unconscious re-description process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1412–1423. https://doi.org/10.1098/rstb.2011.0421

Tobias, S., & Everson, H. (1996). Assessing Metacognitive Knowledge Monitoring. *College Board Report*, *96–01*.

Toglia, J., & Kirk, U. (2000). Understanding awareness deficits following brain injury. *NeuroRehabilitation*, *15*(1), 57–70. https://doi.org/10.3233/nre-2000-15104

Tovar, K. R., McGinley, M. J., & Westbrook, G. L. (2013). Triheteromeric NMDA Receptors at Hippocampal Synapses. *The Journal of Neuroscience*, *33*(21), 9150–9160. https://doi.org/10.1523/jneurosci.0829-13.2013

Tyler, M. W., Yourish, H. B., Ionescu, D. F., & Haggarty, S. J. (2017). Classics in Chemical Neuroscience: Ketamine. *ACS Chemical Neuroscience*, *8*(6), 1122–1134. https://doi.org/10.1021/acschemneuro.7b00074

Uludağ, K., Müller-Bierl, B., & Uğurbil, K. (2009). An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage*, *48*(1), 150–165.

https://doi.org/10.1016/j.neuroimage.2009.05.051

Vaccaro, A. G., & Fleming, S. M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and Neuroscience Advances*, *2*, 1–14. https://doi.org/10.1177/2398212818810591

Van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. F. (2010). Unconscious activation of the prefrontal no-go network. *The Journal of Neuroscience*, *30*(11), 4143–4150. https://doi.org/10.1523/jneurosci.2992-09.2010

van Geuns, R.-J. M., Wielopolski, P. A., de Bruin, H. G., Rensing, B. J., van Ooijen, P. M. A., Hulshoff, M., Oudkerk, M., & de Feyter, P. J. (1999). Basic principles of magnetic resonance imaging. *Progress in Cardiovascular Diseases*, *42*(2), 149–156. https://doi.org/10.1016/S0033-0620(99)70014-9

Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: a discussion. *Metacognition and Learning*, *6*(2), 205–211. https://doi.org/10.1007/s11409-011-9080-x

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*(1), 3–14. https://doi.org/10.1007/s11409-006-6893-0

Vega, D., Torrubia, R., Marco-Pallarés, J., Soto, A., & Rodriguez-Fornells, A. (2020). Metacognition of daily self-regulation processes and personality traits in borderline personality disorder. *Journal of Affective Disorders*, *267*, 243–250. https://doi.org/10.1016/j.jad.2020.02.033

Villani, V., Tsakiris, M., & Azevedo, R. T. (2019). Transcutaneous vagus nerve stimulation improves interoceptive accuracy. *Neuropsychologia*, *134*, 107201. https://doi.org/10.1016/j.neuropsychologia.2019.107201

Vitzthum, F. B., Veckenstedt, R., & Moritz, S. (2014). Individualized Metacognitive Therapy Program for Patients with Psychosis (MCT+): Introduction of a Novel Approach for Psychotic Symptoms. *Behavioural and Cognitive Psychotherapy*, *42*(1), 105–110. https://doi.org/10.1017/S1352465813000246

Vlisides, P. E., Bel-Bahar, T., Nelson, A., Chilton, K., Smith, E., Janke, E., Tarnal, V., Picton, P., Harris, R. E., & Mashour, G. A. (2018). Subanaesthetic ketamine and altered states of consciousness in humans. *British Journal of Anaesthesia*, *121*(1), 249–259. https://doi.org/10.1016/j.bja.2018.03.011

Watkins, M. W., Lei, P.-W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, *35*(1), 59–68. https://doi.org/10.1016/j.intell.2006.04.005

Watts, M. E., Pocock, R., & Claudianos, C. (2018). Brain Energy and Oxygen Metabolism: Emerging Role in Normal Function and Disease. *Frontiers in Molecular Neuroscience*, *11*. https://doi.org/10.3389/fnmol.2018.00216

Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., Dolan, R. J., & Blakemore, S. J. (2013). The development of metacognitive ability in adolescence. *Consciousness and Cognition*, *22*(1), 264–271. https://doi.org/10.1016/j.concog.2013.01.004

Weiler, M. A., Thaker, G. K., Lahti, A. C., & Tamminga, C. A. (2000). Ketamine effects on eye movements. *Neuropsychopharmacology*, *23*(6), 645–653. https://doi.org/10.1016/S0893-133X(00)00156-1

Weingartner, H. J., Joyce, E. M., Sirocco, K. Y., Adams, C. M., Eckardt, M. J., George, T., & Lister, R. G. (1993). Specific memory and sedative effects of the benzodiazepine triazolam. *Journal of*

Psychopharmacology, *7*(4), 305–315. https://doi.org/10.1177/026988119300700401

Wells, A. (1995). Meta-cognition and worry: A cognitive model of generalized anxiety disorder. *Behavioural and Cognitive Psychotherapy*, *23*(3), 301–320. https://doi.org/10.1017/S1352465800015897

Wells, A. (1999). A Cognitive Model of Generalized Anxiety Disorder. *Behavior Modification*, *23*(4), 526–555. https://doi.org/10.1177/0145445599234002

Wells, A. (2005). The metacognitive model of GAD: Assessment of meta-worry and relationship with DSM-IV generalized anxiety disorder. *Cognitive Therapy and Research*, *29*(1), 107–121. https://doi.org/10.1007/s10608-005-1652-0

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: Properties of the MCQ-30. *Behaviour Research and Therapy*, *42*(4), 385–396. https://doi.org/10.1016/S0005-7967(03)00147-5

Wells, A., & Fisher, P. (2011). Meta-Cognitive Therapy Without Metacognition: A Case of ADHD. *American Journal of Psychiatry*, *168*(3), 327. https://doi.org/10.1176/appi.ajp.2010.10101467

Wells, A., Fisher, P., Myers, S., Wheatley, J., Patel, T., & Brewin, C. R. (2009). Metacognitive therapy in recurrent and persistent depression: A multiple-baseline study of a new treatment. *Cognitive Therapy and Research*, *33*(3), 291–300. https://doi.org/10.1007/s10608-007-9178-2

Wells, A., & Papageorgiou, C. (1998). Relationships between worry, obsessive–compulsive symptoms and meta-cognitive beliefs. *Behaviour Research and Therapy*, *36*(9), 899–913. https://doi.org/10.1016/S0005-7967(98)00070-9

Wells, A., & Sembi, S. (2004). Metacognitive therapy for PTSD: a preliminary investigation of a new brief treatment. *Journal of Behavior Therapy and Experimental Psychiatry*, *35*(4), 307–318. https://doi.org/10.1016/j.jbtep.2004.07.001

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, *18*(1), 1–10. https://doi.org/10.1111/j.2044-8317.1965.tb00689.x

Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*. https://doi.org/10.1016/j.cub.2022.07.010

Wierzchoń, M., Asanowicz, D., Paulewicz, B., & Cleeremans, A. (2012). Subjective measures of consciousness in artificial grammar learning task. *Consciousness and Cognition*, *21*(3), 1141–1153. https://doi.org/10.1016/j.concog.2012.05.012

Wilding, T. J., Lopez, M. N., & Huettner, J. E. (2016). Chimeric Glutamate Receptor Subunits Reveal the Transmembrane Domain Is Sufficient for NMDA Receptor Pore Properties but Some Positive Allosteric Modulators Require Additional Domains. *The Journal of Neuroscience*, *36*(34), 8815–8825. https://doi.org/10.1523/jneurosci.0345-16.2016

Wilsey, B., Deutsch, R., & Marcotte, T. D. (2016). Maintenance of Blinding in Clinical Trials and the Implications for Studying Analgesia Using Cannabinoids. *Cannabis and Cannabinoid Research*, *1*(1), 139–148. https://doi.org/10.1089/can.2016.0016

Wohleb, E. S., Wu, M., Gerhard, D. M., Taylor, S. R., Picciotto, M. R., Alreja, M., & Duman, R. S. (2016). GABA interneurons mediate the rapid antidepressant-like effects of scopolamine. *Journal of Clinical Investigation*, *126*(7), 2482–2494. https://doi.org/10.1172/JCI85033

Wokke, M. E., Achoui, D., & Cleeremans, A. (2020). Action information contributes to metacognitive decision-making. *Scientific Reports*, *10*(1), 3632. https://doi.org/10.1038/s41598-020-60382-y

Wokke, M. E., Cleeremans, A., & Ridderinkhof, K. R. (2017). Sure i'm sure: Prefrontal oscillations support metacognitive monitoring of decision making. *The Journal of Neuroscience*, *37*(4), 781–789. https://doi.org/10.1523/jneurosci.1612-16.2016

Wolff, K., & Winstock, A. R. (2006). Ketamine. *CNS Drugs*, *20*(3), 199–218. https://doi.org/10.2165/00023210-200620030-00003

Wright, N. D., Edwards, T., Fleming, S. M., & Dolan, R. J. (2012). Testosterone induces off-line perceptual learning. *Psychopharmacology*, *224*(3), 451–457. https://doi.org/10.1007/s00213-012-2769-y

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

Yoo, P. E., John, S. E., Farquharson, S., Cleary, J. O., Wong, Y. T., Ng, A., Mulcahy, C. B., Grayden, D. B., Ordidge, R. J., Opie, N. L., O'Brien, T. J., Oxley, T. J., & Moffat, B. A. (2018). 7T-fMRI: Faster temporal resolution yields optimal BOLD sensitivity for functional network imaging specifically at high spatial resolution. *NeuroImage*, *164*, 214–229. https://doi.org/10.1016/j.neuroimage.2017.03.002

Yu, A., & Lau, A. Y. (2018). Glutamate and Glycine Binding to the NMDA Receptor. *Structure*, *26*(7), 1035-1043.e2. https://doi.org/10.1016/j.str.2018.05.004

Zhang, M. W. B., & Ho, R. C. M. (2016). Controversies of the Effect of Ketamine on Cognition. *Frontiers in Psychiatry*, *7*, 47. https://doi.org/10.3389/fpsyt.2016.00047

Zhou, W., Wang, N., Yang, C., Li, X.-M., Zhou, Z.-Q., & Yang, J.-J. (2014). Ketamine-induced antidepressant effects are associated with AMPA receptors-mediated upregulation of mTOR and BDNF in rat hippocampus and prefrontal cortex. *European Psychiatry*, *29*(7), 419–423. https://doi.org/10.1016/j.eurpsy.2013.10.005

Zorumski, C. F., Izumi, Y., & Mennerick, S. (2016). Ketamine: NMDA Receptors and Beyond. *The Journal of Neuroscience*, *36*(44), 11158–11164. https://doi.org/10.1523/jneurosci.1547-16.2016

Zylberberg, A., Barttfeld, P., & Sigman, M. (2012). The construction of confidence in a perceptual decision. *Frontiers in Integrative Neuroscience*, *6*(79), 1–10. https://doi.org/10.3389/fnint.2012.00079

VI. APPENDIX A: LIST OF FIGURES AND TABLES

Figures

Tables

VII. APPENDIX B: LIST OF ABBREVIATIONS

2AFC	Two-alternative-forced-choice
5D-ASC	Five-dimensional altered state of consciousness rating scale
AIC	Akaike information criterion
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
aPFC	Anterior prefrontal cortex
AUROC2	Area under the Type II receiver-operating-characteristic curve
BCIS	Beck cognitive insight scale
BIC	Bayesian information criterion
BOLD	Blood-oxygen-level-dependent
BPD	Borderline personality disorder
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CFQ	Cognitive failures questionnaire
CNS	Central nervous system
DXM	Dextromethorphan
fMRI	Functional magnetic resonance imaging
GABA	γ-aminobutyric acid
GAD	Generalized anxiety disorder
GMV	Gray matter volume
IFG	Inferior frontal gyrus
MAI	Metacognitive awareness inventory
MCQ-30	Metacognitions Questionnaire-30
MR	Magnetic resonance
MRI	Magnetic resonance imaging
NCCs	Neural correlates of consciousness
NMDA	*N*-methyl-D-aspartate
NMDAR	*N*-methyl-D-aspartate receptor
OCD	Obsessive-compulsive disorder
PCP	Phencyclidine
PFC	Prefrontal cortex
PPC	Posterior parietal cortex
RMSEA	Root mean square error of approximation
ROC	Receiver operating characteristic
SDT	Signal detection theory
SEM	Structural equation modeling
SRMR	Standardized root mean square residual
TE	Echo time
ToM	Theory of mind
TR	Repetition time

VIII. APPENDIX C: PUBLICATIONS INCLUDED IN THIS THESIS

The current thesis is based on three original publications (see Table 2), which are attached to this dissertation in a consecutive manner. To avoid violations of copyright, the article of Study 1 was removed from the appendix section of the electronic version. It can be found online using the following reference or in hardcopies of the present dissertation at the Universitäts- und Landesbibliothek Bonn.

Study 1	**Lehmann, M.**, Hagen, J., & Ettinger, U. (2022). Unity and diversity of metacognition. *Journal of Experimental Psychology: General, 151*(10), 2396–2417. doi: 10.1037/xge0001197

Study 2	**Lehmann, M.**, Neumann, C., Wasserthal, S., Schultz, J., Delis, A., Trautner, P., Hurlemann, R., & Ettinger, U. (2021). Effects of ketamine on brain function during metacognition of episodic memory. *Neuroscience of Consciousness, 7*(1), niaa028. doi: 10.1093/nc/niaa028

Neuroscience of Consciousness, 2021, 7(1): niaa028

doi: 10.1093/nc/niaa028
Research Article

Effects of ketamine on brain function during metacognition of episodic memory

Mirko Lehmann [1], Claudia Neumann[2], Sven Wasserthal[3], Johannes Schultz[4,5], Achilles Delis[2], Peter Trautner[4,5,6], René Hurlemann[3,7,8] and Ulrich Ettinger[1,*]

[1]Department of Psychology, University of Bonn, Bonn, Germany; [2]Department of Anesthesiology and Intensive Care Medicine, University Hospital Bonn, Bonn, Germany; [3]Department of Psychiatry and Division of Medical Psychology, University Hospital Bonn, Bonn, Germany; [4]Center for Economics and Neuroscience, University of Bonn, Bonn, Germany; [5]Institute for Experimental Epileptology and Cognition Research, University of Bonn Medical Center, Bonn, Germany; [6]Department for NeuroCognition, Life & Brain Center, Bonn, Germany; [7]Department of Psychiatry, School of Medicine & Health Sciences, University of Oldenburg, Oldenburg, Germany; [8]Research Center Neurosensory Science, University of Oldenburg, Oldenburg, Germany

*Correspondence address. Department of Psychology, Kaiser-Karl-Ring 9, 53111 Bonn, Germany. Tel: +49-228-73-4208; Fax: +49-228-73-62323, E-mail: ulrich.ettinger@uni-bonn.de

Abstract

Only little research has been conducted on the pharmacological underpinnings of metacognition. Here, we tested the modulatory effects of a single intravenous dose (100 ng/ml) of the *N*-methyl-D-aspartate-glutamate-receptor antagonist ketamine, a compound known to induce altered states of consciousness, on metacognition and its neural correlates. Fifty-three young, healthy adults completed two study phases of an episodic memory task involving both encoding and retrieval in a double-blind, placebo-controlled fMRI study. Trial-by-trial confidence ratings were collected during retrieval. Effects on the subjective state of consciousness were assessed using the 5D-ASC questionnaire. Confirming that the drug elicited a psychedelic state, there were effects of ketamine on all 5D-ASC scales. Acute ketamine administration during retrieval had deleterious effects on metacognitive sensitivity (meta-d′) and led to larger metacognitive bias, with retrieval performance (d′) and reaction times remaining unaffected. However, there was no ketamine effect on metacognitive efficiency (meta-d′/d′). Measures of the BOLD signal revealed that ketamine compared to placebo elicited higher activation of posterior cortical brain areas, including superior and inferior parietal lobe, calcarine gyrus, and lingual gyrus, albeit not specific to metacognitive confidence ratings. Ketamine administered during encoding did not significantly affect performance or brain activation. Overall, our findings suggest that ketamine impacts metacognition, leading to significantly larger metacognitive bias and deterioration of metacognitive sensitivity as well as unspecific activation increases in posterior hot zone areas of the neural correlates of consciousness.

Keywords: metacognition; confidence; ketamine; episodic memory; glutamate

Received: 22 May 2020; **Revised:** 14 November 2020. **Accepted:** 23 November 2020

© The Author(s) 2021. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Many of our thoughts in everyday life revolve around other thoughts, about something we said or a decision we made. It has been postulated that these *meta*-thoughts constitute a distinct feature of consciousness. According to Block (1995), consciousness can be divided into phenomenal consciousness, access consciousness, self-consciousness, and monitoring consciousness. The latter concerns metacognition, i.e., the ability to reflect upon our own thoughts and knowledge and to monitor the quality of ongoing cognitive or perceptual processes (Grimaldi *et al.* 2015). The link between metacognition and consciousness is based on the intuition that, if an individual is unable to reflect on a particular mental state, this state cannot be conscious and consequently, some kind of metacognition should accompany all conscious representations (Shea and Frith 2019).

Metacognition is frequently measured on a trial-by-trial-basis as participants indicate their level of confidence about the accuracy of a perceptual or mnestic judgment (Grimaldi *et al.* 2015). A second-order confidence rating (Type 2 response) is therefore based on a first-order judgment (Type 1 response). Measures of metacognitive sensitivity tap how well participants introspectively assess or monitor their own cognitive processes (Fleming and Lau 2014). By applying signal-detection-theory (SDT) methodology, metacognitive *sensitivity* (as meta-d′) can be quantified independently of interindividual differences in response tendencies (Maniscalco and Lau 2012). The meta-d′-framework also allows to control for the influence of primary task performance on metacognitive sensitivity (Maniscalco and Lau 2014): metacognitive *efficiency* (meta-d′/d′) represents the amount of signal strength available for the metacognitive process, expressed as a fraction of the amount of signal strength available for the Type 1 task (McCurdy *et al.* 2013). Finally, it is important to consider the general tendency for higher or lower confidence ratings, the so-called metacognitive *bias* (Fleming and Lau 2014).

But what is the neural basis of metacognition? By drawing on evidence from no-report paradigms, Koch *et al.* (2016) argue that the neural correlates of consciousness are primarily localized in a posterior cortical network labeled a “hot zone” for conscious functions. However, neuroimaging and lesion studies suggest that higher-order conscious functions such as metacognition may also engage a frontoparietal network (Rouault *et al.* 2018; Vaccaro and Fleming 2018).

A more complete understanding of the neural mechanisms of metacognition also requires insight into the underlying neurotransmitter systems. To date, very little is known about the pharmacology of metacognition. Recently, Hauser *et al.* (2017) revealed that blockade of noradrenaline led to increased metacognitive sensitivity with unchanged perceptual decision-making performance, whereas both perceptual discrimination and metacognition remained unaffected by dopamine blockade.

One neurotransmitter likely to mediate aspects of consciousness is the glutamatergic system. Antagonists at the *N*-methyl-D-aspartate (NMDA) glutamate receptor, such as phencyclidine or ketamine, provoke psychedelic states which are clearly distinct from a normal waking state of consciousness (Anis *et al.* 1983; Umbricht *et al.* 2002; Morris and Wallach 2014), characterized by dissociative experiencing including vigilance reduction, ego transcendence, disembodiment, and visual and sensory disturbances (Vlisides *et al.* 2018). The noncompetitive NMDA-receptor antagonist ketamine is dose-dependently used for the treatment of depression (Murrough *et al.* 2013) and general anesthesia (Kurdi *et al.* 2014; Sarasso *et al.* 2015); in addition, it is a well-established research tool with an excellent safety record in both clinical and experimental applications (Javitt *et al.* 2012; Doyle *et al.* 2013). Ketamine-induced psychotropic effects such as distorted sense of space and time, euphoria and out-of-body experiences have contributed to its abuse as a recreational drug (Schifano *et al.* 2008; Giorgetti *et al.* 2015). Based on findings that acute ketamine administration temporarily and reversibly induces a range of both positive (hallucinations, thought disorder, delusions) and negative (social withdrawal, emotional blunting) psychosis-like symptoms in otherwise healthy volunteers, the compound is also a widely used pharmacological model of schizophrenia (Krystal *et al.* 1994; Malhotra *et al.* 1996).

Ketamine effects on cognition include a selective degradation of episodic memory (Hetem *et al.* 2000; Morgan *et al.* 2004). In episodic memory tasks, participants typically encode word items, and later retrieve those items by writing down as many words as they can remember (*recall*) or indicate whether a given item had previously been encoded or not (*recognition*) (Honey *et al.* 2005b). Previous findings suggest that retrieval performance is disturbed when ketamine is administered during encoding but remains unimpaired when only recognition, but not encoding, takes place under the influence of ketamine (Oye *et al.* 1992; Hetem *et al.* 2000; Honey *et al.* 2005b). This effect may, however, also depend on the depth of semantic processing of the encoded items: Honey *et al.* (2005b) found that ketamine reduced retrieval performance only when items were encoded at an intermediate level of processing (LoP), not on deep or shallow levels. A functional magnetic resonance imaging (fMRI) study by Honey *et al.* (2005a) reported that ketamine affects brain function during retrieval even if encoding occurred prior to ketamine administration: ketamine was associated with attenuated left prefrontal cortical response to deeply encoded items, whereas anterior cingulate activation was reduced for incorrect compared to correct responses.

Even though growing research effort is directed towards identifying the neural underpinnings of metacognition, and previous studies have aimed at specifying the role of glutamate in various cognitive functions, the involvement of this neurotransmitter system in metacognition has not yet been examined. In this double-blind, placebo-controlled fMRI study, the primary aim was to investigate the role of the glutamate system in metacognition and its underlying neural activity by applying a psychotomimetic dose of ketamine. Confidence ratings were collected in an episodic memory framework, based on the dissociation of ketamine effects on encoding and retrieval as operationalized by Honey *et al.* (2005a).

Specifically, we applied a task in which differences in Type 2 responses should not be due to altered Type 1 performance, since ketamine was previously shown to leave episodic memory performance in deep and shallow encoding conditions unaffected (Honey *et al.* 2005b). Metacognitive sensitivity was quantified using the meta-d′-framework, which was previously shown to be sensitive to the effects of pharmacological challenges (Clos *et al.* 2019) and drug consumption (Sadeghi *et al.* 2017). We expected metacognitive sensitivity to be altered by ketamine in both study phases and further predicted ketamine to affect neural activity during both metacognitive confidence ratings and encoding. The secondary study aims included investigation of LoP effects on retrieval performance and metacognitive accuracy as well as confirmation of the subjective, phenomenological effects of ketamine by including a self-report measure of altered states of consciousness.

Materials and Methods

Participants

Fifty-three healthy, non-smoking, right-handed volunteers (aged 18–34, $M = 23.47$, $SD = 3.24$; 29 female) with normal or corrected to normal vision and native speaker level command of German language were recruited for this study. Exclusion criteria were as follows: prior experience with ketamine, history of psychiatric or neurological disorder, claustrophobia, metalliferous implants, pregnancy, positive drug test, under- or overweight (Body Mass Index: <17; ≥ 30), or consumption of any medication. Further medical contraindications for the administration of ketamine included hypertension and hyperthyroidism.

The study was approved by the Research Ethics Committee at the Department of Psychology, University of Bonn (approval number: 18-03-28). In accordance with this approval, data of the study are not stored on public repositories, but behavioral data are available as Supplementary materials, and fMRI data will be made available upon request. Materials, analysis scripts, and preregistration of the study are available in Open Science Framework (https://osf.io/numxs/).

Screening procedure

An online prescreening interview was conducted with individuals who responded to study advertisements. Those who met all inclusion criteria were invited for a screening visit in the laboratory, where the German version of the 5.0.0 MINI-International Neuropsychiatric Interview (Ackenheil *et al.* 1999), a urine drug screen (Drug-Screen Multi-5T, nal von minden GmbH) and, for females, a pregnancy test (NADAL hCG Pregnancy Test, nal von minden GmbH) were carried out to screen for exclusion criteria of psychiatric illness, drug abuse, and pregnancy. Measurements of height, bodyweight, and blood pressure were obtained. A medical questionnaire was used to exclude any current or past medical conditions, or any diagnosis of psychotic disorders among first-degree relatives. Additionally, the first five questions of the Structured Instrument for Prodromal Syndromes (SIPS 5.0) were included to rule out prodromal symptoms of schizophrenia (McGlashan *et al.* 2001). Suitable individuals were invited for assessment visits.

Study design

A double-blind, randomized, placebo-controlled between-subjects design was employed. Randomization lists were created independently for females and males. The study team carrying out the assessments was not involved in the process of randomization. An unblinded study anesthesiologist prepared the infusion solution and constantly monitored oxygen saturation and heart rate of the participants during the infusion. Twenty-four participants were administered a subanesthetic dose of racemic ketamine (Ketamin-Ratiopharm 500 injection solution, Ratiopharm, Ulm, Germany), 29 participants received a saline solution (0.9% sodium chloride).

Ketamine was administered as a 2 mg/ml solution with a constant target plasma level of 100 ng/ml by a bolus and continuous infusion using a computerized infusion pump (Graseby 3500, Smith Medical Int. Ltd, Luton, UK). The solutions were administered using the STANPUMP program (Steven Shafer, M.D., Anesthesiology Service, PAVAMC 3801 Miranda Ave., Palo Alto, USA) based on the three-compartment model described by Domino *et al.* (1982). Previous studies of our group (Steffens *et al.* 2016, 2018) using the same infusion equipment and procedure confirmed that ketamine concentrations were close to the targeted plasma level and no residual traces of ketamine solution from the infusion site contaminated the results; therefore, no blood samples were drawn in this study.

General procedure

On assessment days, participants were required to refrain from solid food for 6 h and clear fluids for 2 h before the infusion. Within 24 h before, participants were also instructed to take no medication and to stay abstinent from alcohol. Female participants took another pregnancy test on the day of assessment. After participants arrived, they completed the first study task (see below) before an additional medical screening was performed by the study anesthesiologist. Participants were then fitted with intravenous access into the nondominant arm and positioned in the MRI scanner. Following an individual adjustment of the field of view and an initial high-resolution structural imaging scan, the infusion was started.

Ketamine effects on metacognition, encoding, and retrieval in an episodic memory task were assessed in two separate study-test phases. Stimuli were selected from the Berlin Affective Word List (Võ *et al.* 2009); word class, frequency, emotionality, arousal level, number of syllables, and vividness were counterbalanced between conditions.

In Study Phase I, items were presented on a computer screen outside the MRI scanner, prior to drug infusion. Retrieval was tested ~60 min after the end of the first encoding task, while BOLD data were acquired during infusion. In this first retrieval task, participants responded to stimuli by categorizing them either as "old items", if they had previously been presented in the encoding task, or "new items", if they had not been presented, and afterwards reported their metacognitive confidence (Type 2 response). Subsequently, in Study Phase II, another word list consisting entirely of novel items was encoded, as participants were still undergoing infusion in the MRI scanner. Retrieval of these items was tested ~60 min after the infusion was terminated and participants had left the scanner. Immediately upon leaving the scanner, participants completed the 5D-ASC questionnaire to assess altered states of consciousness (Dittrich 1998). In the second retrieval task, items encoded in the second encoding task ("old items") were again presented on a computer screen alongside "new items", again requiring participants to state their confidence after each Type 1 response. Figure 1 provides an overview of the general procedure of assessment days.

Task design

Study Phase I

Participants were presented with a total of 120 word items displayed in the center of a computer screen and were instructed to make one of two types of judgments about these items, which served as a manipulation of the depth of processing. We aimed for two levels of processing (deep/shallow) and selected a manipulation that could be expected to yield a pronounced LoP effect (Honey *et al.* 2005b). For each of 60 word items, participants indicated their subjective judgment of the pleasantness (pleasant/unpleasant) of the word (leading to deep encoding), whereas the other 60 items were encoded in a shallow manner, by participants reporting the number of syllables of each word (even/odd). Participants were not told that the retrieval of these items would be tested afterwards. These encoding tasks alternated blockwise, with each of four blocks comprising 30 items;

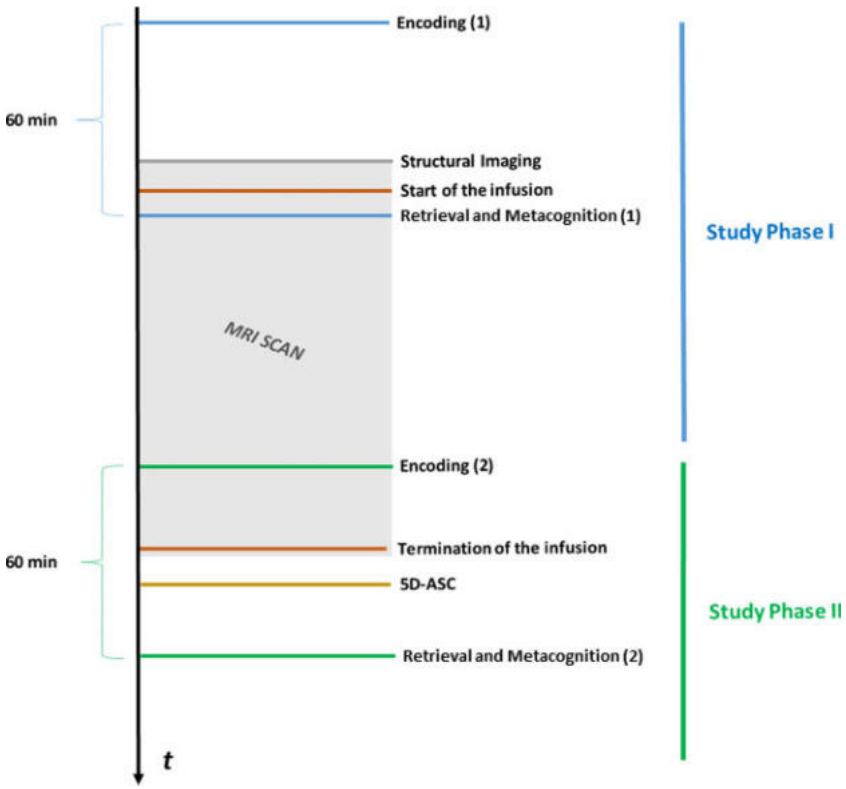

Figure 1. Study protocol. In Study Phase I (shown in blue), participants first encoded word items in the absence of infusion and before entering the MRI scanner. After a medical screening (in purple), participants completed a structural scan (in gray). Following the start of the infusion (in brown), retrieval of encoded items and corresponding metacognitive confidence was tested. As participants were still undergoing infusion in the MRI scanner, in Study Phase II (in green), participants encoded a second word list which was later retrieved outside the scanner, after termination of the infusion. Questionnaire data were collected using the 5D-ASC (in yellow). The MRI scanning period is represented by the grey box.

the starting condition was determined randomly. Items were presented until keypress for a maximum of 3 s, with an interstimulus interval (ISI) of 0.5 s.

The fMRI retrieval task was implemented in an event-related design. Participants responded to items presented on the center of a monitor behind the MRI scanner via a mirror by predefined button presses. A total of 180 word items were used, including the 120 items that had been encoded in the previous task as well as 60 new items. The 2:1 ratio of old to new items was based on previous studies (Honey *et al.* 2005a). Items were presented in randomized order for a duration of 2.5 s followed by an ISI that varied randomly between 2 s and 6 s; participants were instructed to respond to items which they considered to be old, i.e., having previously been presented, with a left index finger button press and to items which they labeled as new with a right index finger button press.

There were two types of second-order ratings: subsequent to 120 of these Type 1 responses, participants rated their subjective confidence regarding the judgment on a 6-point Likert scale (1 = "not confident at all", 6 = "very confident"). In this "Report" condition, designed to tap metacognitive processes, participants moved a cursor along the scale, using their index fingers, until they reached the position on the scale that most accurately matched their subjective confidence, which they were instructed to confirm by a left or right thumb press. During the 60 "Follow" trials which served as a control condition not involving the actual process of confidence formation (Yokoyama *et al.* 2010; Fleming *et al.* 2012), participants were instructed to navigate the cursor towards a predefined number on the scale, highlighted in blue. The initial position of the cursor was random in each condition; there were no written labels to either point of the scale to avoid extreme responding bias (Overgaard *et al.* 2006). "Report" and "Follow" trials alternated in randomized order; exactly two-third of each of the episodic memory condition trials (deep/shallow/new) were followed by the "Report" condition. The duration of the decision window for this second-order response was 3.5 s, followed by a 0.5 s screen where a change in cursor color from white to red highlighted the participant's response. Another variable ISI (2–6 s) preceded the onset of the next trial. In order to minimize exhaustion, the experiment was paused halfway through the task and a separate scan was started for the second half of the experiment. Figure 2 provides an illustration of the task.

Study Phase II

Following the completion of this first retrieval task, participants remained in the scanner and performed a second encoding task. Here, they were presented 100 novel word items in a block design; again, 50% of the items were encoded deeply by rating the subjective pleasantness of each word, whereas 50% of the items were encoded in a shallow manner by reporting the number of syllables. Again, encoding tasks alternated blockwise, with 10 blocks each comprising 10 items. At the beginning of each block, an instruction about the upcoming task was shown for 2 s. Participants responded via left or right button presses within a 3 s window (ISI = 0.5 s) for each item.

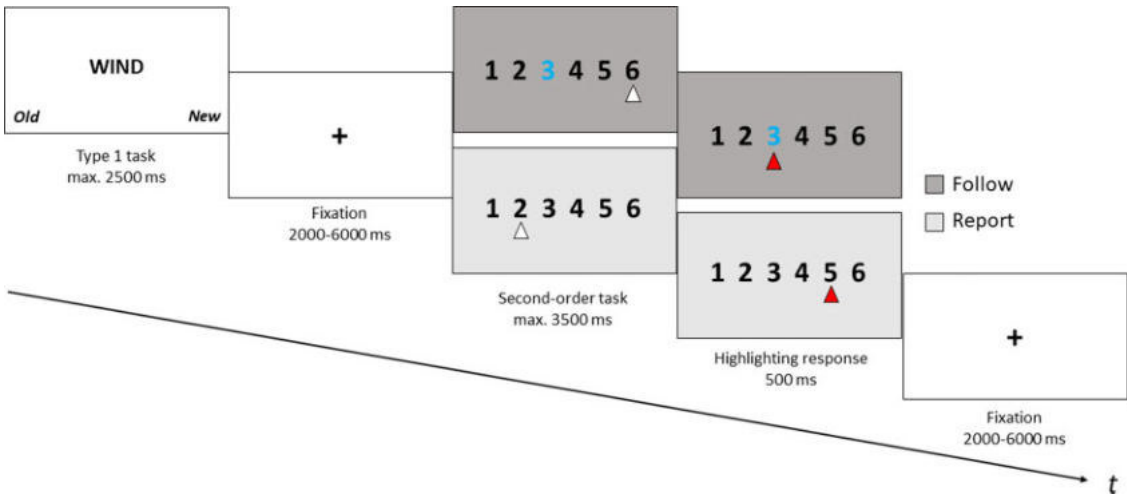

Figure 2. Schematic trial representation for the first retrieval task (stimuli are not to scale). Each trial consisted of two parts: first, participants categorized a presented word stimulus either as old (presented in the previous encoding task) or new (not having been presented before) (Type 1 task). Subsequently, they either indicated their subjective confidence ("Report" condition, shown in white) or placed the cursor at a color-coded position on the scale ("Follow" condition, grey) (Second-order task). The second retrieval task was similar, only here, the second-order task consisted entirely of "Report" trials, and the fixation period between task screens was shorter (1000 ms).

After termination of infusion and leaving the scanner, participants filled in the 5D-ASC, marking their extent of agreement with statements regarding various phenomenal experiences (Dittrich 1998). The 5D-ASC is a self-report questionnaire to retrospectively assess five dimensions of altered states of consciousness. These include three primary, etiology-independent scales, "Oceanic Boundlessness", "Dread of Ego Dissolution", and "Visionary Restructuralization", which can be conflated to a global measure of altered consciousness, and two secondary, etiology-specific scales comprising further aspects of altered experiences, "Auditory Alterations" and "Vigilance Reduction". 5D-ASC scale scores were formed following guidelines by Dittrich *et al.* (2006).

One hour after completion of the second encoding task, retrieval of those items was tested in a second retrieval task, without infusion at a time when plasma levels of ketamine are significantly reduced (Honey *et al.* 2005b). The design of the second retrieval task was almost identical to the first one, with two exceptions: ISI was constant (1 s), and there was no "Follow" condition, so participants had to report their confidence on each of the 150 trials (100 old, 50 new).

fMRI data acquisition and analysis

Imaging was conducted using a 1.5 T Avanto MRI scanner (Siemens, Erlangen, Germany). High-resolution structural images were acquired to optimize normalization of functional imaging data using a T1-weighted 3D MPRAGE sequence [Repetition time (TR) = 1660 ms, echo time (TE) = 3.09, inversion time = 800 ms, matrix size = 256 × 256, slice thickness = 1.0 mm, FoV = 256 mm, flip angle = 15°, voxel size = 1 × 1 × 1 mm^2, 160 sagittal slices]. Task-related BOLD fMRI data were acquired using a T2*-weighted echo-planar imaging sequence (TR = 2500 ms; TE = 45 ms, matrix size = 64 × 64, slice thickness = 3.0 mm, FoV = 192 mm, flip angle = 90°, voxel size = 3 × 3 × 3 mm, 31 slices). A standard 12-channel head coil was used for radio frequency transmission and reception.

fMRI data were analyzed using Statistical Parametric Mapping 12 software (Wellcome Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm) implemented in Matlab R2014a (The MathWorks Inc., Natick, USA). To allow for T1 equilibration, the first five volumes of each functional time series were discarded. Each participant's structural image was segmented into gray matter, white matter, and cerebro-spinal fluid using a forward deformation field to map it onto template tissue probability maps (Ashburner and Friston 2005). Functional images were realigned to the first image of each time series to correct for head movement, using a six-parameter rigid body transformation. The realigned functional images were then coregistered to the anatomical image. For spatial normalization, functional scans were transformed into standard stereotaxic space of the Montreal Neurological Institute (MNI) template (Evans *et al.* 1992; Holmes *et al.* 1998) and resampled at 2 × 2 × 2 mm voxel size. Finally, images were spatially smoothed using an 8 mm full-width-at-half-maximum Gaussian kernel.

Following pre-processing, at the first (single-subject) level for Study Phase I, the onset of each stimulus was defined as the onset of the event; for Type 1 responses, the duration was set to be the reaction time from stimulus presentation to button press. For second-order responses, the function spanned the time from onset of scale presentation to the first movement participants made on the scale. This was done as the decisive metacognitive processes during Report trials were expected to take place during that time, and to eliminate motion-related activation. The realignment parameters were added to the model as covariates of no interest. Correctly retrieved deep, shallow and new items were included as Type 1 regressors; since there were too few cases of incorrect answers in the majority of participants, an overall residual regressor of no interest was formed for incorrect answers, thereby departing from our preregistered analysis plan.

Overall, there were four Type 1 regressors: "Deep" (*mean number of trials across participants*: 49, SD = 8.6); "Shallow" (M = 27.19, SD = 11.34); "New" (M = 47.02, SD = 11.65); and "Incorrect" (M = 41.85, SD = 9.33). For each of these four regressors, two separate regressors were included for second-order ratings, resulting in a total of eight second-order regressors: "DeepReport" (M = 29.92, SD = 8.05); "DeepFollow" (M = 17.68,

SD = 7.58); "ShallowReport" (M = 18.96, SD = 8.95); "ShallowFollow" (M = 7.53, SD = 4.7); "NewReport" (M = 26.72, SD = 6.69); "NewFollow" (M = 13.17, SD = 3.53); "IncorrectReport" (M = 31.02, SD = 9.04); "IncorrectFollow" (M = 13.94, SD = 5.03). All contrasts were estimated by comparing specific effects against the baseline of the respective first-level-model; consequently, the two separate runs were conflated in this step. Additionally, we set up an exploratory first-level-model, in which "Report" regressors were parametrically modulated by the selected confidence rating in each trial, all other regressors remaining unmodified, as only "Report" ratings were expected to require the engagement of metacognitive Type 2 evaluations.

For Study Phase II, the function spanned the time from onset of word presentation to button press. Here, a simpler model with conditions "Deep" and "Shallow" was specified. Also departing from preregistration, the factor "Retrieval Performance" (later correctly/incorrectly retrieved) could not be applied, as there was an insufficient amount of incorrect answers.

On the second level, a full factorial analysis was carried out on Study Phase I data using between-subjects factor "Drug" (ketamine/placebo) and within-subjects factor "Word Type" (deep/shallow/new) for Type 1 contrasts with an additional within-subjects-factor "Rating Type" (report/follow) for second-order contrasts. A separate full factorial analysis was conducted on Study Phase II data, using between-subjects-factor "Drug" (ketamine/placebo) and within-subjects-factor "Encoding Level" (deep/shallow).

All second-level analyses were conducted at the whole-brain-level. The statistical height threshold was $P < 0.001$, and significant clusters were inferred if the peak voxel of the cluster survived a statistical threshold of $P < 0.05$ family-wise-error (FWE) corrected (cluster-level). In order to assign anatomical labels, the anatomy toolbox was utilized (Eickhoff *et al.* 2005). To determine whether significant clusters of each contrast represented activations or deactivations, mean summary functions were created using MarsBaR (https://sourceforge.net/projects/marsbar).

BOLD data of four participants during Study Phase I and of three participants during Study Phase II had to be excluded from fMRI analysis because normalization failed. Consequently, fMRI data analysis was performed on 49 participants (23 ketamine, 26 placebo) for Study Phase I and on 50 participants (23 ketamine, 27 placebo) for Study Phase II. Behavioral data analysis was carried out on all 53 participants who completed the assessment.

Behavioral data analysis

Following our preregistration, Type 1 (retrieval) and Type 2 (metacognitive) performance was assessed in an SDT framework (Green and Swets 1966; Barrett *et al.* 2013). We applied meta-d′ analysis (Maniscalco and Lau 2012) to quantify metacognitive *sensitivity*—i.e., the individual ability to discriminate between correct and incorrect retrieval judgments. Meta-d′ represents a response-bias free measure of how well confidence ratings track task accuracy and is on the same scale as the Type 1 sensitivity measure d′ (Maniscalco and Lau 2014). Meta-d′ was estimated in a maximum-likelihood-estimation model using code by Maniscalco (http://www.columbia.edu/~bsm2105/type2sdt) in Matlab R2016a (The MathWorks Inc., Natick, USA); only "Report" trials in which participants provided button presses on both retrieval and confidence rating were used for calculation. Additionally, metacognitive *efficiency* was calculated by dividing meta-d′ by d′ to provide an index of Type 2 performance that takes into account differences in Type 1 performance (Fleming and Lau 2014). To evaluate Type 2 performance, we therefore considered both absolute Type 2 sensitivity (meta-d′) and Type 2 efficiency relative to Type 1 performance (meta-d′/d′).

In addition to our preregistered analyses, we also conducted various exploratory analyses to facilitate mechanistic understanding of the outcomes. For example, we decided to expand our analysis to investigate ketamine effects on performance-corrected metacognitive bias (quantified as *mean judgment minus mean performance*) to test for differences in the selected confidence ratings between the two groups while controlling for the confounding influence of performance on confidence levels (Fleming and Lau 2014). Moreover, we explored Pearson's correlations between Type 1 and both Type 2 performance measures as well as metacognitive bias in both study phases with the 5D-ASC global measure of altered consciousness; alpha-level was Bonferroni-corrected (α=.05/8=.006). Finally, we applied an extension of the HMeta-d toolbox (Fleming 2017), a hierarchical Bayesian estimation of metacognitive efficiency (https://github.com/metacoglab/HMeta-d) in Matlab R2016a, which estimates group-level parameters over log(meta-d′/d′) while taking into account uncertainty in model fits at the single-subject level. To test for a true group difference in metacognitive efficiency, we fitted separate models for the ketamine and placebo group and calculated the 95% highest-density intervals (HDIs; the interval containing 95% of the Markov chain Monte Carlo posterior samples) on the difference between the group posterior densities and evaluated their potential overlaps with zero (Kruschke 2014). We ran three chains for estimation and ensured chain convergence (Fleming 2017).

All other behavioral data analyses were conducted using SPSS 22 (IBM Corp., Armonk, USA). Data were tested for violation of statistical assumptions; Kolmogorov–Smirnov tests were applied to test for normality of distribution, Mauchly's tests checked for sphericity, Levene's statistics tested for homogeneity of variances and Box-M-tests for homogeneity of covariances. When normality was violated in only one variable of a group, none of the variables were transformed. Drug effects on 5D-ASC scales, Type 1 and Type 2 reaction times and metacognitive bias were tested via independent samples *t*-tests. Paired *t*-tests were employed to compare Type 1 and Type 2 reaction times and metacognitive bias between deeply vs. shallowly encoded items. Separate mixed-design ANOVAs were employed with factors "Encoding Level" and the "Drug" for Type 1 and Type 2 sensitivity and Type 2 efficiency. Effect sizes for *t*-tests are given in Cohen's *d* (Cohen and Maydeu-Olivares 1992), effect sizes for ANOVAs in partial eta-squared (Cohen 1973).

Results

5D-ASC

There was a significant ketamine effect on the 5D-ASC global measure of altered consciousness [$t(23.7) = 4.69$, $P < 0.001$, $d = 1.35$] and on all scales. Participants who had received ketamine scored significantly higher on the three primary dimensions "Oceanic Boundlessness" [$t(23.23) = 4.04$, $P < 0.001$, $d = 1.17$], "Dread of Ego Dissolution" [$t(25.73) = 4.56$, $P < 0.001$, $d = 1.31$], and "Visionary Restructuralization" [$t(23.43) = 3.48$, $P = 0.002$, $d = 1.01$]. They also achieved significantly higher values on the "Auditory Alterations" [$t(28.17) = 4.55$, $P < 0.001$, $d = 1.29$]

Table 1. Descriptive statistics of 5D-ASC questionnaire scores by drug.

Scale	Placebo ($n = 29$)		Ketamine ($n = 24$)	
	M	SD	M	SD
[Global Index of Altered States]	1.08	1.87	14.51	13.91
Oceanic Boundlessness	0.71	1.48	16.63	19.25
Dread of Ego Dissolution	2.05	3.26	13.71	12.18
Visionary Restructuralization	0.52	1.74	12.27	16.45
Auditory Alterations	1.85	4.63	14.14	12.56
Vigilance Reduction	12.58	14.04	47.71	25.75

Note: Scale values are given in percent. M, mean; SD, standard deviation.

and "Vigilance Reduction" scales [$t(34.01) = 5.99$, $P < 0.001$, $d = 1.69$]. Descriptive statistics are provided in Table 1.

Exploratory analyses

There were no significant correlations of the 5D-ASC global measure of altered consciousness with Type 1 and Type 2 outcomes in either study phase (all $P > 0.006$).

Study Phase I

Descriptive statistics of Type 1 and Type 2 measures for Study Phase I are provided in Table 2. Distribution plots of raw data for all relevant dependent variables can be found in the Supplementary materials.

Type 1 behavioral analyses

The LoP manipulation was successful: participants showed significantly enhanced retrieval performance for deeply compared to shallowly encoded items [main effect of "Encoding Level": $F(1,51) = 241.44$, $P < 0.001$, $\eta_p^2 = 0.83$]. However, there was no main effect of "Drug" [$F(1,51) = 1.78$, $P = 0.188$, $\eta_p^2 = 0.03$]; ketamine did not significantly alter retrieval performance. Type 1 reaction times were significantly shorter for deeply than shallowly encoded items [$t(52) = 9.17$, $P < 0.001$, $d = 0.71$] but were unaffected by ketamine [$t(51) = 0.04$, $P = 0.972$, $d < 0.01$]. There were no significant interactions ($P > 0.05$).

Type 1 fMRI analyses

For BOLD data during retrieval, there was no significant difference between ketamine and placebo ($P > 0.05$). For a detailed summary of LoP and Old vs. New effects, see Supplementary materials.

Type 2 behavioral analyses

Participants showed enhanced metacognitive sensitivity for deeply compared to shallowly encoded items [$F(1,51) = 186.36$, $P < 0.001$, $\eta_p^2 = 0.79$]. Importantly, there was a significant main effect of "Drug" [$F(1,51) = 4.64$, $P = 0.036$, $\eta_p^2 = 0.08$]: metacognitive sensitivity deteriorated under ketamine. However, there was no significant main effect of either "Drug" [$F(1,50) = 1.03$, $P = 0.315$, $\eta_p^2 = 0.02$) or "Encoding Level" [$F(1,50) = 2.17$, $P = 0.147$, $\eta_p^2 = 0.04$] on metacognitive efficiency. Type 2 reaction times were faster for deeply encoded items [$t(52) = 4.25$, $P < 0.001$, $d = 0.41$] but were found to be unaltered by "Drug" [$t(51) = 0.03$, $P = 0.98$, $d < 0.01$]. There were no significant interactions ($P > 0.05$).

Table 2. Descriptive statistics of study phase I sensitivity measures (type 1 and type 2) and reaction times (type 1 and type 2) by drug and encoding level.

Measure	Placebo ($n = 29$)		Ketamine ($n = 24$)	
	M	SD	M	SD
Type 1 performance (d')[a]				
Deep vs. new	2.11	0.63	1.94	0.49
Shallow vs. new	0.85	0.38	0.74	0.37
Type 2 sensitivity (meta-d')[a,b]				
Deep vs. new	2.41	0.95	2.06	0.72
Shallow vs. new	0.89	0.49	0.58	0.39
Type 2 efficiency (meta-d'/d')				
Deep vs. new	1.17	0.38	1.13	0.48
Shallow vs. new	1.15	0.61	0.92	0.69
Type 1 reaction times (in ms)[a]				
Deep	1415.66	216.69	1458.75	189.0
Shallow	1572.63	197.31	1576.0	174.66
New	1600.11	227.04	1549.02	152.33
Type 2 reaction times (in ms)[a]				
Deep	1592.22	210.58	1578.45	221.79
Shallow	1670.7	229.7	1698.43	297.13
New	1755.37	251.38	1731.03	284.39

M, mean; ms, milliseconds; SD, standard deviation.
[a]Significant effects of encoding level.
[b]Significant effects of drug.

Exploratory analyses. Hierarchical Bayesian estimation of group-level meta-d′/d′ confirmed that we cannot be certain that there is a true difference in metacognitive efficiency between the two groups, even though the estimated difference between groups was relatively high [mean: 0.23 (highest-density interval: −0.04 to 0.58)]. Figure 3 provides an illustration of the estimated group-level parameters of metacognitive efficiency.

There was also a significant effect of "Drug" on metacognitive bias scores [$t(51) = 2.15$, $P = 0.037$, $d = 0.59$), with participants under ketamine being overconfident. In addition, there was a significant effect of "Encoding Level" on metacognitive bias, with ratings for shallowly encoded items reflecting overconfidence [$t(48) = 7.25$, $P < 0.001$, $d = 1.24$].

Second-order fMRI analyses

Report vs. follow effects. Higher BOLD responses during Report than Follow were found in a right visual cluster of right calcarine and lingual gyrus (Figure 4, Table 3). The cluster furthermore encompassed left and right cuneus, as well as bilateral superior occipital gyrus. A second, left-hemispheric, cluster was located in the posterior medial frontal cortex (pMFC).

The reverse effect (Follow>Report, indicating BOLD responses that were higher when participants had to select a predefined specification on the scale) revealed a total of 11 clusters (Figure 4, Table 4). These correspond to the default-mode network (DMN) that is active in the absence of task demands (Andrews-Hanna 2012), which encompasses angular gyrus, precuneus, posterior cingulate cortex (PCC), superior frontal areas, and parahippocampal gyrus, all of which were activated in the contrast.

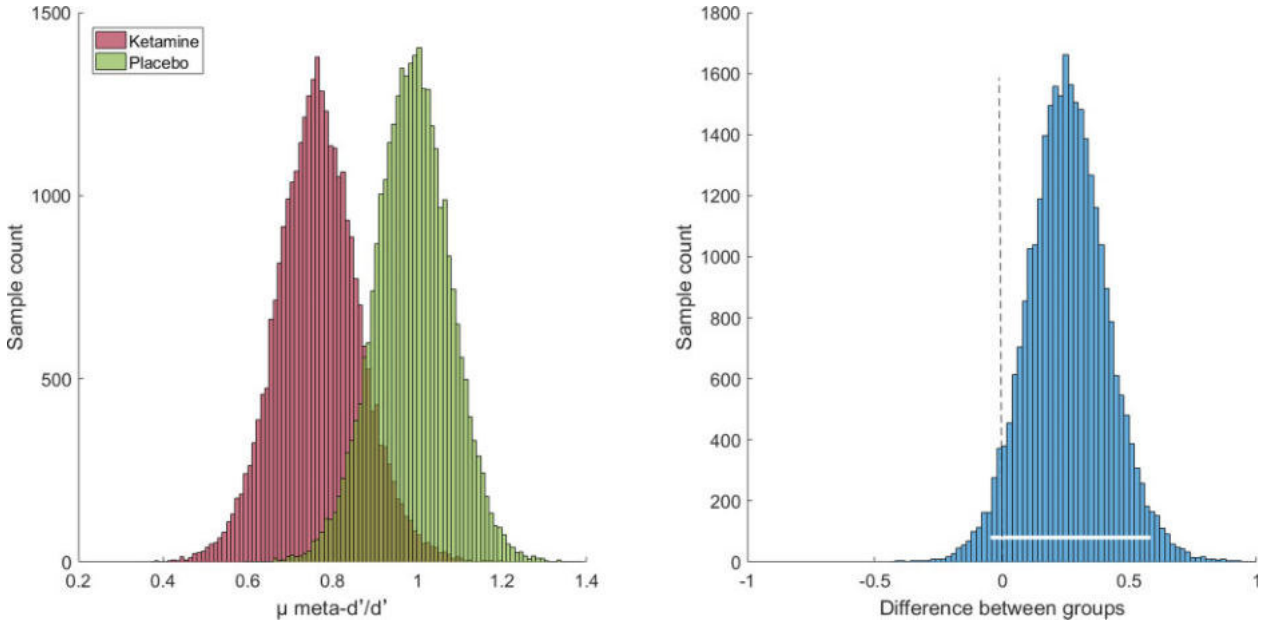

Figure 3. Hierarchical Bayesian estimation of metacognitive efficiency ($\mu_{\text{meta-d'/d'}}$) in Study Phase I. Left panel: Group-level values for the ketamine group (red histogram) and the placebo group (green histogram). Right panel: Difference in group posteriors (in log units). The white bar indicates the 95% highest-density interval which narrowly overlaps with zero.

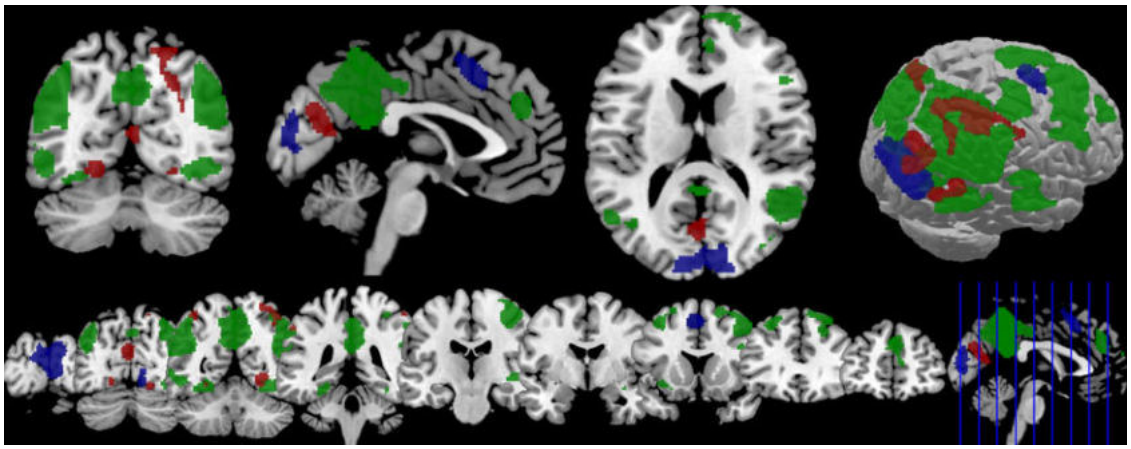

Figure 4. Second-order fMRI results. Significant activation in the contrasts Ketamine>Placebo (red), Report>Follow (blue) and Follow>Report (green) at significance level $P < 0.001$ (uncorrected).

Table 3. Summary of significant clusters for the report > follow contrast.

Anatomical label	Laterality	Cluster size [k]	T-value	Peak voxel MNI coordinates		
Calcarine gyrus	R	1599	5.33	8	−86	4
Lingual gyrus	R		5.24	12	−80	−8
Cuneus	R		4.88	8	−86	26
Cuneus	L		4.61	−6	−94	22
Superior occipital gyrus	L		4.58	−10	−96	20
Superior occipital gyrus	R		3.81	18	−96	18
pMFC	L	352	6.08	−4	16	48

Combined sample. Only unique anatomical labels are reported for each cluster at one laterality. Whole-brain cluster-level FWE corrected ($P < 0.001$ uncorrected).
FWE, family-wise error; L, left; MNI, Montreal Neurological Institute; pMFC, posterior medial frontal cortex; R, right.

Table 4. Summary of significant clusters for the follow > report contrast.

Anatomical label	Laterality	Cluster size [k]	*T*-value	Peak voxel MNI coordinates		
Angular gyrus	R	3662	9.36	56	−52	36
Superior parietal lobule	R		4.8	36	−44	58
Middle occipital gyrus	R		4.53	36	−80	10
Postcentral gyrus	R		3.36	24	−44	66
Precuneus	R	3384	7.18	10	−50	38
PCC	R		6.96	4	−48	28
MCC	R		6.85	10	−44	32
Precuneus	L		5.46	−6	−52	44
MCC	L		5.11	−4	−46	48
Superior frontal gyrus	R	2630	7.22	4	46	30
Middle frontal gyrus	R		5.54	30	24	54
IFG (p. Triangularis)	R		4.53	48	24	24
Inferior parietal lobule	L	2548	7.14	−54	−54	36
Angular gyrus	L		5.97	−40	−72	38
Supramarginal gyrus	L		4.37	−62	−36	38
Middle occipital gyrus	L		4.05	−36	−80	28
Fusiform gyrus	L	1094	7.34	−30	−52	−16
Inferior temporal gyrus	L		4.61	−54	−54	−8
Middle temporal gyrus	L		4.07	−60	−50	−2
Parahippocampal gyrus	L		3.32	−22	−28	−18
Fusiform gyrus	R	807	7.5	30	−52	−16
Inferior occipital gyrus	R		5.35	36	−72	−10
Inferior temporal gyrus	R		3.88	52	−64	−8
Precentral gyrus	R	753	5.67	38	−22	54
Middle temporal gyrus	R	682	5.5	60	−20	−10
Posterior insula	R		4.49	34	−6	−12
Insula lobe	R		4.41	40	−18	−2
Superior temporal gyrus	R		3.45	50	−12	−10
Middle frontal gyrus	L	674	5.57	−32	24	50
Superior frontal gyrus	L		4.3	−22	22	56
Superior frontal gyrus	R	203	4.48	14	66	16
Anterior insula	L	162	4.99	−28	6	−14
Insula lobe	L		4.49	−32	16	−12

Combined sample. Only unique anatomical labels are reported for each cluster at one laterality. Whole-brain cluster-level FWE corrected ($P < 0.001$ uncorrected).
FWE, family-wise error; IFG, inferior frontal gyrus; L, left; MCC, midcingulate cortex; MNI, Montreal Neurological Institute; PCC, posterior cingulate cortex; R, right.

Table 5. Summary of significant clusters for the ketamine > placebo contrast.

Anatomical label	Laterality	Cluster size [k]	*T*-value	Peak voxel MNI coordinates		
Superior parietal lobule	R	642	5.51	36	−52	64
Supramarginal gyrus	R		3.56	60	−28	48
Middle occipital gyrus	R		3.32	30	−64	33
Inferior parietal lobule	R		3.26	40	−54	48
Angular gyrus	R		3.23	36	−56	48
Calcarine gyrus	L	257	4.59	−2	−72	18
Lingual gyrus	R	212	4.42	18	−70	−10
Inferior parietal lobule	L	188	4.24	−40	−52	60
Lingual gyrus	L	172	5.21	−18	−68	−8
Fusiform gyrus	L		3.98	−28	−52	−12

Only unique anatomical labels are reported for each cluster at one laterality. Whole-brain cluster-level FWE corrected ($P < 0.001$ uncorrected).
FWE, family-wise error; L, left; MNI, Montreal Neurological Institute; R, right.

Table 6. Descriptive statistics of study phase II sensitivity measures (type 1 and type 2) and reaction times (type 1 and type 2) by drug and encoding level.

Measure	Placebo ($n = 29$)		Ketamine ($n = 24$)	
	M	SD	M	SD
Type 1 performance (d')[a]				
Deep vs. new	1.79	0.58	1.59	0.54
Shallow vs. new	0.68	0.41	0.57	0.31
Type 2 sensitivity (meta-d')[a]				
Deep vs. new	1.97	0.78	1.77	0.68
Shallow vs. new	0.44	0.41	0.4	0.58
Type 2 efficiency (meta-d'/d')[a]				
Deep vs. new	1.08	0.35	1.2	0.6
Shallow vs. new	0.69	0.65	0.71	1.19
Type 1 reaction times (in ms)[a]				
Deep	1267.99	235.79	1233.3	187.04
Shallow	1368.44	245.62	1306.19	186.04
New	1416.41	268.68	1286.1	187.95
Type 2 reaction times (in ms)[a]				
Deep	1132.52	213.1	1077.98	300.41
Shallow	1194.43	253.64	1112.9	301.18
New	1228.7	290.89	1117.47	272.8

M, mean; ms, milliseconds; SD, standard deviation.
[a]Significant effects of encoding level.

Drug effects. During second-order ratings (both Report and Follow), there was larger BOLD with ketamine than placebo in five clusters (Figure 4, Table 5): The first, right-hemispheric, cluster included superior parietal lobule (SPL), supramarginal gyrus, inferior parietal lobule (IPL), and angular gyrus. A second cluster was located in left calcarine gyrus, a third cluster in right lingual gyrus. The fourth cluster included left IPL, whereas a fifth, left-hemispheric cluster encompassed lingual gyrus and fusiform gyrus. There were no significant effects for the reverse contrast and no significant interactions ($P > 0.05$).

Exploratory analyses. Parametric modulation analysis ("Report" trials parametrically modulated by the selected confidence rating) revealed very similar results, i.e., higher BOLD response for ketamine than placebo in bilateral lingual, fusiform, and calcarine gyrus and right SPL (see Supplementary Table 6). There were no significant effects for the reverse contrast and no significant interactions ($P > 0.05$).

Study Phase II

Encoding: fMRI analyses

There were no significant ketamine effects on BOLD during encoding ($P > 0.05$). For LoP effects, see Supplementary materials.

Type 1 behavioral analyses

Descriptive statistics of Type 1 and Type 2 measures for Study Phase II are provided in Table 6. Distribution plots of raw data for all relevant dependent variables can be found in the Supplementary materials.

Items that had been encoded deeply were recognized more often than shallowly encoded items [significant main effect of "Encoding Level": $F(1,51) = 273.94$, $P < 0.001$, $\eta_p^2 = 0.85$]. There was no significant effect of "Drug" on d′ [$F(1,51) = 1.8$, $P = 0.185$, $\eta_p^2 = 0.04$]. "Drug" also had no effect on Type 1 reaction times [$t(51) = 1.29$, $P = 0.203$, $d = 0.36$]; when deeply encoded items were presented, participants made significantly quicker button presses [$t(52) = 5.7$, $P < 0.001$, $d = 0.4$]. There were no significant interactions ($P > 0.05$).

Type 2 behavioral analyses

There were significant main effects of "Encoding Level" on metacognitive sensitivity [$F(1,50) = 263.38$, $P < 0.001$, $\eta_p^2 = 0.84$] and metacognitive efficiency [$F(1,49) = 18.01$, $P < 0.001$, $\eta_p^2 = 0.27$), but no effects of "Drug" on either meta-d′ [$F(1,50) = 0.655$, $P = 0.422$, $\eta_p^2 = 0.01$] or metacognitive efficiency [$F(1,49) = 0.16$, $P = 0.691$, $\eta_p^2 < 0.01$]. Type 2 reaction times were also significantly shorter for deeply encoded items [$t(51) = 2.68$, $P = 0.01$, $d = 0.19$], but there was no effect of "Drug" [$t(50) = 1.13$, $P = 0.264$, $d = 0.34$). There were no significant interactions ($P > 0.05$).

Exploratory analyses. Corresponding to overlaps of 95% HDIs with zero, we found no significant group difference in metacognitive efficiency between ketamine and placebo [0.03 (−0.35 to 0.043)]. Thus, there was no significant ketamine effect on any measure of Type 2 performance when retrieval took place after termination of the infusion. We did, however, observe a significant effect of "Drug" on metacognitive bias [$t(50) = 2.75$, $P = 0.008$, $d = 0.76$], with participants under ketamine displaying overconfidence. There was also significantly larger metacognitive bias for shallowly than for deeply encoded words [$t(50) = 9.31$, $P < .001$, $d = 1.33$].

Discussion

This study investigated the role of the glutamate system in metacognition and associated brain activity using a ketamine challenge during episodic memory tasks in two study phases.

Study Phase I

Task effects

For a detailed discussion on LoP effects both at the behavioral and the brain functional level, see Supplementary materials.

Two clusters were significantly more active during Report than Follow; the first includes right calcarine gyrus, bilateral cuneus, and right lingual gyrus. The latter structure has been demonstrated to display increased functional connectivity with prefrontal cortex (PFC) in Report compared to Follow trials (Fleming *et al.* 2012). The second cluster in left pMFC provides further evidence for its role in metacognition and resembles the anatomically adjacent dorsal anterior cingulate cortex cluster which Fleming *et al.* (2012) found to be involved in reporting confidence in a similar task design. A recent meta-analysis (Vaccaro and Fleming 2018) identified bilateral pMFC as one of the prime neural correlates of metacognitive judgments, representing the biggest cluster in the composite meta-analysis of all metacognition-related activity and the second-biggest cluster associated with metacognitive ratings following memory judgments.

In the reverse contrast (Follow>Report), increased activation was found in brain regions involved in the DMN, which has been linked to introspective mental activities such as mind

wandering (Andrews-Hanna 2012). Again, this confirms Fleming *et al.* (2012), who reported similar patterns in this contrast.

Drug effects

As expected, subjective measures (5D-ASC) revealed that ketamine caused phenomenological experiences significantly deviating from a normal state of consciousness on all scales of the questionnaire. This finding confirms the known psychotomimetic effects of ketamine (Anis *et al.* 1983; Vlisides *et al.* 2018) and validates the rationale for using this pharmacological challenge to investigate the glutamatergic basis of metacognition.

Our study is one of only very few to indicate a potential pharmacological modulation of metacognitive performance (Lou *et al.* 2011; Hauser *et al.* 2017) and the first to investigate ketamine effects on metacognition. We show that disrupting the glutamatergic system by means of ketamine administration may challenge introspective monitoring processes: at the behavioral level, ketamine application during retrieval resulted in deterioration of metacognitive sensitivity (meta-d′) and overconfidence (larger metacognitive bias). Differences in metacognitive bias have been suggested to reflect genuine differences in awareness (Schwiedrzik *et al.* 2011), suggesting a role of various conscious processes giving rise to this ketamine effect on metacognitive bias. Furthermore, as overconfidence has been reported in patients with schizophrenia (Moritz *et al.* 2014), this finding provides another piece of evidence for use of ketamine as a model system of schizophrenia. Importantly, ketamine did not affect retrieval (Type 1) performance, in line with previous reports (Honey *et al.* 2005b), even though some group-heterogeneity has to be considered in Type 1 performance. Additionally, both Type 1 and Type 2 reaction times were unaffected by ketamine, which also indicates that the drug did not lead to a general deterioration of cognitive performance.

However, when controlling for the influence of Type 1 performance (d′) on metacognitive sensitivity (meta-d′) by calculating metacognitive efficiency (meta-d′/d′), there was no significant group difference. It is advised to apply metacognitive efficiency measures when comparing different groups (Fleming and Lau 2014; Vaccaro and Fleming 2018) although the theoretical assumption of the relationship of Type 1 and Type 2 performance measures (Galvin *et al.* 2003; Maniscalco and Lau 2012) is frequently violated in cases of "hyper"-metacognitive efficiency (meta-d′/d′ $>$ 1), potentially arising as a consequence of post-decisional and/or second-order computation (Fleming and Daw 2017) as evidence continues to be accumulated after the Type 1 response (Murphy *et al.* 2015; Rausch and Zehetleitner 2016). In general, meta-d′ represents a measure of an individual's ability to discriminate between their own correct and incorrect responses independently of differences in response bias (Fleming and Lau 2014) and prior studies have reported meta-d′ either as the only measure of metacognitive sensitivity (Rausch *et al.* 2015) or alongside the meta-d′/d′ ratio (Beck *et al.* 2019).

While it is necessary to keep in mind that the ketamine-associated deterioration of Type 2 sensitivity might be influenced by non-significant group-heterogeneity in Type 1 performance, rather than reflecting a general deficit in the underlying metacognitive processes (Maniscalco and Lau 2012), it is still important to understand ketamine effects on meta-d′ in Study Phase I. This is based on the absence of group effects on Type 1 performance in our study but also on the fact that 95% HDIs only narrowly overlapped with zero in two-sided testing for group differences in metacognitive efficiency. The group-level estimation in a hierarchical Bayesian framework offers several methodological advantages over previous estimation methods for metacognitive efficiency (Fleming 2017). As illustrated in Figure 3, there was an almost perfect fit of the ideal observer model in the placebo group (group-level meta-d′/d′: 0.99), whereas the ketamine group (0.76) substantially deviated from the ideal observer model implied in the meta-d′-framework (Fleming 2017).

Furthermore, we observed a pronounced up-regulation of activity in posterior brain regions with ketamine. This effect was observed only during second-order ratings (including both metacognitive reports and the control condition), whereas Type 1 BOLD showed no difference in activation between the groups. Specifically, there was increased activity in the right-hemispheric superior-posterior cortex compared to placebo. The superior parietal lobe is mainly associated with spatial attention and plays a pivotal role in somatosensory and visuomotor integration (Culham and Valyear 2006; Iacoboni 2006), motor learning (Weiss *et al.* 2003; Wenderoth *et al.* 2004), mental rotation (Wolbers *et al.* 2003; Gogos *et al.* 2010), with a mosaic of specialized subregions (Wang *et al.* 2015). Increased BOLD with ketamine also occurred in left calcarine gyrus, where the primary visual cortex is concentrated (Goebel *et al.* 1998; Seghier *et al.* 2000); bilateral lingual gyrus, which has been linked to processing vision (especially letter-reading) and encoding visual memories (Mechelli *et al.* 2000); and left IPL, which is involved in language processing, mathematical operations and body image (Radua *et al.* 2010), agency (Chaminade and Decety 2002), and working memory (Ravizza *et al.* 2004). Importantly, these ketamine effects on BOLD were observed for both second-order rating types (Report/Follow) and are therefore not specific to genuine metacognitive processes. It should be noted, however, that Report trials were overall more frequent (2:3) than Follow trials and thus had a greater overall contribution to the ketamine effects on second-order BOLD.

Overall, it appears that ketamine affects brain function during second-order ratings by means of an up-regulation of posterior visuospatial cortical brain areas. The visual, affective word stimuli employed in this study may have evoked vivid, imaginative processes in all participants, irrespective of drug, during retrieval. In participants experiencing the altered state of consciousness induced by ketamine, these imaginative processes may yet have persisted well beyond the retrieval process and consequently perturbed the signal available for the second-order task, irrespective of its specific demands, which could account for both the deterioration in metacognitive sensitivity as well as the increased activation in visuospatial areas during second-order ratings. However, it should be reiterated that it is uncertain to what extent the observed effects are related to metacognition, or whether they do not simply reflect neural responses to the presentation of the rating scale.

It is intriguing, however, that the anatomical location of our results is of interest with regards to the "hot zone" for conscious functions proposed by Koch *et al.* (2016): As this hot zone primarily encompasses sensory areas, it is mainly associated with phenomenal qualities of conscious experiences, which self-reported 5D-ASC measures confirmed to be altered by ketamine. Thus, as individuals under the influence of ketamine processed the demands of the second-order task (including introspective assessments of their internal mental world), phenomenal qualities of their normal waking-state experience may be distinctly altered. The posterior parietal cortical areas found in this study have been proposed to encode decision confidence (Kiani and Shadlen 2009), but recent studies suggest that activity in these areas tracks reliability of the sensory input rather than the core process of confidence formation (Bang and Fleming 2018).

Accordingly, our findings suggest that not confidence formation itself, but early aspects of the metacognitive process could be impacted by ketamine as individuals struggle to make sense of a distorted input signal which results in an up-regulation of neural activity, whereas episodic memory or processing speed remain largely unaffected.

This interpretation is supported by evidence that ketamine increases bilateral temporoparietal functional connectivity (Höflich *et al.* 2015) and causes a significant alpha current reduction in posterior cortical areas such as precuneus and temporoparietal junction, which may reflect efforts to maintain ego integrity (Carhart-Harris *et al.* 2014; Vlisides *et al.* 2018). The ketamine-induced psychedelic state is characterized by elevated entropy in certain aspects of brain function, thereby collapsing the highly organized, low-entropy activity within the DMN (Carhart-Harris *et al.* 2014). This is in line with the notion by Carhart-Harris *et al.* (2014) that DMN integrity is a key foundation for accurate metacognition: upon perturbing DMN activity by inducing a psychedelic state, the functionality of metacognitive processes should hence be reduced, whereas the retrieval process may in many cases be based on a notion of familiarity with the word item, and therefore depend less on DMN integrity.

To achieve a comprehensive understanding of the findings, there are additional aspects to be considered. First, the lack of correlation between the 5D-ASC index of altered consciousness and ketamine effects on metacognitive sensitivity makes it difficult to draw a direct connection between the ketamine-induced altered subjective state and the observed objective effects on metacognition—although it may not be adequate to assume both effects to take place on the same conscious level, since the impairment of metacognition represents unconscious effects on conscious decisions (such as ratings given on the 5D-ASC). Second, it has to be considered that different causes might result in a deterioration of metacognitive sensitivity. Both a reduction in the sensory reliability of the input to the metacognitive process (i.e. increased noise in the evidence on which confidence formation is based) as well as trial-to-trial variability in the placement of confidence criteria might account for this effect. A clear interpretation remains difficult, but exploratory analysis of metacognitive bias, which revealed significantly higher bias (i.e. overconfidence) for the ketamine group, offers potential insights into the underlying mechanisms: fluctuations across individual trials in participants' confidence indicate that participants under the influence of ketamine based their confidence ratings on certain conscious experiences, which could be due to changes in conscious access as well as altered, hallucinatory-like experiences, and which are ultimately unknown to the experimenter (Fleming and Lau 2014). Ultimately, it is possible that the unspecific up-regulation of the posterior parietal areas during second-order ratings reflect either the disturbances in signal input or alterations in conscious experience, or even both.

Study Phase II

Drug effects

There were no ketamine effects on Type 1 sensitivity or Type 2 sensitivity and efficiency of items encoded during maintained drug infusion. This was confirmed by exploratory hierarchical Bayesian estimation of group-level metacognitive efficiency; unexpectedly, there was no group difference in metacognitive performance for Study Phase II. The absence of ketamine effects on retrieval is in accordance with previous studies (Honey *et al.* 2005a,b) using a very similar LoP manipulation. We found no drug-related group differences in functional activity during encoding in the continued presence of drug infusion and were thus unable to reproduce the increased activation for deeply encoded items in left PFC with ketamine reported by Honey *et al.* (2005a). Moreover, there were no effects of ketamine on either Type 1 or Type 2 reaction times, again indicating that ketamine did not affect reaction speed. However, metacognitive bias (overconfidence) was again significantly higher in the ketamine group, as was the case in Study Phase I. Even when ketamine was absent at retrieval, ketamine participants were overconfident about their mnestic judgments, suggesting that ketamine evokes substantial distortions in the placement of confidence criteria, irrespective of whether encoding or retrieval took place under the influence of ketamine. While it not possible to retroactively rule out a baseline difference in confidence level between the groups, an overall diffuse memory trace might account for the observed overconfidence, as ketamine affects source memory (Honey *et al.* 2005b). Therefore, ketamine effects on metacognitive bias could be driven by shared and distinct mechanisms for the two study phases.

Limitations

The employment of a between-subjects-design might be a potential shortcoming, as homogeneity in all relevant individual factors can never be achieved across the groups. However, the advantage of this design is that expectancy biases based on experience with the first of two assessments in a within-subjects-design are eliminated.

Whilst the infusion protocol served to keep plasma levels of ketamine constant, it cannot be ruled out that participants became accustomed to the ketamine-induced state of consciousness and developed mechanisms to stabilize higher-order cognitive functions over the course of the infusion. This potential habituation effect may account for the observation that encoding processes in Study Phase II were less affected by ketamine than previously observed (Honey *et al.* 2005a,b).

As participants were not informed about the subsequent retrieval task at encoding in either study phase, it is important to point out that during the encoding task in Study Phase II, participants might have been more likely to infer the subsequent memory testing, which could have altered their encoding strategy. This introduces an additional difference between the two study phases, which complicates a direct comparison of ketamine effects between the phases.

Another limitation is that only trials with correctly retrieved items could be included in the fMRI analyses, due to the fact that the majority of participants produced an insufficient amount of incorrect answers in the Type 1 task. Finally, even though the combined sample size of both groups corresponded to sample sizes of previous within-subject designs (Steffens *et al.* 2016, 2018; Van Loon *et al.* 2016), it is possible that the study lacked sufficient power to detect a statistically significant difference between groups not only on metacognitive sensitivity but also on efficiency.

Generally, additional research is required to gain further understanding of ketamine effects on metacognition. Such potential future research efforts could encompass the application of advanced modeling capable of contrasting theories, such as the Stochastic Detection and Retrieval Model (Jang *et al.* 2012), which could help disentangle the underlying mechanisms of the observed effects and allow to discriminate between increased noise in the sensory evidence accumulation and trial-

by-trial variability in the placement of confidence criteria. Furthermore, dynamic causal modeling of fMRI results could also help to clarify the extent to which the vivid, imaginative processes affect brain activity during second-order ratings.

Conclusions

In summary, we present evidence for a role of the NMDA-glutamate-receptor antagonist ketamine in metacognition, including significantly larger metacognitive bias and deterioration of metacognitive sensitivity with ketamine. We also observed unspecific up-regulation of activity in posterior brain areas during second-order ratings compared to placebo. Importantly, ketamine did not affect metacognitive efficiency as estimated in a hierarchical Bayesian framework. The reported effects are neither sufficiently strong nor specific enough to attribute metacognition solely to the function of the glutamatergic system. Our results do, however, suggest that ketamine impacts on metacognition, which could be due to a reduction in the sensory reliability of the input to the metacognitive process as well as alterations in conscious experience. Further research is required in order to expand our understanding of the neural and pharmacological underpinnings of metacognition.

Supplementary data

Supplementary data is available at *NCONSC Journal* online.

Acknowledgments

We would like to thank Dr. Elisabeth Behrens-Callsen and Laura Augsburg for their excellent support on assessment days. We are grateful to all the participants who volunteered to take part in this study.

Conflict of interest statement. None declared.

Data Availability Statement

Behavioral data available in supplementary material; fMRI data will be made available upon request.

References

Ackenheil M, Stotz G, Dietz-Bauer R, *et al*. *Mini International Interview–German Version 5.0. 0*. München, Germany: Psychiatrische Universitätsklinik München,1999.

Andrews-Hanna JR. The brain's default network and its adaptive role in internal mentation. Neuroscientist. 2012;**18**(3):251–70. doi: 10.1177/1073858411403316.

Anis NA, Berry SC, Burton NR, *et al*. The dissociative anaesthetics, ketamine and phencyclidine, selectively reduce excitation of central mammalian neurones by N-methyl-aspartate. *Br. J. Pharmacol.* 1983;**79**(2):565–75. doi: 10.1111/j.1476-5381.1983.tb11031.x.

Ashburner J, Friston KJ. Unified segmentation. *NeuroImage* 2005;**26**:839–51. doi: 10.1016/j.neuroimage.2005.02.018.

Bang D, Fleming SM. Distinct encoding of decision confidence in human medial prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 2018;**115**:6082–7. doi: 10.1073/pnas.1800795115

Barrett AB, Dienes Z, Seth AK. Measures of metacognition on signal-detection theoretic models. *Psychol. Methods* 2013;**18**:535. doi: 10.1037/a0033268

Beck B, Peña-Vivas V, Fleming S, *et al*. Metacognition across sensory modalities: vision, warmth, and nociceptive pain. *Cognition* 2019;**186**:32–41. doi: 10.1016/j.cognition.2019.01.018

Block N. On a confusion about a function of consciousness. *Behav. Brain Sci.* 1995;**18**:227–47. doi: 10.1017/S0140525X00038188.

Carhart-Harris RL, Leech R, Hellyer PJ, *et al*. The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front. Hum. Neurosci.* 2014;**8**:1–22. doi: 10.3389/fnhum.2014.00020

Chaminade T, Decety J. Leader or follower? Involvement of the inferior parietal lobule in agency. *NeuroReport* 2002;**13**:1975–8. doi: 10.1097/00001756-200210280-00029

Clos M, Bunzeck N, Sommer T. Dopamine is a double-edged sword: dopaminergic modulation enhances memory retrieval performance but impairs metacognition. *Neuropsychopharmacology* 2019;**44**:555–63. doi: 10.1038/s41386-018-0246-y

Cohen J. Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educ. Psychol. Measure.* 1973;**33**:107–12. doi: 10.1177/001316447303300111

Cohen J, Maydeu-Olivares A. A power primer. *Psychol. Bull.* 1992;**112**(1):155–9. doi: 10.1038/141613a0

Culham JC, Valyear KF. Human parietal cortex in action. *Curr. Opin. Neurobiol.* 2006;**16**:205–12. doi: 10.1016/j.conb.2006.03.005

Dittrich A. The standardized psychometric assessment of altered states of consciousness (ASCs) in humans. *Pharmacopsychiatry* 1998;**31**:80–4. doi: 10.1055/s-2007-979351

Dittrich A, Lamparter D, Maurer M. *5D-ABZ: Fragebogen zur Erfassung Aussergewöhnlicher Bewusstseinszustände. Eine kurze Einführung.[5D-ASC: Questionnaire for the Assessment of Altered States of Consciousness. A Short Introduction]*. Zürich, Switzerland: PSIN PLUS,2006.

Domino EF, Zsigmond EK, Domino LE, *et al*. Plasma levels of ketamine and two of its metabolites in surgical patients using a gas chromatographic mass fragmentographic assay. *Anesth. Analg.* 1982;**61**(2):87–92 .doi: 10.1213/00000539-198202000-00004

Doyle OM, De Simoni S, Schwarz AJ, *et al*. Quantifying the attenuation of the ketamine pharmacological magnetic resonance imaging response in humans: a validation using antipsychotic and glutamatergic agents. *J. Pharmacol. Exp. Ther.* 2013;**345**(1):151–60. doi: 10.1124/jpet.112.201665

Eickhoff SB, Stephan KE, Mohlberg H, *et al*. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 2005;**25**:1325–35. doi: 10.1016/j.neuroimage.2004.12.034.

Evans AC, Marrett S, Neelin P, *et al*. Anatomical mapping of functional activation in stereotactic coordinate space. *Neuroimage* 1992;**1**:43–53. doi: 10.1016/1053-8119(92)90006-9.

Fleming SM. HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci. Conscious.* 2017;**2017**:1–14. doi: 10.1093/nc/nix007

Fleming SM, Daw ND. Self-evaluation of decision-making: a general Bayesian framework for metacognitive computation. *Psychol. Rev.* 2017;**124**:91–114. doi: 10.1037/rev0000045

Fleming SM, Huijgen J, Dolan RJ. Prefrontal contributions to metacognition in perceptual decision making. *J. Neurosci.* 2012;**32**:6117–25. doi: 10.1523/JNEUROSCI.6489-11.2012

Fleming SM, Lau HC. How to measure metacognition. *Front. Hum. Neurosci.* 2014;**8**:1–9. doi: 10.3389/fnhum.2014.00443

Galvin SJ, Podd JV, Drga V, *et al*. Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions. *Psychon. Bull. Rev.* 2003;**10**:843–76. doi: 10.3758/BF03196546

Giorgetti R, Marcotulli D, Tagliabracci A, *et al*. Effects of ketamine on psychomotor, sensory and cognitive functions relevant for

driving ability. *Forensic Sci. Int.* 2015;**252**:127–42. doi: 10.1016/j.forsciint.2015.04.024

Goebel R, Khorram-Sefat D, Muckli L, *et al.* The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur. J. Neurosci.* 1998;**10**:1563–73. doi: 10.1046/j.1460-9568.1998.00181.x

Gogos A, Gavrilescu M, Davison S, *et al.* Greater superior than inferior parietal lobule activation with increasing rotation angle during mental rotation: an fMRI study. *Neuropsychologia* 2010;**48**:529–35. doi: 10.1016/j.neuropsychologia.2009.10.013.

Green DG, Swets JA. *Signal Detection Theory and Psychophysics*. New York, NY: Wiley & Sons, Inc., 1966: doi: 10.1901/jeab.1969.12-475.

Grimaldi P, Lau H, Basso MA. 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. *Neurosci. Biobehav. Rev.* 2015;**55**:88–97. doi: 10.1016/j.neubiorev.2015.04.006.

Hauser TU, Allen M, Purg N, *et al.* Noradrenaline blockade specifically enhances metacognitive performance. *ELife* 2017;**6**:1–13. doi: 10.7554/eLife.24901

Hetem LAB, Danion JM, Diemunsch P, *et al.* Effect of a subanesthetic dose of ketamine on memory and conscious awareness in healthy volunteers. *Psychopharmacology* 2000;**152**:283–8. doi: 10.1007/s002130000511

Höflich A, Hahn A, Küblböck M, *et al.* Ketamine-induced modulation of the thalamo-cortical network in healthy volunteers as a model for schizophrenia. *Int. J. Neuropsychopharmacol.* 2015;**18**:1–11. doi: 10.1093/ijnp/pyv040

Holmes CJ, Hoge R, Collins L, *et al.* Enhancement of MR images using registration for signal averaging. *J. Comput. Assist. Tomogr.* 1998;**22**(2):324–33. doi: 10.1097/00004728-199803000-00032

Honey GD, Honey RAE, O'Loughlin C, *et al.* Ketamine disrupts frontal and hippocampal contribution to encoding and retrieval of episodic memory: an fMRI study. *Cerebral Cortex* 2005a;**15**:749–59. doi: 10.1093/cercor/bhh176

Honey GD, Honey RAE, Sharar SR, *et al.* Impairment of specific episodic memory processes by sub-psychotic doses of ketamine: the effects of levels of processing at encoding and of the subsequent retrieval task. *Psychopharmacology* 2005b;**181**:445–57. doi: 10.1007/s00213-005-0001-z

Iacoboni M. Visuo-motor integration and control in the human posterior parietal cortex: evidence from TMS and fMRI. *Neuropsychologia* 2006;**44**:2691–9. doi: 10.1016/j.neuropsychologia.2006.04.029

Jang Y, Wallsten TS, Huber DE. A stochastic detection and retrieval model for the study of metacognition. *Psychol. Rev.* 2012;**119**:186–200. doi: 10.1037/a0025960

Javitt DC, Zukin SR, Heresco-Levy U, *et al.* Has an angel shown the way? Etiological and therapeutic implications of the PCP/NMDA model of schizophrenia. *Schizophr. Bull.* 2012;**38**:958–66. doi: 10.1093/schbul/sbs069

Kiani R, Shadlen MN. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 2009;**324**:759–64. doi: 10.1126/science.1169405

Koch C, Massimini M, Boly M, *et al.* Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* 2016;**17**:307–21. doi: 10.1038/nrn.2016.22

Kruschke J. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. New York, NY: Academic Press, 2014.

Krystal JH, Karper LP, Seibyl JP, *et al.* Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch. Gen. Psychiatry* 1994;**51**:199–214. doi: 10.1001/archpsyc.1994.03950030035004.

Kurdi M, Theerth K, Deva R. Ketamine: current applications in anesthesia, pain, and critical care. *Anesth. Essays Res.* 2014;**8**:283–90.doi: 10.4103/0259-1162.143110

Lou HC, Skewes JC, Thomsen KR, *et al.* Dopaminergic stimulation enhances confidence and accuracy in seeing rapidly presented words. *J. Vis.* 2011;**11**:1–6. doi: 10.1167/11.2.1

Malhotra AK, Pinals DA, Weingartner H, *et al.* NMDA receptor function and human cognition: The effects of ketamine in healthy volunteers. *Neuropsychopharmacology* 1996;**14**:301–7. doi: 10.1016/0893-133X(95)00137-3

Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious. Cogn.* 2012;**21**:422–30. doi: 10.1016/j.concog.2011.09.021

Maniscalco B, Lau H. Signal detection theory analysis of type 1 and type 2 data: meta-d′, response-specific meta-d′, and the unequal variance SDT model. In Fleming, S.M., Frith, C.D. (Eds.): *The Cognitive Neuroscience of Metacognition*. Basel, Switzerland: Springer Verlag Publishing, 2014, 25–66.

McCurdy LY, Maniscalco B, Metcalfe J, *et al.* Anatomical coupling between distinct metacognitive systems for memory and visual perception. *J. Neurosci.* 2013;**33**:1897–906. doi: 10.1523/JNEUROSCI.1890-12.2013

McGlashan TH, Miller TJ, Woods SW, *et al.* Instrument for the assessment of prodromal symptoms and states. In Miller, T.J., Mednick, S.A., McGlashan, T.H., Liberger, J., Johannessen, J.O. (Eds.): Early Intervention in Psychotic Disorders. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2001, 135–49. doi: 10.1007/978-94-010-0892-1_7.

Mechelli A, Humphreys GW, Mayall K, *et al.* Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. *Proc. R. Soc. B* 2000;**267**:1909–13. doi: 10.1098/rspb.2000.1229

Morgan CJA, Mofeez A, Brandner B, *et al.* Acute effects of ketamine on memory systems and psychotic symptoms in healthy volunteers. *Neuropsychopharmacology* 2004;**29**:208–18. doi: 10.1038/sj.npp.1300342

Moritz S, Ramdani N, Klass H, *et al.* Overconfidence in incorrect perceptual judgments in patients with schizophrenia. *Schizophr. Res. Cogn.* 2014;**1**:165–70. doi: 10.1016/j.scog.2014.09.003

Morris H, Wallach J. From PCP to MXE: a comprehensive review of the non-medical use of dissociative drugs. *Drug Test. Anal.* 2014;**6**:614–32. doi: 10.1002/dta.1620

Murphy PR, Robertson IH, Harty S, & O'Connell, R.G. *et al.* Neural evidence accumulation persists after choice to inform metacognitive judgments. *ELife* 2015;**4**:e11946. doi: 10.7554/eLife.11946.

Murrough JW, Iosifescu DV, Chang LC, *et al.* Antidepressant efficacy of ketamine in treatment-resistant major depression: a two-site randomized controlled trial. *Am. J. Psychiatry* 2013;**170**:1134–42. doi: 10.1176/appi.ajp.2013.13030392.

Overgaard M, Rote J, Mouridsen K, *et al.* Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Conscious. Cogn.* 2006;**15**:700–8. doi: 10.1016/j.concog.2006.04.002

Oye I, Paulsen O, Maurset A. Effects of ketamine on sensory perception: evidence for a role of N- methyl-D-aspartate receptors. *J. Pharmacol. Exp. Ther.* 1992;**260**:1209–13.

Radua J, Phillips ML, Russell T, *et al.* Neural response to specific components of fearful faces in healthy and schizophrenic adults. *NeuroImage* 2010;**49**:939–46. doi: 10.1016/j.neuroimage.2009.08.030

Rausch M, Müller HJ, Zehetleitner M. Metacognitive sensitivity of subjective reports of decisional confidence and visual experience. *Conscious. Cogn.* 2015;**35**:192–205. doi: 10.1016/j.concog.2015.02.011.

Rausch M, Zehetleitner M. Visibility is not equivalent to confidence in a low contrast orientation discrimination task. *Front. Psychol*. 2016;**7**:591. doi: 10.3389/fpsyg.2016.00591.

Ravizza SM, Delgado MR, Chein JM, *et al*. (2004). Functional dissociations within the inferior parietal cortex in verbal working memory. *NeuroImage* 2004;**22**(2):562–73. doi: 10.1016/j.neuroimage.2004.01.039

Rouault M, McWilliams A, Allen MG, *et al*. Human metacognition across domains: insights from individual differences and neuroimaging. *Personal Neurosci*. 2018;**1**:e17: doi: 10.1017/pen.2018.16

Sadeghi S, Ekhtiari H, Bahrami B, *et al*. Metacognitive deficiency in a perceptual but not a memory task in methadone maintenance patients. *Sci. Rep*. 2017;**7**:1–8. doi: 10.1038/s41598-017-06707-w

Sarasso S, Boly M, Napolitani M, *et al*. Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine. *Curr. Biol*. 2015;**25**:3099–105. doi: 10.1016/j.cub.2015.10.014

Schifano F, Corkery J, Oyefeso A, *et al*. Trapped in the "K-hole": overview of deaths associated with ketamine misuse in the UK (1993-2006). *J. Clin. Psychopharmacol*. 2008;**28**:114–6. doi: 10.1097/JCP.0b013e3181612cdc.

Schwiedrzik CM, Singer W, Melloni L. Subjective and objective learning effects dissociate in space and in time. *Proc. Natl. Acad. Sci. USA* 2011;**108**:4506–11. doi: 10.1073/pnas.1009147108.

Seghier M, Dojat M, Delon-Martin C, *et al*. Moving illusory contours activate primary visual cortex: an fMRI study. *Cerebral Cortex* 2000;**10**(7):663–70. doi: 10.1093/cercor/10.7.663

Shea N, Frith CD. The global workspace needs metacognition. *Trends Cogn. Sci*. 2019;**23**:560–71. doi: 10.1016/j.tics.2019.04.007

Steffens M, Becker B, Neumann C, *et al*. Effects of ketamine on brain function during smooth pursuit eye movements. *Hum. Brain Map*. 2016;**37**:4047–60. doi: 10.1002/hbm.23294

Steffens M, Neumann C, Kasparbauer AM, *et al*. Effects of ketamine on brain function during response inhibition. *Psychopharmacology* 2018;**235**:3559–71. doi: 10.1007/s00213-018-5081-7

Umbricht D, Koller R, Vollenweider FX, *et al*. Mismatch negativity predicts psychotic experiences induced by NMDA receptor antagonist in healthy volunteers. *Biol. Psychiatry* 2002;**51**:400–6. doi: 10.1016/S0006-3223(01)01242-2

Vaccaro AG, Fleming SM. Thinking about thinking: a coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain Neurosci. Adv*. 2018;**2**:239821281881059.doi: 10.1177/2398212818810591

Van Loon AM, Fahrenfort JJ, Van Der Velde B, *et al*. NMDA receptor antagonist ketamine distorts object recognition by reducing feedback to early visual cortex. *Cerebral Cortex* 2016;**26**:1986–96. doi: 10.1093/cercor/bhv018

Vlisides PE, Bel-Bahar T, Nelson A, *et al*. Subanaesthetic ketamine and altered states of consciousness in humans. *Br. J. Anaesth*. 2018;**121**:249–59. doi: 10.1016/j.bja.2018.03.011

Võ MLH, Conrad M, Kuchinke L, *et al*. The Berlin Affective Word List Reloaded (BAWL-R). *Behav. Res. Methods* 2009;**41**:534–8. doi: 10.3758/BRM.41.2.534.

Wang J, Yang Y, Fan L, *et al*. Convergent functional architecture of the superior parietal lobule unraveled with multimodal neuroimaging approaches. *Hum. Brain Map*. 2015;**36**:238–57. doi: 10.1002/hbm.22626

Weiss PH, Marshall JC, Zilles K, *et al*. Are action and perception in near and far space additive or interactive factors? *NeuroImage* 2003;**18**:837–46. doi: 10.1016/S1053-8119(03)00018-1

Wenderoth N, Debaere F, Sunaert S, *et al*. Parieto-premotor areas mediate directional interference during bimanual movements. *Cerebral Cortex* 2004;**14**:1153–63. doi: 10.1093/cercor/bhh075

Wolbers T, Weiller C, Büchel C. Contralateral coding of imagined body parts in the superior parietal lobe. *Cerebral Cortex* 2003;**13**(4):392–9. doi: 10.1093/cercor/13.4.392

Yokoyama O, Miura N, Watanabe J, *et al*. Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neurosci. Res*. 2010;**68**:199–206. doi: 10.1016/j.neures.2010.07.2041

Study 3	**Lehmann, M.**, Neumann, C., Wasserthal, S., Delis, A., Schultz, J., Hurlemann, R., & Ettinger, U. (2022). Ketamine increases fronto-posterior functional connectivity during meta-perceptual confidence ratings. *Behavioural Brain Research, 430*, 113925. doi: 10.1016/j.bbr.2022.113925

Behavioural Brain Research 430 (2022) 113925

Contents lists available at ScienceDirect

Behavioural Brain Research

journal homepage: www.elsevier.com/locate/bbr

Ketamine increases fronto-posterior functional connectivity during meta-perceptual confidence ratings☆

Mirko Lehmann [a,1], Claudia Neumann [b], Sven Wasserthal [c], Achilles Delis [b], Johannes Schultz [d,e], René Hurlemann [f,g], Ulrich Ettinger [a,*]

[a] Department of Psychology, University of Bonn, Kaiser-Karl-Ring 9, 53111 Bonn, Germany
[b] Department of Anesthesiology, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany
[c] Division of Medical Psychology, Department of Psychiatry and Psychotherapy, University Hospital of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany
[d] Center for Economics and Neuroscience, University of Bonn, Am Hofgarten 8, 53113 Bonn, Germany
[e] Institute for Experimental Epileptology and Cognition Research, Medical Faculty, University of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany
[f] Department of Psychiatry, School of Medicine & Health Sciences, University of Oldenburg, Hermann-Ehlers-Str. 7, 26129 Oldenburg, Germany
[g] Research Center Neurosensory Science, University of Oldenburg, 26129 Oldenburg, Germany

ARTICLE INFO

Keywords:
Metacognition
Ketamine
Perceptual decision-making
Confidence
Glutamate
Functional connectivity

ABSTRACT

Recent advances in the neuropsychopharmacology of metacognition indicate a constituent role of glutamate for the integrity of metamnestic processes. However, the extent to which previous results can be generalized across functional domains to characterize the relationship between glutamate and metacognition remains unclear. Here, in a randomized, double-blind, placebo-controlled, preregistered fMRI study, we tested the effects of a psychotomimetic dose (target plasma concentration 100 ng/mL) of the *N*-methyl-D-aspartate (NMDA) glutamate receptor antagonist ketamine on metacognition in a perceptual decision-making framework. We collected trial-by-trial metacognitive reports as participants performed a two-alternative forced-choice perceptual task during functional magnetic resonance imaging (fMRI). Results indicated ketamine-induced deterioration in metacognitive performance, whereas no significant effects were observed for perceptual performance, response times and – unexpectedly – metacognitive bias. Whilst there were no detectable ketamine effects on mean BOLD activation, exploratory psychophysiological interaction (PPI) analysis revealed alterations in functional connectivity during metacognitive confidence ratings under ketamine. Specifically, there was increased task-specific connectivity for ketamine compared to placebo between right anterior dorsolateral prefrontal cortex and left middle temporal, supramarginal and precentral gyrus, as well as between right insula/inferior frontal gyrus and left lingual gyrus, possibly indicating re-representations of object-level features supplied for metacognitive evaluations. Overall, these findings contribute towards the emerging picture of the substructures underlying metacognitive operations at the neurotransmitter level and may shed light on a neural pattern characteristic of pharmacologically challenged metacognition.

1. Introduction

The term *metacognition*, albeit a notoriously heterogeneous concept, is most commonly described as "thinking about thinking" and refers to the human ability to reflect about one's own cognition and the use of these reflections to regulate cognitive processes [1,2]. The concept was first introduced into the psychological literature by Flavell [3,4] and involves two major functions: monitoring and control of cognition [5,6]. Metacognition serves behavioral optimization, as it guides adaptive decisions e.g. in conditions when external feedback is absent or ambiguous [7]. Furthermore, it is useful in that it can provide a representation of the absence of knowledge [8]. The modification of metacognitive processes is a focus in various psychological therapies, e.g. in the treatment of depression, generalized anxiety disorder, or

☆ This study was preregistered; a link to preregistration, analysis scripts and other relevant materials can be found in the Section 2.

* Correspondence to: Department of Psychology, Kaiser-Karl-Ring 9, 53111 Bonn, Germany.
E-mail addresses: mirko_lehmann@uni-bonn.de (M. Lehmann), claudia.neumann@ukbonn.de (C. Neumann), sven.wasserthal@ukbonn.de (S. Wasserthal), achilles.delis@ukbonn.de (A. Delis), johannes.schultz@ukbonn.de (J. Schultz), rene.hurlemann@uol.de (R. Hurlemann), ulrich.ettinger@uni-bonn.de (U. Ettinger).
[1] ORCID: 0000-0003-0685-9214.

https://doi.org/10.1016/j.bbr.2022.113925
Received 22 October 2021; Received in revised form 12 April 2022; Accepted 9 May 2022
Available online 14 May 2022
0166-4328/© 2022 Elsevier B.V. All rights reserved.

schizophrenia [9–11].

The measurement of metacognition typically requires participants to report their subjective confidence following a cognitive or perceptual judgment; this confidence rating is then evaluated in relation to objective task performance. Therefore, a second-order, metacognitive judgment is adjusted to a first-order judgment, based on a trial-by-trial introspection of the underlying process [12]. Traditionally, two aspects are of major interest: metacognitive *bias*, which expresses an individual's tendency to be generally under- or overconfident, and metacognitive *sensitivity*, which expresses the individual's ability to appropriately discriminate between own correct and incorrect judgments by means of confidence ratings.

Moreover, it is relevant to consider some confounding factors affecting the accuracy of conventional approaches to quantify metacognitive ability. Apart from first-order type 1 (task-related) and second-order type 2 (metacognitive) response tendencies, this primarily concerns the influence of type 1 sensitivity on metacognition [13]. Type 1 task performance may therefore be fixed at a predetermined level using a staircase procedure [14–16], which can be implemented before and/or during collection of metacognitive judgments and which accounts for important sources of bias in the estimation of metacognitive performance [17]. However, staircase procedures may introduce another problem, as they were recently shown to lead to inflated estimates of metacognitive ability due to the mixing of low and high contrast stimuli, which – among other potential solutions – makes it advisable to control for stimulus variability [18]. Finally, the meta-d′ framework [19,20] allows to correct for confounding factors: In addition to "absolute" metacognitive sensitivity (meta-d′), an index of performance-corrected "relative" metacognitive sensitivity or *efficiency* (meta-d′/d′) can be obtained. Metacognition studies employing staircase procedures have therefore used both absolute and relative indices of metacognitive sensitivity to obtain a reliable estimate of an individual's metacognitive ability [18].

Despite a strong increase in efforts and insights over recent years, such as identifying the importance of specific subregions of the prefrontal cortex (PFC) for different metacognitive requirements [21], there are still various remaining questions regarding the functional and biological architecture of metacognition. In particular, there is only a handful of pharmacological challenge studies of metacognition, including a demonstration of increased metacognitive performance after noradrenaline blockade [22] as well as observations of selectively impaired metacognitive efficiency after hydrocortisone administration [23] or, following dopaminergic modulation, for 'New' decisions in a memory paradigm [24].

In a previous functional magnetic resonance imaging (fMRI) study of our group [25], the *N*-methyl-D-aspartate (NMDA) receptor antagonist ketamine led to a deterioration of metacognitive sensitivity (meta-d′) and larger overall confidence in an episodic memory paradigm. This effect was accompanied by unspecific activation increases in posterior brain areas linked with the "posterior hot zone" of the neural correlates of consciousness, a postulate proposed by Koch et al. [26] and recently substantiated by use of dynamic causal modeling [27]. Likewise, activations of visual areas were observed prominently both in ketamine-associated and metacognition-related BOLD contrasts in our previous study. This is unsurprising given the importance of the visual system for cortical organization and, ultimately, various aspects of consciousness [28], but could further be argued to represent increased processing of visual input in relation to hallucination-like percepts [25]. Moreover, there is evidence for increased functional connectivity between a core area of confidence formation anchored in right rostrolateral PFC and visual cortex (i.e., lingual gyrus) during metacognitive reports about visual percepts [29].

Given the results of our previous study [25], ketamine represents a promising candidate for pharmacological modulations of metacognition. The primary pharmacological mechanism of ketamine, which for many years has been used clinically for its anesthetic effects [30], appears to be its role as an uncompetitive antagonist at the NMDA receptor, an ionotropic glutamate receptor type [31–33]. Recently, a growing body of studies has focused on the glutamatergic system as a viable target for the treatment of mood disorders, as ketamine evokes rapid and sustained antidepressant effects in patients with treatment-resistant depression [34–36]. Modulations of the glutamatergic system also offer promising insights into mechanisms and treatments of schizophrenia [37,38]. Notably, at subanesthetic levels, ketamine possesses psychotomimetic properties [39], which include dissociative bodily experiences with spatiotemporal distortions [40] and a general "broadening" of the scope of conscious contents, enhancing the vividness of imagination and modulating the flexibility of cognition [41–43].

The observation that subanesthetic doses of ketamine reliably induce the so-called *psychedelic state* [39,44] is of particular interest to consciousness research, as it allows to modulate different aspects of consciousness in fundamentally different ways [45], e.g. by eliciting a state of *ego dissolution* [46]. Within the framework of Integrated Information Theory (IIT) [47,48], regarded as one of the preeminent contemporary theories of consciousness (although Hanson and Walker [49] recently outlined deficiencies in its falsifiability), the psychedelic state is attributed with various alterations in awareness: Potentially increased cognitive flexibility, creativity, and imagination, which, however, comes at functional costs, such as a degradation of the brain's ability to organize, categorize, and differentiate the constituents of conscious experiences, as well as an inflation of possible cause-effect mechanisms [50]. In the psychedelic state, the brain is thus characterized by a higher state of entropy, experientially richer and more flexible, but less informative than normal waking consciousness [42,50]. Focusing specifically on subanesthetic doses of ketamine, this state of elevated entropy is associated with reduced resting-state connectivity between anterior and posterior parts of the brain's default-mode network (DMN) [51].

Furthermore, ketamine was shown to lead to a reduction of brain activity in regions involved in self-monitoring while increasing activity in regions associated with reward processing and emotional blunting [52]. Specifically, pregenual and subgenual aspects of anterior cingulate cortex (ACC) have been subject to closer scrutiny in previous studies; for the former, a ketamine-associated, region-specific increase in BOLD response was argued to implicate a role of ketamine in attenuating an inordinate self-focus during negative experiences [53]. Interestingly, these regions are functionally associated with posterior medial frontal cortex (pMFC), which was identified as a central hub of metacognition-related activations in our previous study as well as a meta-analysis of MRI studies on metacognitive judgments [54]. Whereas much research focused on BOLD activation or resting-state connectivity, it is in some contexts more informative to consider context-dependent connectivity of brain areas in relation to specific task conditions [55]; here, ketamine was shown to increase coupling between medial prefrontal and parahippocampal areas in an emotional memory task [56].

The identification of the neural correlates and neurotransmitter systems underlying metacognition is meanwhile complicated by the fact that distinct metacognitive subsystems may exist for different tasks and requirements. Multiple studies [14,57,58] failed to obtain significant correlations regarding the accuracy of metacognitive judgments across experimental domains, which was substantiated in a meta-analysis [59]. At the neural level, Baird et al. [15] reported evidence for spatial specialization within the anterior PFC for different types of metacognitive processes, namely in relation to perceptual decision-making ("meta-perception") and mnemonic retrieval ("meta-memory"). Consequently, meta-perception and meta-memory may represent distinct processes with distinct neural correlates, and findings obtained about metacognition with respect to one specific domain are thus not necessarily applicable to other functional domains.

However, recent studies demonstrated that domain-general contributions to the structure of metacognition can be revealed under optimized methodological conditions and with sufficient statistical power

[60,61], and beyond functional specializations, there is also some unity in the neural profiles of different metacognitive processes [54]. Although little is known about the domain-generality or domain-specificity of the effects of pharmacological challenges on metacognition, one might reasonably assume a shared reliance of metacognitive processes on specific neuronal mechanisms. Detecting congruent effects of the same pharmacological intervention across tasks would suggest a domain-general neurophysiological substrate at the level of (partially) shared neuronal mechanisms that could subserve the computation of metacognitive processes, which would contribute towards a fundamental account of the biological substructures that constitute the functional architecture of metacognition.

Building upon our previous study [25], which suggested ketamine impacts meta-memory, we conducted a double-blind, placebo-controlled, preregistered fMRI study to investigate the effects of a psychotomimetic dose of ketamine on meta-perception and associated brain activity. As participants' primary task performance was maintained at a constant level by use of a staircase procedure, they provided confidence ratings on their trial-by-trial decisions on a two-alternative forced-choice (2AFC) perceptual magnitude comparison task with static visual stimuli. Induction of a psychedelic state was assessed using a self-report questionnaire. In accordance with preregistration, we hypothesized that we would find evidence for ketamine-induced alterations in metacognitive performance as well as neural activity during metacognitive reports. Thereby, we aimed to contribute to the emergent understanding of metacognition under pharmacological challenges.

2. Materials and methods

2.1. Participants

Seventy young adult volunteers were recruited via mailing lists and online advertisements. They provided written, informed consent and received financial reimbursement for their participation. Volunteering individuals were excluded if they met any of the following criteria: Serious physical illness; history of neurological or psychiatric illness; hyperthyroidism; hypo- or hypertension; under- or overweight; prior experience with ketamine; history of alcohol or drug abuse within the last twelve months or complications during anesthesia; concurrent medication, MRI incompatibility (metalliferous implants, claustrophobia), positive urine drug test, and positive urine pregnancy test. An extensive screening procedure was carried out, as detailed in a previous publication employing a different study sample, but the same equipment and infusion protocol [25]. Participants arrived at the testing facility after a minimum of 2 h fasting clear fluids, 6 h fasting solid food and 24 h fasting alcohol. On the assessment day, an on-site physical examination was performed by medical professionals prior to MRI testing. Consistent with ethical and anesthesiological standards, participants received pre-experimental information about the possibility of ketamine application and potential side effects of the drug. All participants were treated with equal care, and the double-blind protocol was maintained at all times. Two participants failed to complete the full course of the study (dropouts), and twenty-three participants were excluded due to a technical error, which led to a large deviation of their responses from the targeted percentage of correct responses. Data of forty-five healthy, right-handed, non-smoking participants (21 female, 24 male; aged 19–34 years; M=23.96, SD=4.06) with normal or corrected-to-normal vision were included in data analysis.

In accordance with the study's Research Ethics Committee approval (Department of Psychology, University of Bonn; approval number: 19–03–29), behavioral data are provided as supplementary materials and MRI data will be made available upon request; analysis scripts, preregistration and other relevant materials can be accessed via OSF (https://osf.io/gucm2/).

2.2. Experimental design and infusion protocol

This study employed a randomized, double-blind, placebo-controlled between-subjects design. As in previous studies of our group [25,62,63], drug administration was carried out via an intravenous access in the non-dominant arm. Of the included participants, a subset of 19 individuals (8 female) received a placebo infusion (0.9% sodium chloride saline solution, Ratiopharm©, Ulm, Germany), while the other 26 participants (13 female) were administered a subanesthetic dose of racemic ketamine (Ketamin-Ratiopharm 500 injection solution, Ratiopharm©, Ulm, Germany) as a 2 mg/mL solution with a constant target plasma level of 100 ng/mL by a bolus and continuous intravenous infusion through a computer-controlled infusion pump (Graseby 3500, Smith Medical Int. Ltd, Luton, UK). Upon termination of the infusion, participants were asked to report their internal states and subjective experiences during drug administration on the Altered States of Consciousness (5D-ASC) rating scale [64,65], a 94-item inventory assessing five dimensions by which altered states of consciousness can be characterized via ratings on a visual-analogue scale (VAS). These encompass three oblique primary dimensions, "Oceanic Boundlessness", "Dread of Ego Dissolution" and "Visionary Restructuralization", which can be summed to form a global measure of altered consciousness, and two ancillary dimensions, "Vigilance Reduction" and "Auditory Alterations".

2.3. Stimuli

Meta-perception was investigated in a 2AFC magnitude comparison task (MC-T). Presentation of the experiment and recording of behavioral responses were performed using *Presentation*® software (Version 17.2, Neurobehavioral Systems). Stimuli were presented on a 32-inch NordicNeuroLab LCD monitor (1920 ×1080 pixels, 120 Hz refresh rate) and viewed via a head coil-mounted mirror; eye gaze was not monitored. Participants gave predefined button-presses on ResponseGrip hardware (NordicNeuroLab, Bergen, Norway), using fingers of both hands.

The MC-T (Fig. 1) was modified from Fleming et al. [66,67] and implemented in a block design. A 2:1 staircase procedure was applied to maintain individual task performance at a constant level, so that all relevant between-group differences could reliably be attributed to group differences in metacognition [18,68]. The following event sequence was reiterated during the task: For 1000 ms, participants were initially presented with two white circles (diameter: 3.96° of visual angle) on dark background with white central crosshairs (global x-shift from center of the screen: ± 2.64°). Subsequently, the crosshairs were removed and randomly distributed white dots (diameter: 0.11°) were

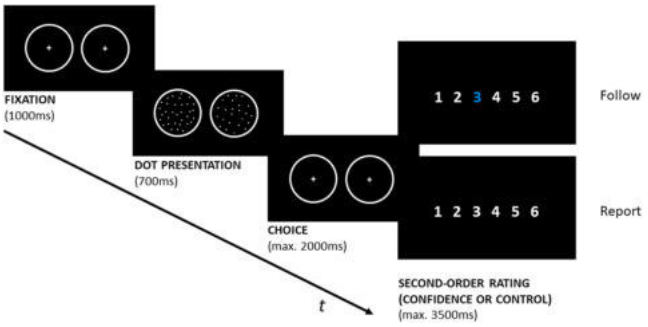

Fig. 1. Schematic trial representation for the 2AFC magnitude comparison task (MC-T). After a fixation period and presentation of dots, participants were required to make binary judgments about which circle (left/right) contained the higher number of dots. On each trial during the experimental phase, they subsequently stated either their confidence in their decision ("Report") or moved the cursor to a color-coded position on the scale ("Follow"). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

presented inside both circles for 700 ms before being replaced by identical crosshairs as before. Within 2000 ms after dot offset, participants were asked to indicate which circle (left/right) had contained the higher number of dots. On each trial during the experimental phase, they were then required to provide a second-order rating.

Corresponding to the experimental design of our previous study [25], there were two second-order (type 2) rating conditions, one requiring the employment of genuine metacognitive processes ("Report", 100 trials) and one serving as a matched control condition with identical motor, but different cognitive demands ("Follow", 50 trials). Each block of 10 "Report" trials was followed by 5 "Follow" trials, which were then succeeded by a new "Report" block. On "Report" trials, participants were required to state their subjective confidence in having rendered a correct perceptual judgment within 3500 ms on a 6-point-Likert-scale with discrete levels of confidence (1 = "no confidence at all", 6 = "very high confidence"). Participants moved a cursor along the scale by index finger button-presses, starting from a randomly determined initial cursor position, until they arrived at the position on the scale most consistent with their subjective feeling of confidence, which they were asked to confirm via thumb button-presses. In the "Follow" condition, participants were required to navigate towards and confirm a randomly determined color-coded number on the scale and withhold from reporting their subjective confidence.

Before initiation of the experimental phase, participants performed 100 training trials without second-order ratings, which were used to calibrate task difficulty for the experimental phase. In line with Fleming et al. [66], one randomly assigned circle on each trial (left/right) contained a variable number of dots ("variable circle") depending on the participant's performance, whereas the other circle always contained exactly 50 dots ("fixed circle"). At the beginning of staircase calibration, a dot number between 35 and 49 or between 51 and 65 was randomly determined for the variable circle, so either circle type could represent the target circle (i.e., the one with the greater number of dots). Equating the difficulty of the MC-T was achieved by titrating the difference in dot number (Δd) between the two circles; the Δd value in the final trial of the training phase was entered as the starting point for the staircase during the experimental phase. After two consecutive correct responses in the training or experimental phase, Δd was decreased by one dot; after one incorrect response, however, Δd was increased by one dot [66]. The variable circle was not allowed to contain exactly 50 dots (Δd=0), so in the case of two consecutive correct responses at Δd= 1, Δd was not decreased.

2.4. Imaging protocol

Imaging data were collected using a 3-Tesla field strength MAGNETOM Tim-Trio MRI system (Siemens, Erlangen, Germany) equipped with a standard 12-channel head coil for signal transmission and reception. Participants' heads were fitted with foam pads to minimize motion-related artifacts. Functional MRI time-series with blood-oxygenation-level-dependent (BOLD) contrast of the whole brain were acquired using a T_2*-weighted gradient-echo planar image (EPI) sequence (repetition time=2500 ms; echo time=30 ms, matrix size=96×96, slice thickness=3.0 mm, field of view=192 mm, flip angle=90°, voxel size=2×2×3 mm, 37 transversal slices). A T_1-weighted gradient-echo sequence with inversion recovery (repetition time=1660 ms, echo time=2.54, inversion time=850 ms, matrix size=320×320, slice thickness=0.8 mm, field of view=256 mm, flip angle=9°, voxel size=0.8 × 0.8 × 0.8 mm, 208 sagittal slices) was used to acquire whole-brain high-resolution anatomical images for normalizing functional imaging data and detecting participants with apparent brain pathologies.

2.5. fMRI analyses

fMRI data were preprocessed and analyzed using *SPM8* and *SPM12* (Statistical Parametric Mapping; Wellcome Centre for Neuroimaging, London, UK) implemented in *Matlab R2014a* and *MatlabR2020b* (The MathWorks Inc., Natick, USA), respectively. Preprocessing was carried out using standardized protocols [69]. First, origins were set manually to anterior commissure to facilitate co-registration [70,71]. Anatomical images were segmented into grey matter, white matter, and cerebral spinal fluid using mutual information and a priori tissue probability maps [72]. After discarding the first five volumes of each functional time-series to ensure steady-state magnetization, functional images were motion-corrected during realignment using a least-squares approach and a six-parameter rigid body transformation; the segmented structural image was co-registered to the mean individual T_2*-weighted image. Furthermore, functional images were spatially normalized into standard stereotaxic Montreal Neurological Institute (MNI) space [73] via non-linear transformations, resampled at 2×2×2 mm resolution, and spatially smoothed using an isotropic full-width-at-half-maximum Gaussian kernel of 8 mm.

On the 1st (participant-wise) level, fMRI time-series were regressed onto the general linear model (GLM) in *SPM12*. Individual trials were modeled as events, as previously implemented by Fleming et al. [29], containing stick functions representing type 1 stimulus onsets and boxcar functions spanning the time from scale onset until confirmation of the second-order (i.e., report/follow) rating; low-frequency fluctuations in BOLD signal were excluded with a 128-s high-pass filter. Consequently, there were five regressors (Perception, CorrectReport, IncorrectReport, CorrectFollow, IncorrectFollow), the latter four "second-order regressors" parametrically modulated by the selected confidence rating in each trial, enabling discrimination between correct and incorrect responses, levels of confidence and perceptual (type 1) and metacognitive (type 2) judgments. Motion-correction parameters were added to the GLM as covariates of no interest; regressors were convolved with a canonical hemodynamic response function (HRF).

As a first exploratory step on the 2nd (group-wise) level, separate random-effects analyses (one-sample *t*-tests) were carried out on 1st level contrast images for perceptual judgments and combined second-order regressors against zero to identify overall patterns of activation, irrespective of Drug or Rating Type (see below). As preregistered, corresponding contrast images of second-order regressors were subsequently entered into a full factorial analysis using the between-subjects factor "Drug" (ketamine/placebo) and within-subject factors "Rating Type" (report/follow) and "Perceptual Performance" (correct/incorrect). Analyses were conducted on the whole-brain level, not masking for any region of interest (ROI). Anatomical labels were inferred by the SPM anatomy toolbox atlas [74]; all reported activations survived $p < .05$, family-wise-error (FWE) corrected at the cluster-level, with a voxel-level threshold of $p < .001$ (uncorrected). Imaging data of one participant were excluded due to missing parts of the PFC in the anatomical image.

As preregistered full factorial analyses of BOLD activation yielded inconclusive results (see below), exploratory psychophysiological interaction (PPI) analyses were applied to assess ketamine effects on task-specific connectivity, i.e. regional changes in the relationship between activity in different areas of the brain as a function of the experimental manipulation [75]. PPI analyses are a powerful tool to explore task-specific functional connectivity, as they do not rely on a priori definitions of possible models [55]. Importantly, PPI measures explain the regional activity of different brain areas in terms of the interaction between a psychological (the task) and a deconvolved physiological factor (e.g., neural responses in a given seed region). Following Fleming et al. [29], we constructed a separate block-level 1st level design matrix for PPI analyses containing boxcar functions spanning the time from onset of one second-order rating block until onset of the succeeding block; consequently, events were defined as an entire block of Report (10) or Follow (5) trials. PPI analyses thus revealed regions exhibiting significant co-activations with the seed regions during Report compared to Follow trials. The time course vector of the

psychophysiological interaction was entered in a fixed-effect GLM along the Report and Follow vectors, the time course of the seed regions and motion-correction parameters as covariates of no interest, yielding a map of co-activations that systematically increased with genuine metacognition [29,76]. The automated generalized PPI toolbox (gPPI) [77] in *SPM8* was used to carry out PPI analysis based on the deconvolved first eigenvariate of the seed region time series [78], which among other advantages has proven to be particularly suited for analyzing fMRI data in block designs [79,80].

Seed regions were determined as a 6 mm sphere centered around the peak coordinate of clusters identified in the meta-analysis by Vaccaro and Fleming [54], which contained five clusters specifically related to metacognition in perceptual decision-making (coordinates are in MNI space): right anterior dorsolateral PFC [x = 26, y= 48, z = 28], right insula [x= 32, y= 20, z = -12], right insula/inferior frontal gyrus (IFG) [44, 14, 0], as well as two global maxima in bilateral pMFC [6, 38, 42; 2, 20, 38]. In 2nd level analyses for each of the seed voxels, drug-related differences in functional connectivity were assessed using random-effects analyses (two-sample *t*-tests) to investigate the differential co-activation maps in a metacognition network during ketamine vs. placebo. Again, we applied a whole-brain cluster-level FWE-correction ($p < .05$) with a peak-level threshold of .001 (uncorrected).

2.6. Behavioral analyses

Both type 1 (task) and type 2 (metacognitive) performance were assessed in a signal detection theory (SDT) framework [81,82]. As reported previously [25], only confidence ratings given on "Report" trials following a completed perceptual judgment contributed to analysis. Since perceptual performance was equated by use of a staircase procedure, there is substantial interpretative value of absolute metacognitive sensitivity (meta-d′) in and by itself [13]. However, as staircase-related stimulus variability could lead to differential effects on ability estimates [18], we considered measures of both absolute metacognitive sensitivity (meta-d′) and type 1 performance-corrected metacognitive sensitivity or *efficiency* (meta-d′/d′), computed in Matlab using the HMeta-d toolbox, which implements a hierarchical Bayesian framework [83]. Among other advantages, this approach yields a more accurate estimation of subject-level parameters by constraining subject-level fits to the group-level estimate and avoiding the need for edge correction, which may otherwise lead to biased subject-level estimates especially with smaller trial counts. In addition, regularization of efficiency estimates by use of the hierarchical approach consistently improves their test-retest reliability [84]. Analysis of the difference between the group posterior densities of independently fitted models for ketamine and placebo groups can be found in the supplementary materials. The HMeta-d toolbox uses Markov-Chain-Monte-Carlo sampling from the posterior distributions [83]; three chains were run for estimation and parameter convergence was assessed by inspection of scale-reduction statistics [85].

Although the variability-based inflation of metacognitive performance estimates may be negligible for studies (a) employing very small step sizes ($\Delta d \pm 1$) and (b) which achieve staircase calibration prior to actual data collection [18], two conditions satisfied in the present study, we monitored stimulus variability in our staircase by testing for group differences in variability (normalized *SD*) and conducting an analysis of covariance (ANCOVA) for drug effects on metacognition measures, controlling for variability (see supplementary materials). In addition, we considered the absolute perceptual threshold, i.e., the stimulus value (Δd) at the end of staircase calibration and its mean value during the experiment, as systematic stimulus differences between groups would suggest perceptual impacts by ketamine despite equating performance. Following a reviewer's comment, this was also investigated in a Bayesian model comparison framework (see supplementary materials) beyond the analyses reported here, using the BayesFactor package [86] in *R* (Version 4.0.1, The R Foundation).

As in our previous publication [25], we extended our preregistered analysis plan to metacognitive bias (quantified as average confidence rating minus average performance) to test for group differences in level of confidence which cannot be explained by group differences in performance, and conducted Pearson's correlations between ability estimates and metacognitive bias with the individual 5D-ASC scores, while correcting for multiple comparisons (Bonferroni-corrected $\alpha = 0.05$, divided by number of correlations). Due to the potential ambiguity of interpretations regarding the lower end of the confidence scale in a 2AFC task [60], we also report group differences in average confidence level, not corrected for performance. Finally, mean beta-values for peak-voxels of significant clusters obtained in the two-sample *t*-tests on PPI contrasts were extracted using the *MarsBar* toolbox in Matlab [87] by transforming clusters into binary mask images. We consequently explored the relationship of behavioral outcomes with regions significantly co-activated with core areas of metacognition during ketamine compared to placebo via separate Pearson's correlations for ketamine and placebo groups.

All analyses of behavioral data were carried out in *SPSS* 22 (IBM Corp., Armonk, USA). As preregistered, data points outside three interquartile ranges of a boxplot were considered to be extreme outliers and not included in data analysis. To ensure that all requirements for statistical analyses were met, data were screened for normality of distribution using histograms, skewness scores and Kolmogorov-Smirnov tests ($\alpha = 0.05$); Levene's statistics were inspected to ensure homoscedasticity. Two-sample *t*-tests were employed to test for drug effects on 5D-ASC scales, stimulus value and variability, metacognitive bias, type 1 (d′) and type 2 (meta-d′; meta-d′/d′) performance as well as perceptual and second-order response times, the latter separately for Report and Follow; Cohen's *d* [88] was calculated for effect sizes. Raincloud plots [89] were created in *R* to visualize data distributions.

3. Results

3.1. Behavioral results

Descriptive statistics of dependent variables per group are in Table 1.

Table 1
Descriptive statistics of 5D-ASC measures, behavioral outcome measures, and response times, per group.

Measure	Ketamine ($n = 26$)		Placebo ($n = 19$)	
	M	*SD*	*M*	*SD*
5D-ASC				
[Global Index of Altered State]*	10.29	9.67	0.68	1.14
Oceanic Boundlessness*	13.89	15.53	0.49	1.22
Dread of Ego Dissolution*	8.53	9.55	0.86	1.36
Visionary Restructuralization*	6.94	7.83	0.74	1.34
Auditory Alterations*	5.85	6.64	1.27	2.24
Vigilance Reduction*	28.67	19.25	5.76	6.41
Stimulus properties				
Initial stimulus value	3.85	3.08	4.42	2.46
Mean stimulus value	4.13	1.04	3.98	1.08
Stimulus variability	2.07	0.46	1.96	0.58
Behavioral outcome measures				
Type 1 sensitivity (d′)	0.87	0.23	0.99	0.25
Type 2 sensitivity (meta-d′)*	0.33	0.39	0.61	0.52
Type 2 efficiency (meta-d′/d′)	0.39	0.45	0.63	0.50
Average confidence level	3.74	0.95	4.02	0.48
Metacognitive bias	-0.10	0.20	-0.05	0.10
Response times (RT, in ms)				
Type 1 RT	671	150	625	168
Report RT	1637	285	1643	281
Follow RT	1494	236	1450	225

Note: Scale values are in percent. M, mean; ms, milliseconds; RT, response time; SD, standard deviation. *significant effect of Drug ($P < .05$).

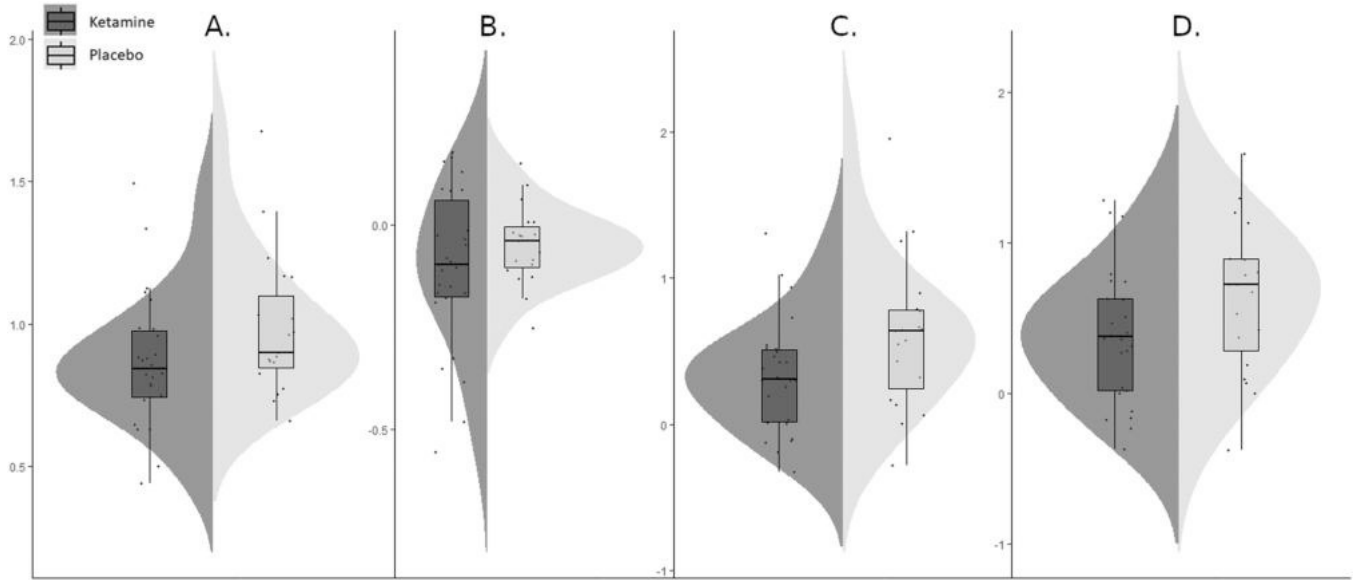

Fig. 2. Raincloud plots for behavioral outcome measures, per group (ketamine, dark grey; placebo, light grey). **A.** Type 1 sensitivity (d′), **B.** performance-corrected metacognitive bias, **C.** Type 2 sensitivity (meta-d′), and **D.** Type 2 efficiency (meta-d′/d′).

Raincloud plots for key outcome measures per group are in Fig. 2; additional raincloud plots for average confidence ratings, response times and stimulus properties can be found in the supplementary materials.

5D-ASC: Highly significant group differences emerged on all scales, with the ketamine group exhibiting higher values on "Oceanic Boundlessness" ($t(25.42) = 4.38$, $p < .001$, $d = 1.22$); "Dread of Ego Dissolution" ($t(26.37) = 4.04$, $p < .001$, $d = 1.12$); "Visionary Restructuralization" ($t(26.94) = 3.91$, $p = .001$, $d = 1.09$); "Auditory Alterations" ($t(32.32) = 3.26$, $p = .003$, $d = 0.92$) and "Vigilance Reduction" ($t(32.14) = 5.66$, $p < .001$, $d = 1.6$) as well as on the composite score of altered consciousness ($t(25.94) = 4.3$, $p < .001$, $d = 1.4$).

However, the 5D-ASC scores did not correlate significantly with type 1 and type 2 outcome measures (all $P > .001$), not even when applying a less conservative correction threshold (all $P > .05$).

Behavioral outcome measures: No significant group difference was observed in stimulus variability ($t(43) = 0.71$, $p = .482$, $d = 0.21$); therefore, a confounding influence of this factor on behavioral outcome variables was not assumed. There were also no significant between-group differences for d′ as the measure of task performance ($t(43) = 1.64$, $p = .109$, $d = 0.49$), or for the initial ($t(43) = 0.70$, $p = .506$, $d = 0.21$) and mean ($t(43) = 0.45$, $p = .65$, $d = 0.14$) stimulus value during experimental blocks. These results were validated by ANCOVA and Bayesian model comparisons (see supplementary materials).

At the type 2 level, meta-d′ or absolute metacognitive sensitivity ($t(43) = 2.04$, $p = .047$, $d = 0.6$) significantly deteriorated under ketamine, whilst the group difference on meta-d′/d′, the measure of relative metacognitive sensitivity/efficiency, was only marginally significant ($t(43) = 1.7$, $p = .096$, $d = 0.51$). There was no significant difference between groups in average confidence level ($t(38.89) = 1.29$, $p = .206$, $d = 0.37$) or performance-corrected metacognitive bias ($t(38.19) = 1.23$, $p = .225$, $d = 0.35$). See supplementary materials for further hierarchical Bayesian analyses on group-level values of metacognitive efficiency.

Response times: Neither type 1 ($t(43) = 0.96$, $p = .343$, $d = 0.29$), nor report ($t(43) = 0.07$, $p = .947$, $d = 0.02$), nor follow response times ($t(43) = 0.62$, $p = .538$, $d = 0.19$) significantly differed between groups.

3.2. BOLD results

One-sample *t*-tests: Random-effects analyses revealed increased activation across groups during perceptual judgments compared to baseline in bilateral pMFC and superior frontal gyrus. One-sample *t*-tests furthermore revealed increased activation during second-order reports (Report and Follow) compared to baseline in left fusiform gyrus; bilateral occipital cortex; superior, inferior and middle temporal gyrus; and motor cortex; decreases in activation were mainly observed in bilateral angular gyrus. For detailed information on clusters and peak-voxels in one-sample *t*-tests, see supplementary tables 1–3.

Full factorial analysis: Our preregistered full factorial analysis with factors "Drug", "Rating Type" and "Perceptual Performance" failed to reveal significant differences in BOLD activation between ketamine and placebo. In contrast to results of our previous study, we were also unable to observe clusters significantly more activated during Report than Follow trials. However, we found increased BOLD signal for the reverse contrast (Follow > Report) in two clusters centered around peak-voxels in bilateral cuneus and precuneus, the latter representing a core structure of the DMN [90]; details are given in Table 2. There were no significant effects in either direction of the factor "Perceptual Performance" (correct/incorrect) and no significant interactions (all $P > .05$).

Functional connectivity analyses: Two-sample *t*-tests revealed significantly higher task-specific connectivity under ketamine compared to placebo (Fig. 3) between the seed voxel in right anterior dorsolateral PFC and two left-hemispheric clusters (Table 3) during Report compared to Follow ratings: one centered in middle temporal gyrus, with additional local maxima in supramarginal, superior temporal and angular gyrus, the other cluster centered in precentral gyrus with an additional local maximum in mid cingulum.

There was also significantly higher functional connectivity under

Table 2
Summary of significant clusters for the Follow > Report contrast.

Anatomical label	Laterality	Cluster size [*k*]	*T*-Value	Peak-voxel MNI coordinates		
Cuneus	L/R	398	3.94	0	-82	36
Cuneus	R		3.80	6	-82	36
Precuneus	R		3.56	16	-78	46
Cuneus	L		3.55	-6	-88	24
Precuneus	L	174	4.43	-4	-46	74
Precuneus	R		4.14	6	-46	74

Note: Combined sample. Only unique anatomical labels are reported for each cluster at one laterality. Whole-brain cluster-level FWE-corrected ($P < .05$). FWE, familywise error; L, left; MNI, Montreal Neurological Institute; R, right.

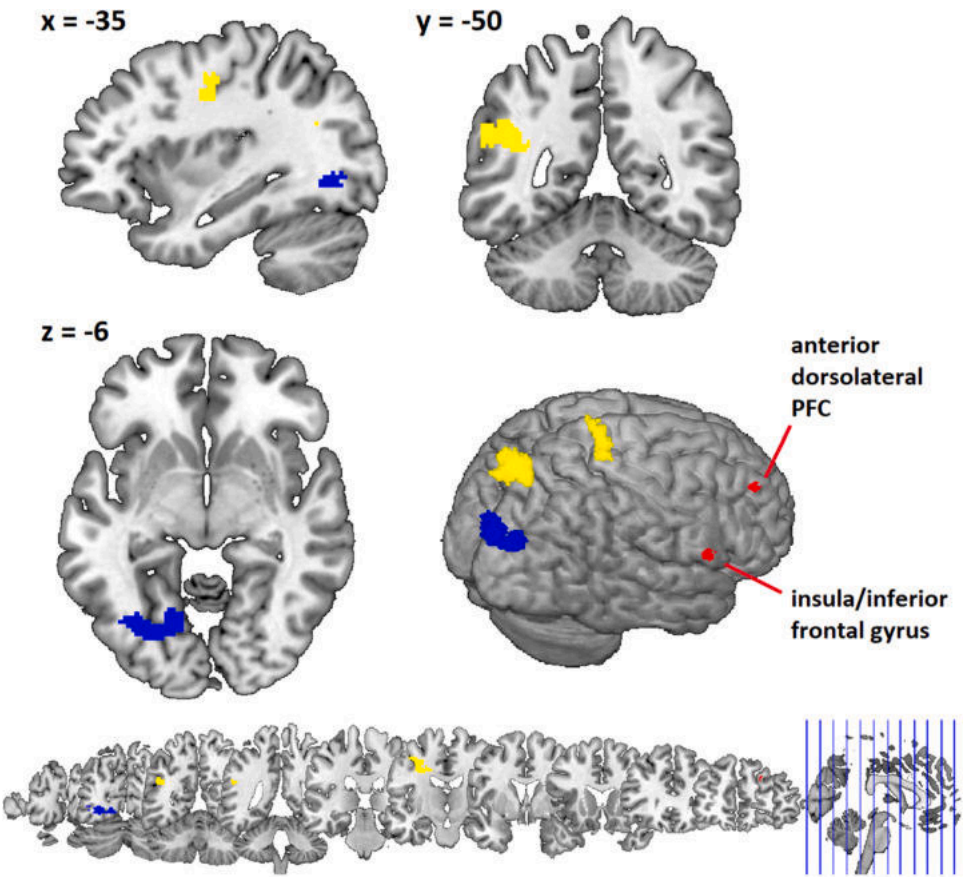

Fig. 3. Increases in functional connectivity during metacognition under ketamine, significant at $P < .05$ (FWE-corrected). Blue and yellow colors indicate areas with significantly higher task-specific connectivity under ketamine compared to placebo with the seed voxel in anterior dorsolateral PFC (yellow) and with the seed voxel in right insula/IFG (blue) during Report trials. Positions of seed voxels are highlighted in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Summary of areas displaying significantly higher task-specific connectivity with the seed voxel in right anterior dorsolateral PFC under ketamine compared to placebo.

Anatomical label	Laterality	Cluster size [*k*]	*T*-Value	Peak-voxel MNI coordinates		
Middle temporal gyrus	L	396	4.82	-46	-48	22
Supramarginal gyrus	L		4.49	-54	-50	24
Superior temporal gyrus	L		4.41	-44	-46	18
Angular gyrus	L		3.99	-42	-64	26
Precentral gyrus	L	150	4.65	-32	-10	38
Mid cingulum	L		3.95	-20	-14	36

Note: Only unique anatomical labels are reported for each cluster at one laterality. Whole-brain cluster-level FWE-corrected ($P < .05$). FWE, familywise error; L, left; MNI, Montreal Neurological Institute.

Table 4
Summary of areas displaying significantly higher task-specific connectivity with the seed voxel in right insula/IFG under ketamine compared to placebo.

Anatomical label	Laterality	Cluster size [*k*]	*T*-Value	Peak-voxel MNI coordinates		
Lingual gyrus	L	299	4.74	-16	-66	-4
Occipital fusiform gyrus	L		4.36	-30	-72	-6
Middle occipital gyrus	L		4.08	-40	-68	-2

Note: Only unique anatomical labels are reported for each cluster at one laterality. Whole-brain cluster-level FWE-corrected ($P < .05$). FWE, familywise error; L, left; MNI, Montreal Neurological Institute.

ketamine compared to placebo between the seed voxel in right insula/IFG and one left-hemispheric cluster in lingual gyrus, with local maxima in occipital fusiform gyrus and middle occipital gyrus (Fig. 3, Table 4).

Correlations with behavioral outcomes: There were no significant correlations between behavioral outcome measures and mean beta-values extracted for peak-voxels of significant clusters in the two-sample *t*-tests on PPI maps (all $P > .001$).

4. Discussion

This study aimed to improve our understanding of the role of the glutamatergic system in metacognition by means of a ketamine challenge, as participants provided trial-by-trial confidence ratings on a perceptual decision-making task during functional brain imaging. Our data indicate that the NMDA-glutamate-receptor antagonist ketamine impacted on the accuracy of meta-perceptual judgments, as it led to significantly lower absolute metacognitive sensitivity. Together with our previous study [25], we conclude that the precision of metacognitive evaluations is attenuated during the ketamine-induced psychedelic state. Since there was no clear evidence of ketamine-induced alterations in other task-related components in either study, neither at the behavioral nor at the neural level, there also appears to be some specificity in the effects of glutamatergic modulation of metacognition, although several factors warrant caution about such conclusions, as outlined below. We were surprised to find that metacognitive bias was unaffected by the drug and that the neural correlates of the ketamine effect on metacognitive performance remained obscure in the preregistered full factorial analysis; however, exploratory PPI analysis revealed ketamine-induced increases in fronto-posterior functional connectivity, thereby providing firm evidence for specific ketamine-induced alterations in metacognitive processes at the neural level. Finally, the induction of substantial alterations of consciousness was confirmed by

significant ketamine effects on all scales of the 5D-ASC questionnaire, although these were somewhat smaller overall than previously observed [25]. As can be inferred from Table 1, effects on sleepiness (vigilance reduction) as well as joyful aspects of subjective experience (oceanic boundlessness) were most prominent, whereas visual or auditory perceptual alterations were reported to a lesser extent.

Whereas our previous study demonstrated stable overconfidence with ketamine, independent of whether or not the drug was currently administered at the time of metacognitive reports, the present analysis was unable to reveal significant group differences in metacognitive bias, with average bias even descriptively lower for ketamine than placebo (Table 1). As discussed previously [25], it cannot be ruled out that baseline differences between groups may have caused the effect reported there; after all, experimental control of such factors is inevitably limited in a between-subjects design. Nevertheless, an interesting notion arises with regard to the noticeably wider spread of bias values in the ketamine group compared to placebo, as evident from Fig. 2B, suggesting that ketamine may in fact be associated with both under- and overconfidence. Such response-heterogeneity might reflect individual differences potentially amplified by ketamine application: Since metacognitive bias has been demonstrated to possess domain-general properties [91,92] that can partially be tapped by self-report measures [61], it may well represent a trait-like quality which could be differentially affected by pharmacological stimulation. Even so, the present finding is not readily placed within the context of the clinical metacognition literature, which e.g. focuses on ketamine as a model system of schizophrenia [93], since overconfidence in incorrect responses has repeatedly been demonstrated in patients with schizophrenia [94,95]. It should be noted, however, that metacognitive phenomena such as jumping-to-conclusions in perceptual decision-making have been argued to depend on the type of schizophrenic symptomatology [96] and that a recent meta-analysis attributed the reported global metacognitive deficit in schizophrenia to methodological shortcomings of multiple studies, such as failures to account for the influence of task performance on metacognitive performance estimates [97].

Although the link between fluctuations in conscious awareness and metacognitive bias is not straightforward, the present finding may also provide clues on whether alterations in conscious experiences and/or a reduction of the sensory reliability of the input to the metacognitive process give rise to the ketamine-associated deterioration of metacognitive accuracy, as outlined previously [25]. Indeed, it could be suggested that the neural correlates reported there may qualify more as a neural correlate of subjective awareness [13] and thus reflect alterations in conscious experience (such as hallucinations etc.), which would be consistent with the association of involved brain areas with the posterior hot zone of conscious functions [26]. Ultimately, the unspecific nature of our previous finding (as the activation reported there was observed in a contrast aggregating over both second-order rating conditions, [25]) restrains confidence in such considerations.

To interpret ketamine effects on metacognitive evaluations and their underlying neuronal mechanisms as thoroughly as possible, it is important to ensure that these are not biased by potential effects on perceptual processes. Given the absence of statistically significant group differences in perceptual performance and response times, this could reasonably be assumed. However, not only descriptively lower average accuracy and slower type 1 response times in the ketamine group warrant a cautionary note regarding our interpretations for the behavioral and neural effects; most importantly, the group difference in relative metacognitive sensitivity or efficiency was slightly above the required significance level. Although it appears ultimately likely that this comes a result of insufficient statistical power following the extensive data exclusion due to a technical error, the possibility must be acknowledged that this non-significance may reflect the influence of relevant group-heterogeneity in perceptual performance. Importantly, however, neither staircase variability nor the initial or mean stimulus value were significantly different between the groups, which augments the relative informativeness of the ketamine effect on absolute metacognitive sensitivity.

At the neural level, the preregistered full factorial analysis yielded neither a significant main effect of Drug nor significant interactions with other factors. Likewise, we were unable to replicate the previously observed activation pattern in the Report > Follow contrast [25]. Instead, the functional connectivity patterns observed in our exploratory PPI analysis may offer a complementary explanation to the fluctuations in conscious awareness for the involvement of posterior brain areas during the ketamine challenge. Using two-sample *t*-tests, we observed increased task-specific connectivity with ketamine compared to placebo between frontal and posterior regions, namely between anterior dorsolateral PFC and temporal and posterior frontal structures, as well as between insula/IFG and a left-hemispheric occipital cluster centered in lingual gyrus.

The latter finding offers an intriguing association with Fleming et al. [29], who demonstrated increased task-specific connectivity in Report > Follow between right rostrolateral PFC and left lingual gyrus. Benedek et al. [98] also found increased connectivity between right anterior inferior parietal lobe and bilateral lingual gyrus, which they associated with internally directed attention and a potential perceptual decoupling process that shields ongoing internal processes from distracting sensory stimulation. Although no connectivity analyses were performed in our previous study, lingual gyrus was bilaterally activated more strongly under ketamine than placebo across both second-order ratings, and also displayed higher right-hemispheric activations during Report than Follow, independent of drug.

Notably, connectivity effects reported here were contralateral in both cases (see Tables 3–4), i.e. with the seed voxel located in the right and the significantly co-activated clusters in the left hemisphere. Based on this convergent evidence, one might suggest this ketamine-induced increase in task-specific connectivity to be the clearest neural correlate of impacted metacognition under ketamine obtained so far, as it could reveal a potential pattern within the neurocircuitry underlying operations of a pharmacologically challenged metacognitive system. Reframing previous arguments about a perceptual decoupling process [98], said connectivity could also be regarded as the manifestation of a compensatory mechanism to counteract the ketamine-induced loss of cause-effect information associated with each concept [50], as the brain explores an expanded repertoire of dynamical states in an unconstrained and hyper-associative fashion [99]. As it could be argued that metacognitive reports in standard experimental paradigms essentially tap such concepts or (meta-)representations, the diminished behavioral performance observed here could be indicative of this loss of information. Such an interpretation would be consistent with brain networks being less anti-correlated in the psychedelic state, according to resting-state connectivity analyses [100], perhaps accompanied by a shift from cortically centered to subcortically centered patterns of connections [101]. A more recent study suggested ketamine-associated increases in resting-state connectivity within the executive network, but decreases in salience network connectivity [102]; see Cavanna et al. [103] for a thorough account of how altered states of consciousness affect meta-stability in brain dynamics. It is worth noting, however, that results obtained in resting-state analyses of functional connectivity should be regarded as complementary to task-specific connectivity patterns as illustrated in the present study due to the limited comparability of both methods, because changes in connectivity during the resting state can indicate either alterations in connectivity between the nodes of the network or changes in activity within the network [55].

Interestingly, Fleming et al. [29] suggested their finding of increased connectivity between rostrolateral PFC and lingual gyrus to be indicative of "neural representations of object-level task uncertainty that may be then re-represented for use in metacognitive report" (p. 6123). The precision of perceptual decisions is determined by a flow of information processing from early posterior (in particular, occipital) sources, signaling a representation of accumulated decision evidence, to anterior

regions, which track internal evidence for metacognitive confidence throughout perceptual decision-making [104]. Accordingly, one could argue with respect to IIT that such re-representation is increasingly invoked under the ketamine challenge, as core areas of confidence formation rely more on information provided e.g. by the lingual gyrus, a structure known to be involved in the encoding and recollection of complex visual memories [105]. This could encompass neural representations of words in our previous study or of quantitative sets in the present study. Such an explanatory approach could also accommodate increased functional connectivity under ketamine compared to placebo between right anterior dorsolateral PFC and left middle temporal, supramarginal and precentral gyrus, as areas dedicated to higher-order metacognitive monitoring may feed off an evidence accumulation process integrating information on inter-sensory conflict during action-feedback monitoring [106] or other relevant somatosensory information, e.g. on space and limbs location [107].

4.1. Limitations

Several shortcomings of the present study have to be acknowledged. First, the neuroanatomical specificity of glutamatergic modulations is inevitably limited, as glutamate is the primary excitatory neurotransmitter of the central nervous system [108]. Another limitation concerns the study's sample size. Whilst each group was within the range or exceeded sample sizes of previous studies employing within-subject designs [109,110], the sample size may yet have been too small to detect ketamine effects beyond those reported here. This can be attributed to the extensive exclusion of participants with unsuccessful staircase calibration, and may account not only for the failure to reproduce the main effects on BOLD in the Ketamine > Placebo and Report > Follow contrasts as reported previously, but in particular to ketamine effects on relative metacognitive sensitivity or efficiency, for which we only observed a marginally significant difference between the groups. Finally, it should be noted that comparability with previous findings is limited by factors unrelated to genuine metacognition. In particular, this concerns differences in task requirements, which may generally obscure a latent domain-general factor [111]. In our previous study [25], a meta-memory task was conducted using a Yes-No response format for the first-order task, whereas the MC-T employed a 2AFC response format. Although we were nonetheless able to provide evidence that glutamatergic modulations may tap an at least partially shared neurophysiological substrate at the neurotransmitter level of both metacognitive subsystems, confidence in conclusions about variations in result patterns is limited due to this heterogeneity. In the future, direct comparisons should be carried out by applying both tasks in a single session within the same sample, and should eliminate differences in task requirements, timing of task application during infusion, and other relevant factors.

5. Conclusions

Our findings suggest that the accuracy of metacognitive evaluations in a perceptual decision-making framework is impacted as a consequence of acute ketamine administration. Building on these findings as well as previous evidence, we suggest that the integrity of the glutamatergic system at least represents a precondition for preserved metacognition. Nevertheless, given the moderate effect sizes of the reported findings, contributions from other neurotransmitter systems seem eminently plausible. The observed increases in fronto-posterior task-specific connectivity under ketamine might be indicative of re-representations of object-level features for use in metacognitive report. The generalizability of such conclusions should be elucidated by future research to help compose a fundamental account of the biological substructures that constitute the functional architecture of metacognition.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Mirko Lehmann: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. **Claudia Neumann:** Investigation, Resources, Writing – review & editing. **Sven Wasserthal:** Investigation, Writing – review & editing. **Achilles Delis:** Investigation, Resources, Writing – review & editing. **Johannes Schultz:** Supervision, Writing – review & editing. **René Hurlemann:** Resources, Writing – review & editing. **Ulrich Ettinger:** Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Acknowledgments

The authors wish to thank Paul Jung for excellent programming support and Dr. Peter Trautner for technical assistance. We are grateful to all the participants who volunteered to take part in this study.

Declarations of interest

None.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.bbr.2022.113925.

References

[1] B. Timmermans, L. Schilbach, A. Pasquali, A. Cleeremans, Higher order thoughts in action: consciousness as an unconscious re-description process, Philos. Trans. R. Soc. B Biol. Sci. 367 (2012) 1412–1423, https://doi.org/10.1098/rstb.2011.0421.
[2] J.E. Szczepanik, H. Brycz, P. Kleka, A. Fanslau, C.A. Zarate, A.C. Nugent, Metacognition and emotion – how accurate perception of own biases relates to positive feelings and hedonic capacity, Conscious. Cogn. 82 (2020), 102936, https://doi.org/10.1016/j.concog.2020.102936.
[3] J.H. Flavell, Metacognitive aspects of problem solving, in: L. Resnick (Ed.), Nat. Intell., Lawrence Erlbaum Associates, Hillsdale, NJ, 1976, pp. 231–235.
[4] J.H. Flavell, Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry, Am. Psychol. 34 (1979) 906.
[5] T.O. Nelson, L. Narens, Why investigate metacognition, in: J. Metcalfe, A. P. Shimamura (Eds.), Metacognition: Knowing about Knowing, The MIT Press, Cambridge, MA, US, 1994, pp. 1–25, https://doi.org/10.7551/mitpress/4561.001.0001.
[6] T.O. Nelson, Consciousness and metacognition, Am. Psychol. 51 (1996) 102–116, https://doi.org/10.1037//0003-066x.51.2.102.
[7] M. Allen, D. Frank, D.S. Schwarzkopf, F. Fardo, J.S. Winston, T.U. Hauser, G. Rees, Unexpected arousal modulates the influence of sensory noise on confidence, Elife 5 (2016) 1–17, https://doi.org/10.7554/eLife.18103.
[8] B. Bahrami, K. Olsen, D. Bang, A. Roepstorff, G. Rees, C. Frith, What failure in collective decision-making tells us about metacognition, Philos. Trans. R. Soc. B Biol. Sci. 367 (2012) 1350–1365, https://doi.org/10.1098/rstb.2011.0420.
[9] C. Eichner, F. Berna, Acceptance and efficacy of metacognitive training (MCT) on positive symptoms and delusions in patients with schizophrenia: a meta-analysis taking into account important moderators, Schizophr. Bull. 42 (2016) 952–962, https://doi.org/10.1093/schbul/sbv225.
[10] S. Moritz, C. Andreou, B.C. Schneider, C.E. Wittekind, M. Menon, R.P. Balzan, T. S. Woodward, Sowing the seeds of doubt: a narrative review on metacognitive training in schizophrenia, Clin. Psychol. Rev. 34 (2014) 358–366, https://doi.org/10.1016/j.cpr.2014.04.004.
[11] A. Wells, P. Fisher, S. Myers, J. Wheatley, T. Patel, C.R. Brewin, Metacognitive therapy in recurrent and persistent depression: a multiple-baseline study of a new treatment, Cogn. Ther. Res. 33 (2009) 291–300, https://doi.org/10.1007/s10608-007-9178-2.
[12] P. Grimaldi, H. Lau, M.A. Basso, 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, Neurosci. Biobehav. Rev. 55 (2015) 88–97, https://doi.org/10.1016/j.neubiorev.2015.04.006.

[13] S.M. Fleming, H.C. Lau, How to measure metacognition, Front. Hum. Neurosci. 8 (2014) 1–9, https://doi.org/10.3389/fnhum.2014.00443.
[14] L.M. Fitzgerald, M. Arvaneh, P.M. Dockree, Domain-specific and domain-general processes underlying metacognitive judgments, Conscious. Cogn. 49 (2017) 264–277, https://doi.org/10.1016/j.concog.2017.01.011.
[15] B. Baird, J. Smallwood, K.J. Gorgolewski, D.S. Margulies, Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception, J. Neurosci. 33 (2013) 16657–16665, https://doi.org/10.1523/jneurosci.0786-13.2013.
[16] S.M. Fleming, R.S. Weil, Z. Nagy, R.J. Dolan, G. Rees, Relating introspective accuracy to individual differences in brain structure, Science (80-) 329 (2010) 1541–1543, https://doi.org/10.1126/science.1191883.
[17] S.J. Galvin, J.V. Podd, V. Drga, J. Whitmore, Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions, Psychon. Bull. Rev. 10 (2003) 843–876, https://doi.org/10.3758/BF03196546.
[18] D. Rahnev, S.M. Fleming, How experimental procedures influence estimates of metacognitive ability, Neurosci. Conscious (2019) (2019), https://doi.org/10.1093/nc/niz009.
[19] B. Maniscalco, H. Lau, A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings, Conscious. Cogn. 21 (2012) 422–430, https://doi.org/10.1016/j.concog.2011.09.021.
[20] B. Maniscalco, H. Lau, Signal detection theory analysis of type 1 and type 2 data: meta-d′, response-specific meta-d′, and the unequal variance SDT model, in: Cogn. Neurosci. Metacognition, Springer, 2014: pp. 25–66. ⟨https://doi.org/10.1007/978-3-642-45190-4_3⟩.
[21] S.M. Fleming, R.J. Dolan, The neural basis of metacognitive ability, Philos. Trans. R. Soc. B Biol. Sci. 367 (2012) 1338–1349, https://doi.org/10.1098/rstb.2011.0417.
[22] T.U. Hauser, M. Allen, N. Purg, M. Moutoussis, G. Rees, R.J. Dolan, Noradrenaline blockade specifically enhances metacognitive performance, Elife 6 (2017) 1–13, https://doi.org/10.7554/eLife.24901.
[23] G. Reyes, A. Vivanco-Carlevari, F. Medina, C. Manosalva, V. de Gardelle, J. Sackur, J.R. Silva, Hydrocortisone decreases metacognitive efficiency independent of perceived stress, Sci. Rep. 10 (2020) 14100, https://doi.org/10.1038/s41598-020-71061-3.
[24] M. Clos, N. Bunzeck, T. Sommer, Dopamine is a double-edged sword: dopaminergic modulation enhances memory retrieval performance but impairs metacognition, Neuropsychopharmacology 44 (2019) 555–563, https://doi.org/10.1038/s41386-018-0246-y.
[25] M. Lehmann, C. Neumann, S. Wasserthal, J. Schultz, A. Delis, P. Trautner, R. Hurlemann, U. Ettinger, Effects of ketamine on brain function during metacognition of episodic memory, Neurosci. Conscious 2021 (2021), https://doi.org/10.1093/nc/niaa028.
[26] C. Koch, M. Massimini, M. Boly, G. Tononi, Neural correlates of consciousness: Progress and problems, Nat. Rev. Neurosci. 17 (2016) 307–321, https://doi.org/10.1038/nrn.2016.22.
[27] R. Ihalainen, O. Gosseries, F. Van de Steen, F. Raimondo, R. Panda, V. Bonhomme, D. Marinazzo, H. Bowman, S. Laureys, S. Chennu, How hot is the hot zone? Computational modelling clarifies the role of parietal and frontoparietal connectivity during anaesthetic-induced loss of consciousness, Neuroimage 231 (2021), 117841, https://doi.org/10.1016/j.neuroimage.2021.117841.
[28] R. Kupers, P. Pietrini, E. Ricciardi, M. Ptito, The nature of consciousness in the visually deprived brain, Front. Psychol. 2 (2011) 19, https://doi.org/10.3389/fpsyg.2011.00019.
[29] S.M. Fleming, J. Huijgen, R.J. Dolan, Prefrontal contributions to metacognition in perceptual decision making, J. Neurosci. 32 (2012) 6117–6125, https://doi.org/10.1523/JNEUROSCI.6489-11.2012.
[30] M.W. Tyler, H.B. Yourish, D.F. Ionescu, S.J. Haggarty, Classics in chemical neuroscience: ketamine, ACS Chem. Neurosci. 8 (2017) 1122–1134, https://doi.org/10.1021/acschemneuro.7b00074.
[31] J.H. Krystal, G. Sanacora, R.S. Duman, Rapid-acting glutamatergic antidepressants: the path to ketamine and beyond, Biol. Psychiatry 73 (2013) 1133–1141, https://doi.org/10.1016/j.biopsych.2013.03.026.
[32] L. Barnett, S.D. Muthukumaraswamy, R.L. Carhart-Harris, A.K. Seth, Decreased directed functional connectivity in the psychedelic state, Neuroimage 209 (2020), 116462, https://doi.org/10.1016/j.neuroimage.2019.116462.
[33] J. Sleigh, M. Harvey, L. Voss, B. Denny, Ketamine – more mechanisms of action than just NMDA blockade, Trends Anaesth. Crit. Care 4 (2014) 76–81, https://doi.org/10.1016/j.tacc.2014.03.002.
[34] C.G. Abdallah, L.A. Averill, K.A. Collins, P. Geha, J. Schwartz, C. Averill, K. E. DeWilde, E. Wong, A. Anticevic, C.Y. Tang, D.V. Iosifescu, D.S. Charney, J. W. Murrough, Ketamine treatment and global brain connectivity in major depression, Neuropsychopharmacology 42 (2017) 1210–1219, https://doi.org/10.1038/npp.2016.186.
[35] T.M. Hillhouse, C.R. Merritt, J.H. Porter, Effects of the noncompetitive N-methyl-D-aspartate receptor antagonist ketamine on visual signal detection performance in rats, Behav. Pharmacol. 26 (2015) 495–499, https://doi.org/10.1097/FBP.0000000000000160.
[36] I.A. Paul, P. Skolnick, Glutamate and depression: clinical and preclinical studies, Ann. N.Y. Acad. Sci. 1003 (2003) 250–272, https://doi.org/10.1196/annals.1300.016.
[37] J.W. Olney, N.B. Farber, Glutamate receptor dysfunction and schizophrenia, Arch. Gen. Psychiatry 52 (1995) 998–1007, https://doi.org/10.1001/archpsyc.1995.03950240016004.
[38] T. Schwartz, S. Sachdeva, S. Stahl, Glutamate neurocircuitry: theoretical underpinnings in schizophrenia, Front. Pharmacol. 3 (2012) 195, https://doi.org/10.3389/fphar.2012.00195.
[39] A.T. Bowdle, A.D. Radant, D.S. Cowley, E.D. Kharasch, R.J. Strassman, P.P. Roy-Byrne, Psychedelic effects of ketamine in healthy volunteers: relationship to steady-state plasma concentrations, Anesthesiology 88 (1998) 82–88, https://doi.org/10.1097/00000542-199801000-00015.
[40] F.X. Vollenweider, M. Kometer, The neurobiology of psychedelic drugs: implications for the treatment of mood disorders, Nat. Rev. Neurosci. 11 (2010) 642–651, https://doi.org/10.1038/nrn2884.
[41] M.M. Schartner, R.L. Carhart-Harris, A.B. Barrett, A.K. Seth, S. D. Muthukumaraswamy, Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin, Sci. Rep. 7 (2017) 46421, https://doi.org/10.1038/srep46421.
[42] R.L. Carhart-Harris, R. Leech, P.J. Hellyer, M. Shanahan, A. Feilding, E. Tagliazucchi, D.R. Chialvo, D. Nutt, The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs, Front. Hum. Neurosci. 8 (2014) 1–22, https://doi.org/10.3389/fnhum.2014.00020.
[43] S.D. Muthukumaraswamy, A.D. Shaw, L.E. Jackson, J. Hall, R. Moran, N. Saxena, Evidence that subanesthetic doses of ketamine cause sustained disruptions of NMDA and AMPA-mediated frontoparietal connectivity in humans, J. Neurosci. 35 (2015) 11706, https://doi.org/10.1523/JNEUROSCI.0903-15.2015 (11694 LP –).
[44] N. Farnes, B.E. Juel, A.S. Nilsen, L.G. Romundstad, J.F. Storm, Increased signal diversity/complexity of spontaneous EEG, but not evoked EEG responses, in ketamine-induced psychedelic state in humans, PLoS One 15 (2020), e0242056, https://doi.org/10.1371/journal.pone.0242056.
[45] T. Bayne, O. Carter, Dimensions of consciousness and the psychedelic state, Neurosci. Conscious 2018 (2018), https://doi.org/10.1093/nc/niy008 (niy008).
[46] C. Letheby, P. Gerrans, Self unbound: ego dissolution in psychedelic experience, Neurosci. Conscious (2017) (2017), https://doi.org/10.1093/nc/nix016.
[47] G. Tononi, Consciousness as integrated information: a provisional manifesto, Biol. Bull. 215 (2008) 216–242, https://doi.org/10.2307/25470707.
[48] G. Tononi, M. Boly, M. Massimini, C. Koch, Integrated information theory: from consciousness to its physical substrate, Nat. Rev. Neurosci. 17 (2016) 450–461, https://doi.org/10.1038/nrn.2016.44.
[49] J.R. Hanson, S.I. Walker, Formalizing falsification for theories of consciousness across computational hierarchies, Neurosci. Conscious 2021 (2021), https://doi.org/10.1093/nc/niab014.
[50] A.R. Gallimore, Restructuring consciousness –the psychedelic state in light of integrated information theory, Front. Hum. Neurosci. 9 (2015) 346, https://doi.org/10.3389/fnhum.2015.00346.
[51] M. Scheidegger, M. Walter, M. Lehmann, C. Metzger, S. Grimm, H. Boeker, P. Boesiger, A. Henning, E. Seifritz, Ketamine decreases resting state functional network connectivity in healthy subjects: implications for antidepressant drug action, PLoS One 7 (2012), e44799, https://doi.org/10.1371/journal.pone.0044799.
[52] D.F. Ionescu, J.M. Felicione, A. Gosai, C. Cusin, P. Shin, B.G. Shapero, T. Deckersbach, Ketamine-associated brain changes: a review of the neuroimaging literature, Harv. Rev. Psychiatry 26 (2018) 320–339, https://doi.org/10.1097/HRP.0000000000000179.
[53] M. Lehmann, E. Seifritz, A. Henning, M. Walter, H. Böker, M. Scheidegger, S. Grimm, Differential effects of rumination and distraction on ketamine induced modulation of resting state functional connectivity and reactivity of regions within the default-mode network, Soc. Cogn. Affect. Neurosci. 11 (2016) 1227–1235, https://doi.org/10.1093/scan/nsw034.
[54] A.G. Vaccaro, S.M. Fleming, Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements, Brain Neurosci. Adv. 2, 2018. 239821281881059. ⟨https://doi.org/10.1177/2398212818810591⟩.
[55] J.X. O'Reilly, M.W. Woolrich, T.E.J. Behrens, S.M. Smith, H. Johansen-Berg, Tools of the trade: psychophysiological interactions and functional connectivity, Soc. Cogn. Affect. Neurosci. 7 (2012) 604–609, https://doi.org/10.1093/scan/nss055.
[56] B. Becker, M. Steffens, Z. Zhao, K.M. Kendrick, C. Neumann, B. Weber, J. Schultz, M.A. Mehta, U. Ettinger, R. Hurlemann, General and emotion-specific neural effects of ketamine during emotional memory formation, Neuroimage 150 (2017) 308–317, https://doi.org/10.1016/j.neuroimage.2017.02.049.
[57] J. Morales, H. Lau, S.M. Fleming, Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex, J. Neurosci. 38 (2018) 3534–3546, https://doi.org/10.1523/JNEUROSCI.2360-17.2018.
[58] B. Baird, M. Cieslak, J. Smallwood, S.T. Grafton, J.W. Schooler, Regional white matter variation associated with domain-specific metacognitive accuracy, J. Cogn. Neurosci. 27 (2015) 440–452, https://doi.org/10.1162/jocn_a_00741.
[59] M. Rouault, A. McWilliams, M.G. Allen, S.M. Fleming, Human metacognition across domains: insights from individual differences and neuroimaging, Personal. Neurosci. 1 (2018), https://doi.org/10.1017/pen.2018.16.
[60] A. Mazancieux, S. Fleming, C. Souchay, C. Moulin, Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks, J. Exp. Psychol. Gen. 149 (2020), https://doi.org/10.1037/xge0000746.
[61] M. Lehmann, J. Hagen, U. Ettinger, Unity and diversity of metacognition, J. Exp. Psychol. Gen. (2022), https://doi.org/10.1037/xge0001197.
[62] M. Steffens, B. Becker, C. Neumann, A.M. Kasparbauer, I. Meyhöfer, B. Weber, M. A. Mehta, R. Hurlemann, U. Ettinger, Effects of ketamine on brain function during smooth pursuit eye movements, Hum. Brain Mapp. 37 (2016) 4047–4060, https://doi.org/10.1002/hbm.23294.

[63] M. Steffens, C. Neumann, A.M. Kasparbauer, B. Becker, B. Weber, M.A. Mehta, R. Hurlemann, U. Ettinger, Effects of ketamine on brain function during response inhibition, Psychopharmacology 235 (2018) 3559–3571, https://doi.org/10.1007/s00213-018-5081-7.
[64] A. Dittrich, The standardized psychometric assessment of altered states of consciousness (ASCs) in humans, Pharmacopsychiatry 31 (1998) 80–84, https://doi.org/10.1055/s-2007-979351.
[65] A. Dittrich, S. von Arx, S. Staub, International study on altered states of consciousness (ISASC): Summary of results., Ger. J. Psychol., 1985.
[66] S.M. Fleming, J. Ryu, J.G. Golfinos, K.E. Blackmon, Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions, Brain 137 (2014) 2811–2822, https://doi.org/10.1093/brain/awu221.
[67] S.M. Fleming, S. Massoni, T. Gajdos, J.-C. Vergnaud, Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance, Neurosci. Conscious. 2016 (2016), https://doi.org/10.1093/nc/niw018 (niw018).
[68] M. Allen, J.C. Glen, D. Müllensiefen, D.S. Schwarzkopf, F. Fardo, D. Frank, M. F. Callaghan, G. Rees, Metacognitive ability correlates with hippocampal and prefrontal microstructure, Neuroimage 149 (2017) 415–423, https://doi.org/10.1016/j.neuroimage.2017.02.008.
[69] A.M. Kasparbauer, N. Petrovsky, P.M. Schmidt, P. Trautner, B. Weber, B. Sträter, U. Ettinger, Effects of nicotine and atomoxetine on brain function during response inhibition, Eur. Neuropsychopharmacol. 29 (2019) 235–246, https://doi.org/10.1016/j.euroneuro.2018.12.004.
[70] A.C. Evans, S. Marrett, P. Neelin, L. Collins, K. Worsley, W. Dai, S. Milot, E. Meyer, D. Bub, Anatomical mapping of functional activation in stereotactic coordinate space, Neuroimage 1 (1992) 43–53, https://doi.org/10.1016/1053-8119(92)90006-9.
[71] P.T. Fox, J.S. Perlmutter, M.E. Raichle, A stereotactic method of anatomical localization for positron emission tomography, J. Comput. Assist. Tomogr. 9 (1985) 141–153.
[72] J. Ashburner, K.J. Friston, Unified segmentation, Neuroimage 26 (2005) 839–851, https://doi.org/10.1016/j.neuroimage.2005.02.018.
[73] K.J. Friston, J. Ashburner, C.D. Frith, J. Poline, J.D. Heather, R.S.J. Frackowiak, Spatial registration and normalization of images, Hum. Brain Mapp. 3 (1995) 165–189, https://doi.org/10.1002/hbm.460030303.
[74] S.B. Eickhoff, K.E. Stephan, H. Mohlberg, C. Grefkes, G.R. Fink, K. Amunts, K. Zilles, A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data, Neuroimage 25 (2005) 1325–1335, https://doi.org/10.1016/j.neuroimage.2004.12.034.
[75] K.J. Friston, C. Buechel, G.R. Fink, J. Morris, E. Rolls, R.J. Dolan, Psychophysiological and modulatory interactions in neuroimaging, Neuroimage 6 (1997) 218–229, https://doi.org/10.1006/nimg.1997.0291.
[76] R. Schröder, A.-M. Kasparbauer, I. Meyhöfer, M. Steffens, P. Trautner, U. Ettinger, Functional connectivity during smooth pursuit eye movements, J. Neurophysiol. 124 (2020) 1839–1856, https://doi.org/10.1152/jn.00317.2020.
[77] D.G. McLaren, M.L. Ries, G. Xu, S.C. Johnson, A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches, Neuroimage 61 (2012) 1277–1286, https://doi.org/10.1016/j.neuroimage.2012.03.068.
[78] K. Weidacker, S.-G. Kim, C.L. Nord, C. Rua, C.T. Rodgers, V. Voon, Avoiding monetary loss: a human habenula functional MRI ultra-high field study, Cortex 142 (2021) 62–73, https://doi.org/10.1016/j.cortex.2021.05.013.
[79] J.M. Cisler, K. Bush, J.S. Steele, A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI, Neuroimage 84 (2014) 1042–1052, https://doi.org/10.1016/j.neuroimage.2013.09.018.
[80] J. Kim, B. Horwitz, Investigating the neural basis for fMRI-based functional connectivity in a blocked design: application to interregional correlations and psycho-physiological interactions, Magn. Reson. Imaging 26 (2008) 583–593, https://doi.org/10.1016/j.mri.2007.10.011.
[81] D.G. Green, J.A. Swets, Signal Detection Theory and Psychophysics, 1966, https://doi.org/10.1901/jeab.1969.12-475.
[82] A.B. Barrett, Z. Dienes, A.K. Seth, Measures of metacognition on signal-detection theoretic models, Psychol. Methods 18 (2013) 535, https://doi.org/10.1037/a0033268.
[83] S.M. Fleming, HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings, Neurosci. Conscious 2017 (2017) 1–14, https://doi.org/10.1093/nc/nix007.
[84] M. Guggenmos, Measuring metacognitive performance: type 1 performance dependence and test-retest reliability, Neurosci. Conscious (2021) (2021), https://doi.org/10.1093/nc/niab040.
[85] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, Stat. Sci. 7 (1992) 457–472, https://doi.org/10.1214/ss/1177011136.
[86] R.D. Morey, J.N. Rouder, BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12–4.3., 2021. ⟨https://cran.r-project.org/package=BayesFactor⟩ (Accessed 12 February 2022).
[87] M. Brett, J.-L. Anton, R. Valabregue, J.-B. Poline, Region of interest analysis using an SPM toolbox, in: 8th Int. Conf. Funct. Mapp. Hum. Brain, Sendai, Japan, 2002: p. 497.
[88] J. Cohen, A. Maydeu-Olivares, A. Power Primer, Psychol. Bull. (1992), https://doi.org/10.1038/141613a0.
[89] M. Allen, D. Poggiali, K. Whitaker, T.R. Marshall, R.A. Kievit, Raincloud plots: a multi-platform tool for robust data visualization, Wellcome Open Res. 4 (2019) 63, https://doi.org/10.12688/wellcomeopenres.15191.1.
[90] A.V. Utevsky, D.V. Smith, S.A. Huettel, Precuneus is a functional core of the default-mode network, J. Neurosci. 34 (2014) 932–940, https://doi.org/10.1523/JNEUROSCI.4227-13.2014.
[91] J. Ais, A. Zylberberg, P. Barttfeld, M. Sigman, Individual consistency in the accuracy and distribution of confidence judgments, Cognition 146 (2016) 377–386, https://doi.org/10.1016/j.cognition.2015.10.006.
[92] R.F. West, K.E. Stanovich, The domain specificity and generality of overconfidence: individual differences in performance estimation bias, Psychon. Bull. Rev. 4 (1997) 387–392, https://doi.org/10.3758/BF03210798.
[93] J. Frohlich, J.D.Van Horn, Reviewing the ketamine model for schizophrenia, J. Psychopharmacol. 28 (2014) 287–302, https://doi.org/10.1177/0269881113512909.
[94] J.S. Mayer, M. Stäblein, V. Oertel-Knöchel, C.J. Fiebach, Functional dissociation of confident and not-confident errors in the spatial delayed response task demonstrates impairments in working memory encoding and maintenance in schizophrenia, Front. Psychiatry 9 (2018) 202, https://doi.org/10.3389/fpsyt.2018.00202.
[95] S. Moritz, A.S. Göritz, J. Gallinat, M. Schafschetzy, N. Van Quaquebeke, M.J. V. Peters, C. Andreou, Subjective competence breeds overconfidence in errors in psychosis. A hubris account of paranoia, J. Behav. Ther. Exp. Psychiatry 48 (2015) 118–124, https://doi.org/10.1016/j.jbtep.2015.02.011.
[96] E. Bristow, P. Tabraham, N. Smedley, T. Ward, E. Peters, Jumping to perceptions and to conclusions: specificity to hallucinations and delusions, Schizophr. Res. 154 (2014) 68–72, https://doi.org/10.1016/j.schres.2014.02.004.
[97] M. Rouy, P. Saliou, L. Nalborczyk, M. Pereira, P. Roux, N. Faivre, Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders, Neurosci. Biobehav. Rev. 126 (2021) 329–337, https://doi.org/10.1016/j.neubiorev.2021.03.017.
[98] M. Benedek, E. Jauk, R.E. Beaty, A. Fink, K. Koschutnig, A.C. Neubauer, Brain mechanisms associated with internally directed attention and self-generated thought, Sci. Rep. 6 (2016) 22959, https://doi.org/10.1038/srep22959.
[99] E. Tagliazucchi, R. Carhart-Harris, R. Leech, D. Nutt, D. Chialvo, Enhanced repertoire of brain dynamical states during the psychedelic experience, Hum. Brain Mapp. 35 (2014), https://doi.org/10.1002/hbm.22562.
[100] L. Roseman, R. Leech, A. Feilding, D.J. Nutt, R.L. Carhart-Harris, The effects of psilocybin and MDMA on between-network resting state functional connectivity in healthy volunteers, Front. Hum. Neurosci. 8 (2014) 204, https://doi.org/10.3389/fnhum.2014.00204.
[101] R. Joules, O.M. Doyle, A.J. Schwarz, O.G. O'Daly, M. Brammer, S.C. Williams, M. A. Mehta, Ketamine induces a robust whole-brain connectivity pattern that can be differentially modulated by drugs of different mechanism and clinical profile, Psychopharmacology 232 (2015) 4205–4218, https://doi.org/10.1007/s00213-015-3951-9.
[102] F. Mueller, F. Musso, M. London, P. de Boer, N. Zacharias, G. Winterer, Pharmacological fMRI: Effects of subanesthetic ketamine on resting-state functional connectivity in the default mode network, salience network, dorsal attention network and executive control network, NeuroImage Clin. 19 (2018) 745–757, https://doi.org/10.1016/j.nicl.2018.05.037.
[103] F. Cavanna, M.G. Vilas, M. Palmucci, E. Tagliazucchi, Dynamic functional connectivity and brain metastability during altered states of consciousness, Neuroimage 180 (2018) 383–395, https://doi.org/10.1016/j.neuroimage.2017.09.065.
[104] T. Balsdon, P. Mamassian, V. Wyart, Separable neural signatures of confidence during perceptual decisions, Elife 10 (2021), e68491, https://doi.org/10.7554/eLife.68491.
[105] W.C. Machielsen, S.A. Rombouts, F. Barkhof, P. Scheltens, M.P. Witter, FMRI of visual encoding: reproducibility of activation, Hum. Brain Mapp. 9 (2000) 156–164, https://doi.org/10.1002/(sici)1097-0193(200003)9:3<156::aid-hbm4>3.0.co;2-q.
[106] B.M. van Kemenade, B.E. Arikan, K. Podranski, O. Steinsträter, T. Kircher, B. Straube, Distinct roles for the cerebellum, angular gyrus, and middle temporal gyrus in action–feedback monitoring, Cereb. Cortex 29 (2019) 1520–1531, https://doi.org/10.1093/cercor/bhy048.
[107] C.L. Reed, R.J. Caselli, The nature of tactile agnosia: a case study, Neuropsychologia 32 (1994) 527–539, https://doi.org/10.1016/0028-3932(94)90142-2.
[108] B. Moghaddam, D. Javitt, From revolution to evolution: the glutamate hypothesis of schizophrenia and its implication for treatment, Neuropsychopharmacology 37 (2012) 4–15, https://doi.org/10.1038/npp.2011.181.
[109] A.M. Van Loon, J.J. Fahrenfort, B. Van Der Velde, P.B. Lirk, N.C.C. Vulink, M. W. Hollmann, H. Steven Scholte, V.A.F. Lamme, NMDA receptor antagonist ketamine distorts object recognition by reducing feedback to early visual cortex, Cereb. Cortex 26 (2016) 1986–1996, https://doi.org/10.1093/cercor/bhv018.
[110] M. Scheidegger, A. Henning, M. Walter, H. Boeker, A. Weigand, E. Seifritz, S. Grimm, Effects of ketamine on cognition-emotion interaction in the brain, Neuroimage 124 (2016) 8–15, https://doi.org/10.1016/j.neuroimage.2015.08.070.
[111] E. Ruby, N. Giles, H. Lau, Finding domain-general metacognitive mechanisms requires using appropriate tasks, BioRxiv (2017), 211805, https://doi.org/10.1101/211805.