Metacognition, the ability to reflect on and evaluate our own thoughts and actions, is a crucial component of human behavior and subjective experience (1). A wealth of empirical studies have shown that impaired metacognition is associated with detrimental behavior and poor mental health (2,3). For instance, delusional thinking in patients with schizophrenia is thought to be maintained by metacognitive deficits such as a lack of insight (4) or overconfidence in incorrect models of the world (5–7). In a range of mental health conditions, metacognition shows consistent, yet specific, individual differences (6,9) [see review (2)], findings that generalize across various tasks (10) and cognitive domains (11), and abnormalities that may be heritable (12). As researchers in psychiatry aim to develop reliable neurocognitive markers for identifying current and future mental health problems, metacognitive assessments hold promise (13).

There are several challenges in meeting this aim. First, metacognition is often confounded with cognitive performance, such as the accuracy of visual decisions or memory recollection. Second, metacognition manifests in various hierarchical levels of abstraction, from local confidence in isolated decisions and global self-beliefs about our abilities and skills. Alterations in metacognition are considered foundational to neurologic and psychiatric disorders, but research has mostly focused on local metacognitive computations, missing out on the role of global aspects of metacognition. Here, we first review current behavioral and neural metrics of local metacognition that lay the foundation for this research. We then address the neurocognitive underpinnings of global metacognition uncovered by recent studies. Finally, we outline a theoretical framework in which higher hierarchical levels of metacognition may help identify the role of maladaptive metacognitive evaluation in mental health conditions, particularly when combined with transdiagnostic methods.

Methods for Quantifying Local Metacognition

Behavioral and Computational Metrics of Local Metacognition

Several metrics have been developed to quantify local metacognition in laboratory tasks, most of which rely on examining the correspondence between objective performance and confidence ratings [a subjective report of being correct about a decision/statement (14)] across multiple experimental trials (Figure 1A). Two independent aspects of local metacognition can be distinguished: metacognitive bias and sensitivity (15) (Figure 1B). Metacognitive bias reflects how confident we are irrespective of actual performance, and this is usually estimated as the mean confidence rating averaged over correct and incorrect judgments. In contrast, metacognitive sensitivity reflects an ability to discriminate correct from incorrect judgments. A participant who rates high confidence on incorrect judgments and low confidence on correct judgments is estimated to have high metacognitive sensitivity.

An initial wave of studies relied on simple correlation statistics, which conflated metacognitive bias and sensitivity in one measure, an issue covered previously (16,17). More recent methods (i.e., type 2 signal detection theory) estimate a bias-free assessment of metacognitive sensitivity (18). However, metacognitive sensitivity is typically dependent on task performance, where easier tasks produce greater sensitivity (16). Model-based methods reliant on signal detection theory (e.g.,
meta’ model) correct for such performance confounds, leading to the derivation of summary statistics such as metacognitive efficiency that represent a participant’s level of metacognitive sensitivity corrected for variation in task performance (17). Another approach is to use staircase procedures (19,20) that adjust task performance at a predetermined level, allowing variation in metacognitive sensitivity to be isolated [e.g., (61)], although this method has caveats (21). Failures to replicate metamemory biases toward lowered confidence in obsessive-compulsive disorder (OCD) (22–31) or recent evidence of previously inflated effects (32) of higher confidence in errors from patients with schizophrenia who are delusion prone and have paranoia (5,6,33–38) were ultimately explained by metacognitive sensitivity and bias not being properly separated. Future experiments should aim to minimize potential confounds in estimating metacognitive sensitivity at either the paradigm design or analysis stage.

Neural Bases of Local Metacognition
Beyond behavioral metrics, studies have begun to reveal the neural bases of local metacognition about perception and memory [see review (11)]. Strong convergent evidence highlights the importance of prefrontal cortex (PFC) for metacognition. Lesions (39) or transcranial magnetic stimulation (40) to the PFC affect perceptual metacognitive sensitivity while leaving task performance unaffected. Structural and functional magnetic resonance imaging studies in healthy humans have linked individual differences in anterior PFC volume, function, and connectivity to metacognitive ability (6,20,41–47). Beyond PFC, a distributed network of brain regions including the medial PFC, precuneus, and hippocampus (20,43,44,46–52) are also involved in metacognition. Electrophysiology studies provided convergent evidence of activity associated with metacognition in prefrontal theta oscillations (53), the P3 event-related potential component (54), and the error-related negativity (55–57). Similar neural correlates are observed in relation to aberrant metacognitive processes in some psychiatric disorders. Altered metacognition about perceptual decisions in patients with schizophrenia correlates with hypoaactivity in frontoparietal areas (58) and also hippocampal volume and its gray matter microstructure (59). Drug addiction, which was linked to deficits in error awareness (60) and perceptual metacognitive sensitivity (61), was associated with hypoaactivity and a loss of structural integrity in the anterior cingulate cortex. Overall, the medial PFC and parietal cortex are proposed to play a domain-general role in metacognition, with other nodes of the network contributing in a domain-specific fashion (11) (Figure 2).

FROM LOCAL CONFIDENCE TO GLOBAL SELF-BELIEFS
Many Forms of Metacognition Coexist
While the psychological and neural bases of local metacognition are increasingly well characterized, its functional roles remain less clear. Local confidence has been suggested to regulate subsequent decisions by recruiting cognitive control (62), gathering information (63), controlling exploration (64), and adapting speed-accuracy trade-offs (65). However, these are all limited in scope and on short time scales. In contrast to local confidence in single decisions, global metacognitive evaluations of performance (i.e., self-beliefs) can span several decisions or experimental trials, allowing for a gradual formation of self-performance estimates in numerous aspects: about our ability on a given task, in a specific cognitive domain, or even how capable we feel, broadly (Figure 3). In turn, these self-beliefs may affect future decisions on longer time scales (66,67), such as promoting the initiation of behavioral sequences toward achieving a goal. Individuals with low self-beliefs tend to feel less in control of their environment, are less likely to believe that their decisions will affect future outcomes, and are slower to recover after setbacks (68,69). Accordingly, distorted self-beliefs may have a pervasive impact in educational and clinical settings (70), determining how people see themselves and their capabilities. However, despite their recognized importance for mental health, the cognitive and neural foundations of self-beliefs remain largely unclear.

Self-beliefs are related to the psychological construct of self-esteem, a global notion of self-worth that cuts across many domains (e.g., physical, social, and academic) (71). Low self-esteem is a key predictor of mental health issues such as anxiety and depression (72,73). Low self-esteem has strong theoretical ties to dominant clinical psychology models of depression (74), where depressive symptoms are thought to be grounded in negative schema that persist despite alternative evidence (75). Negative schemas encompass several processes, among which confidence/self-beliefs is one critical aspect, with the proposed neural correlates of negative schemas and confidence partly overlapping, e.g., cingulate cortex (76). However, despite the strong face validity of these negative schema, their measurement with clinical scales precludes a mechanistic understanding of how these self-reports arise (77). In contrast, models of global metacognition constitute a mechanistic framework within which to define testable hypotheses and unpack the mechanisms underpinning low self-beliefs. For instance, we can examine how shifts in processes supporting local decision confidence lead to gradual changes in global self-beliefs that likely unfold over longer timescales. The study of apathy provides a recent example—a single self-report (i.e., apathetic state) could be attributed to various computational mechanisms (reduced reward sensitivity or increased subjective perception of effort), each associated with distinct neurobiological systems (78,79).

Neurocognitive Foundations of Simple Forms of Global Self-beliefs
We have begun to delineate computations underlying the formation of global self-beliefs from local confidence estimates (80,81). In these experiments, participants were asked to perform mini blocks of two interleaved perceptual tasks. At the end of each block, they selected the task which they thought they performed best—a proxy for global self-beliefs about the two tasks. Local subjective confidence ratings were found to predict global self-beliefs over and above objective performance (80). Using functional magnetic
resonance imaging, we further found that ventral striatal activity reflected the level of global self-beliefs (but not local confidence signals), while confidence-related activity in ventromedial PFC was further modulated by the level of global self-belief (81). This is in line with two studies indicating that ventromedial PFC reflects fluctuations in self-performance estimates on mini games performed across several trials when participants monitor expected task success with (64) or without (82) external feedback. Moreover, white matter structural integrity between ventral striatum and ventromedial PFC, estimated using diffusion tensor imaging, shows systematic links with individual self-esteem (83). These results establish an initial link between local and global metacognition (Figure 2) and reveal neural representations of global self-beliefs that go beyond the tracking of local confidence (84).

It is important to acknowledge that global self-beliefs assessed in these studies were limited to the scope of a laboratory experiment and to perceptual (80,84) or color/time estimation (82) tasks. These tasks are well characterized in terms of how local perceptual decisions and confidence estimates are formed [e.g., (85)], which is vital for precisely quantifying how self-beliefs are constructed from local confidence and external feedback (86). However, there is a substantial gap between experimental investigations of so-called global self-beliefs and self-beliefs relevant to real-life decisions, which typically fluctuate over considerably longer time scales than those assessed in the laboratory. In addition, many other factors contribute to the formation of real-life self-beliefs, such as feedback from other people and one’s social environment (87,88). We suggest that we can bridge the gap by examining how self-beliefs generalize across different tasks and across cognitive domains (Figure 3). Such a generalization mechanism should normally support the formation of useful priors about expected ability in closely related tasks, but if this mechanism becomes maladaptive, leading to, e.g., excessive generalization from local experiences, it could create volatile self-beliefs. Conversely, a disruption in updating mechanisms could result in rigid self-beliefs being insufficiently updated in light of new positive experiences.

**Relating Global Self-beliefs to Functional Symptoms**

Adapting a framework for global metacognition may prove useful clinically because it may be more directly relevant to the subjective and functional experiences of patients as compared...
How Local and Global Metacognition Shape Mental Health

Figure 2. Neural correlates of metacognitive evaluation. Schematic sagittal slice and lateral view of the human brain highlighting the role of prefrontal cortex (PFC) in metacognition. Studies of local metacognition have highlighted the ventromedial PFC (vmPFC) and posterior medial frontal cortex (pMFC) as central hubs reflecting confidence estimates [a: (9,44,46,149)] and error detection [b: (146-148)], while the frontopolar cortex (FPC), together with the lateral PFC (laPFC), is involved in mediating explicit metacognitive judgments, (meta)cognitive control, and subsequent behavioral regulation [c: (6,47); d: (40,43)]. Some of the neural substrates linked to local metacognition exhibit cognitive domain-specificity, e.g., the precuneus (PRECU) has mostly been implicated in metamemory [e: (6,95,44,46,149)], while lateral-parietal areas (laPAR) are mostly implicated in metaperception [f: (47,150)]. Recent studies have begun to reveal that neural substrates of global metacognition only partly overlap with those of local metacognition. In particular, in vmPFC and PRECU, local confidence signals were found to be further modulated by the level of global self-belief on a perceptual task (81).

with local confidence in isolated decisions. For example, anosognosia, defined as a lack of awareness of cognitive deficits, particularly about memory, is a common symptom of dementia (86). A lack of self-awareness may lead to a failure to adapt to changes in cognitive abilities, for instance, leading to risky behaviors such as driving long distances or traveling to unfamiliar locations (89). Anosognosia may also affect decisions about appropriate courses of treatment or prevent the implementation of strategies to aid memory such as setting reminders (90,91). Similarly, intact global metacognition may be crucial for treatment adherence as an individual may only be willing to participate in therapeutic interventions if they have insight into their symptoms. Previous work with patients with schizophrenia has indeed shown that clinical insight is predictive of medication compliance (92,93).

At present, only local confidence is routinely measured in experimental studies of metacognition. However, there is likely a complex and largely unexplored interplay between local metacognitive evaluations and global self-beliefs. Notably, anosognosia may coexist with relatively intact local metacognition about performance on individual trials. In these studies (94–96), participants with Alzheimer’s disease underwent assessments of local metacognition on memory and motor tasks, and clinicians evaluated the patients’ global awareness of their deficits (95). While both local memory and motor metacognition were found to be relatively intact (89,95), there was a specific deficit of global awareness in the memory (and not motor) domain (95), suggesting that local and global metacognitive levels may dissociate in some cognitive domains but not others. We note, however, that extended clinical interviews and/or informants’ reports were used as proxies for ground-truth ability; therefore, the data remain disconnected from approaches that seek to model the relationship between performance and confidence.

Global and local metacognition also diverge in Parkinson’s disease. Patients differ from healthy participants in their feeling of knowing accuracy in recognition memory tests at the item level but not in their global prediction of accuracy (97). These examples highlight the value of a neurocognitive framework encompassing local and global metacognition to pinpoint the origins of lack of awareness (80). It could be that symptom severity only affects upper hierarchical levels (Figure 3) or creates imbalances between global and local metacognitive processing within a specific domain. Similar to anosognosia, functional cognitive disorder, a condition characterized by the experience of persistent and distressing subjective cognitive difficulties in the absence of detectable objective cognitive deficit and underlying neurologic disease (98,99), is thought to be explained by changes in metacognitive ability. However, it is unknown which layer(s) of the metacognitive hierarchy, if
any, is affected in this condition. Likewise, patients with motor conversion disorder report difficulties in performing certain motor actions without any apparent neurologic disease. Previous work using a visuomotor task revealed that patients are just as aware and confident in trajectory deviations as control participants, but they engaged distinct brain networks when estimating their confidence (100). In this case, distortions in the formation of global self-beliefs may be central in explaining a mismatch between an internal subjective experience of poor self-ability and otherwise intact objective performance and local metacognition (Figure 3).

The various layers in a putative metacognitive hierarchy are likely to be more fine-grained than the local/global dichotomy highlighted here. For instance, we can make a distinction between “how well did I perform this task today at work?” and “how well am I performing at my job in general?” The levels of metacognition outlined here (Figure 3) partly map onto a previously proposed psychological framework for characterizing global awareness in dementia (101) that distinguishes four levels: sensory preregistration (basic evaluation), performance monitoring (corresponding to so-called local metacognition here), evaluative judgment, and metarepresentation. However, in this model, the latter two constructs were defined in relation to how others see us, rather than in relation to objective experimental measurements.

Interim Conclusion

Building a complete theoretical framework supported by empirical evidence of how various levels of metacognition relate to each other is important because global self-beliefs are a major determinant of our behavior. Unlike local metacognition, which is often tied to a particular task or cognitive domain, changes in global self-beliefs may generalize to other domains and to a range of daily life functions (89). In turn, global self-beliefs may be more directly relevant for understanding the mechanistic and computational bases of global aspects of subjective experience such as low mood or self-esteem characteristic of negative schemas in depression (80).

A TRANSDIAGNOSTIC APPROACH FOR UNCOVERING ASSOCIATIONS BETWEEN METACOGNITION AND MENTAL HEALTH SYMPTOMS

If local and global metacognition are to be neurocognitive markers for psychopathology, their robustness and specificity are important. Psychiatric research suggests that the use of the DSM categories poses a concern for these goals (102). Hence, accounting for comorbid symptoms appears to be crucial for understanding the precise clinical consequences of abnormalities in metacognition and ultimately allows us to map symptoms more closely to behavior and neural circuits (102–104) (Figure 4C).

Figure 3. Reciprocal interactions between multiple hierarchical levels of metacognitive evaluation. Previous work has revealed that local confidence contributes to the formation of global self-beliefs, but global self-beliefs are also likely to in turn influence local confidence. Under this framework, local confidence may reflect a combination of a local component related to decision performance evaluation and a global component formed over the aggregation of multiple experiences across various tasks and domains formed through learning. On the right, examples are given to illustrate each hierarchical level in the domain of memory, although the true distinction between levels is likely to be more fine-grained. Each of these metacognitive levels is associated with dynamics unfolding across different timescales, with higher levels of the hierarchy having slower dynamics than lower levels. Global self-beliefs may shape and be shaped by even more global constructs such as an individual’s level of self-esteem.

Transdiagnostic Studies of Local Metacognition

Recent studies have leveraged transdiagnostic approaches to uncover links between symptom dimensions and metacognition. With self-reported symptoms in nine psychiatric questionnaires (105), we characterized large online general population samples along three symptom dimensions [anxious-depression, compulsive behavior and intrusive thought (henceforth compulsivity), and social withdrawal; replicated from a previous study (106)]. Using a perceptual decision-making task and local confidence ratings, we found that the anxious-depression dimension was associated with lower confidence, whereas the compulsivity dimension was related to higher confidence (Figure 5). These results stand in contrast to classic questionnaire scores showing that OCD symptoms alone were not linked to any alterations in confidence (Figure 5), similar to previous findings (107,108). This is because anxiety and depression, which are both linked to lower local confidence judgments, overlapped with OCD scores (109,110), masking a positive association between confidence and compulsivity. These findings suggest that metacognitive dysfunctions previously observed may be masked by the co-occurrence of other symptoms, particularly if different families of symptoms predict opposing effects on confidence.

A transdiagnostic approach therefore provides context for interpreting previous metacognition findings in case-control studies of OCD. Vaghi et al. (108) employed a reinforcement learning task where participants predicted where a particle will land and report their confidence in catching the particle. They observed a form of decreased metacognitive sensitivity in OCD.
as compared with healthy participants (smaller correlation between confidence and behavioral adjustments of their prediction), without a difference in local confidence or in how sensitive participants’ confidence was to task events (e.g., sudden changes in landing location). Conducting the same paradigm in a large online general population sample, we replicated Vaghi et al.’s (108) finding of an impaired relationship between confidence and behavioral adjustments in OCD (111). Using a dimensional approach, we also found that higher confidence [as in the perceptual task (110)] (Figure 5) and a lower sensitivity of confidence to task events were linked to compulsivity symptoms. These studies demonstrate that transdiagnostic approaches can be crucial in delineating hidden metacognitive relationships and enhancing our understanding of psychopathology.

To our knowledge, the transdiagnostic studies presented above are the only ones applying such approaches to local metacognitive metrics. By using the same three-dimensional structure across multiple studies, we can prevent the overfitting of new psychiatric dimensions to data. Indeed, the same compulsivity dimension linked to metacognitive deficits (105,111) is also associated with goal-directed failures (106), enhanced learning from safety than threat (112), reduced avoidance of cognitive effort (113), and faulty neural representations of task structure knowledge (114). In the case of goal-directed control, deficits are seen in online (106) and in-person (114) samples alike, and work in patients has shown that these deficits are more strongly linked to variation in a compulsive dimension than a diagnosis of OCD (115). Although these findings are suggestive, it remains to be seen if the metacognitive abnormalities associated with these dimensions are also altered in patient samples. We also note that these dimensions may not necessarily describe cognitive alterations better than DSM-defined psychopathology or other transdiagnostic structures (116–118). Alternative dimensional or hierarchical approaches to phenotyping (119) remain to be tested in the context of metacognition (120–122). As psychiatry continues to improve how we define mental health and illness in the population, we can expect cycles of iterative evolution of dimensional phenotypes (both those of interest and those to be controlled for) (123).

**Intersecting Hierarchies of Metacognition With Transdiagnostic Approaches**

Transdiagnostic approaches have revealed that individuals with strong anxious-depression symptoms have lower local confidence, whereas those with compulsivity have higher confidence (105,111). However, the same individual can experience both anxiety and compulsivity symptoms (e.g., OCD). We argue that such opposing effects of confidence between anxious-depression and compulsivity may be unraveled by better distinguishing between local confidence and global self-beliefs. It is likely that an individual’s local belief

Figure 4. Dimensional approaches to psychiatry addressing within- and across-diagnosis homogeneities and heterogeneities. (A, B) Case-control studies comparing diagnosed patient and healthy control groups have often failed to recognize that patients have varying levels of other psychopathologies (e.g., compulsivity, anxiety) beyond the one under study. Comparing such groups (typically, ranging between 15 and 50 participants per group) has often revealed ambiguous or nonspecific effects in relation to metacognition. (C) Mathematical methods of dimensionality reduction allow identification of latent factors underlying various mental health conditions. These dimensions may better reflect the psychopathological complexity underlying traditional psychiatric categories and uncover more consistent relationships with metacognition. Obsessive-compulsive disorder (OCD) and generalized anxiety disorder (GAD) reflect traditional diagnostic categories. In contrast, anxious and compulsive dimensions reflect transdiagnostic symptom dimensions. Typically, transdiagnostic dimensions are estimated using groups of hundreds or thousands of participants.
about performance is not pure and instead involves numerous, and at least partially dissociable, neural and computational processes. Local confidence ratings in anxious-depression may be contaminated, i.e., driven by global estimates of self-performance unrelated to the current task, while local confidence ratings in compulsion could reflect selective abnormalities in local evidence evaluation processes. This explanation is supported by observations that anxiety and depression symptoms are strongly linked to low self-esteem (72,73), while compulsion is associated with difficulties in developing and using models to solve decision-making tasks (114,124). In sum, a local confidence rating could depend both on a global prior about self-ability and a local evaluation of performance.

Patients with schizophrenia have frequently been reported to be overconfident about individual (local) decisions (2,37). However, recent moderation analyses suggest that this metacognitive deficit is based on studies in which other cognitive performance features vary across participants, thereby questioning whether the overconfidence effect is a central deficit (32). This issue is likely exacerbated by the inclusion of variable diagnoses (e.g., bipolar disorder or depression with psychosis) beyond schizophrenia in previous studies (32). Certain forms of schizophrenia also include high levels of apathy, which could be partly linked to low global subjective expectations of success (125). As positive and negative symptoms coexist in schizophrenia, combining a transdiagnostic perspective while considering different levels of metacognition may be fundamental to delineating the underlying psychopathology. For this reason, we advocate that future studies use tasks that can distinguish, and simultaneously control for, multiple levels of metacognition (80). Cross-task comparisons might prove useful too, as we hypothesize that reductions in local confidence in depression, if driven by global self-beliefs, should be relatively impervious to task design, and generalize across domains (10).

In contrast, if local confidence biases in compulsion are the result of an issue with model building, we expect the finding of overconfidence to be highly sensitive to task demands.

**Clinical Implications**

Metacognitive beliefs have long been a therapeutic target. Metacognitive therapy for anxiety, depression (70,126,127), OCD (128,129), and schizophrenia (130,131) focus on modifying intrusive thoughts and cognitive biases to dampen mal-adaptive rumination, compulsive rituals, or delusional ideation.

However, efficacy of metacognitive therapy is not useful for all patients (132–134), and little is known about the underlying neural mechanisms facilitating symptom alleviation (135). Assessing metacognition before and after metacognitive therapy should help formalize a mechanistic and neural model of how clinical gains occur and establish if it is through metacognitive processes. Meanwhile, recent studies have shown that training can improve metacognitive ability (136,137) [although with exceptions (138)]. A next step is to examine if these metacognitive changes have therapeutic benefit, that is, transfer beyond a particular training or therapeutic session and generalize to real-world functioning. Gaining an understanding of the factors promoting generalization will be critical for devising tools for improving metacognition (136,137,139) and modifying self-beliefs through psychotherapy (70,140).

The current evidence for a relationship between mental health and metacognition is correlational. Translating these insights to the clinic requires probing these associations causally and in longitudinal designs. A key question is whether abnormalities in metacognitive bias and sensitivity resolve when symptoms improve or are relatively stable traits that may signal an overall risk for developing a mental health condition. Drawing on adjacent literature, there is some evidence to suggest that negative biases in face perception improve following antidepressant drug administration in patients with...
depression and predict subsequent clinical response (141). If metacognitive bias follows a similar pattern as negative biases, it may similarly constitute a predictor of treatment outcome. Quantifying metacognition could therefore have clinical value if changes in metacognitive parameters help to identify individuals at risk, facilitate early intervention, guide us as to who might respond best to a given treatment, or assist in developing transdiagnostic treatment protocols that target metacognition (142–144).

CONCLUSIONS

Theories about the role of metacognition in mental health may be enriched by adopting quantitative task-based methods for measuring metacognition across different hierarchical levels (Figure 3) together with robust transdiagnostic approaches (Figure 4). Many other aspects of metacognition have yet to be looked at in relation to mental health, and the paradigms and models described here represent a starting point. The current review serves as a framework for thinking about how different levels of metacognition (from local to global) are interrelated, possibly by generalization mechanisms, and outlines hypotheses for how these map onto transdiagnostic dimensions of mental health.

ACKNOWLEDGMENTS AND DISCLOSURES

TXFS is a postdoctoral fellow at the Max Planck UCL Centre for Computational Psychiatry and Ageing Research. The Max Planck UCL Centre for Computational Psychiatry and Ageing Research is a joint initiative supported by UCL and the Max Planck Society. The Wellcome Centre for Human Neuroimaging is supported by core funding from the Wellcome Trust (203147/Z/16/Z). MR is the beneficiary of a postdoctoral fellowship from the AXA Research Fund. His work was also supported by a department-wide grant from the Agence Nationale de la Recherche (ANR-17-EURE-0017, EUR FrontCog). This work has received support under the program «Investissements d’Avenir» launched by the French Government and implemented by ANR (ANR-10-IDEX-0001-02 PSL). CMG is supported by a fellowship from MQ: transforming mental health (MQ16IP13) and holds grant funding from Science Foundation Ireland’s Frontiers for the Future Award (19/FFP/6418). SMF is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and Royal Society (206648/Z/17/Z). The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Max Planck UCL Centre for Computational Psychiatry and Ageing Research (TXFS, SMF), Wellcome Centre for Human Neuroimaging (TXFS, SMF), and Department of Experimental Psychology (SMF), University College London, London, United Kingdom; Institut Jean Nicod (MR), Centre National de la Recherche Scientifique, Département d’études cognitives, École Normale Supérieure, PSL Research University; and Laboratoire de neurosciences cognitives et computationnelles (MR), Institut National de la Santé et de la Recherche Médicale, Département d’études cognitives, École Normale Supérieure, PSL Research University, Paris, France; and School of Psychology (CMG), Trinity College Institute of Neuroscience (CMG), and Global Brain Health Institute (CMG), Trinity College Dublin, Dublin, Ireland. TXFS and MR contributed equally to this work. Address correspondence to Tricia X.F. Seow, Ph.D., at t.seow@ucl.ac.uk, or Marion Rouault, Ph.D., at marion.rouault@gmail.com. Received Nov 18, 2020; revised May 14, 2021; accepted May 16, 2021.

REFERENCES

1. Flavell JH (1979): Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. Am Psychol 34:906–911.
2. Hoven M, Lebretón MMM, Engelmann JB, Denys D, Luigjes J, van Holst RJ (2019): Abnormalities of confidence in psychiatry: An overview and future perspectives. Transl Psychiatry 9:268.
3. Sun X, Zhu C, So SHW (2017): Dysfunctional metacognition across psychopathologies: A meta-analytic review. Eur Psychiatry 45:139–153.
4. Engh JA, Friis S, Birkenaes AB, Jønsdóttir H, Klingosoy A, Ringen PA, et al. (2010): Delusions are associated with poor cognitive insight in schizophrenia. Schizophren Bull 36:830–835.
5. Moritz S, Woodward TS, Ruff CC (2003): Source monitoring and memory confidence in schizophrenia. Psychol Med 33:131–139.
6. Moritz S, Woodward TS, Whitman JC, Cuttlar C (2005): Confidence in errors as a possible basis for delusions in schizophrenia. J Nerv Ment Dis 193:9–16.
7. Moritz S, Ramdani N, Klass H, Andreou C, Jungclausden D, Effier S, et al. (2014): Overconfidence in incorrect perceptual judgments in patients with schizophrenia. Schizophr Res Cogn 1:165–170.
8. Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G (2010): Relating introspective accuracy to individual differences in brain structure. Science 329:1541–1543.
9. Morales J, Lau H, Fleming SM (2018): Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. J Neurosci 38:3534–3546.
10. Als J, Zylberberg A, Barttfeld P, Sigman M (2016): Individual consistency in the accuracy and distribution of confidence judgments. Cognition 146:377–386.
11. Rouault M, McWilliams A, Allen MG, Fleming SM (2018): Human metacognition across domains: Insights from individual differences and neuroimaging. Personal Neurosci 1:617.
12. Cesarini D, Johannesson M, Lichtenstein P, Wallace B (2009): HERIABILITY OF OVERCONFIDENCE. J Eur Econ Assoc 7:617–627.
13. Paulus MP, Huys QJMM, Maia TV (2016): A roadmap for the development of applied computational psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 1:386–392.
14. Pouget A, Drugowitsch J, Kepes A (2016): Confidence and certainty: Distinct probabilistic quantities for different goals. Nat Neurosci 19:366–374.
15. Fleming SM, Lau HC (2014): How to measure metacognition. Front Hum Neurosci 8:443.
16. Maniscalco B, Lau H (2012): A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Conscious Cogn 21:422–430.
17. Fleming SM (2017): HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. Neurosci Conscious 2017:nix007.
18. Guggenmos M (2021): Validity and reliability of metacognitive performance measures. PsyArXiv. https://doi.org/10.31234/osf.io/rkzcm.
19. Song C, Kanai R, Fleming SM, Weis RS, Schwarzkopf DS, Rees G (2011): Relating inter-individual differences in metacognitive performance on different perceptual tasks. Conscious Cogn 20:1787–1792.
20. Allen M, Glen JC, Müllensiefen D, Schwarzkopf DS, Fardo F, Frank D, et al. (2017): Metacognitive ability correlates with hippocampal and prefrontal microstructure. Neuroimage 149:415–423.
21. Rahnev D, Fleming SM (2019): How experimental procedures influence estimates of metacognitive ability. Neurosci Conscious 2019:niz009.
22. Moritz S, Jacobsen D, Willenborg B, Jelinek L, Friese S (2006): A check on the memory deficit hypothesis of obsessive–compulsive checking. Eur Arch Psychiatry Clin Neurosci 256:82–86.
23. Moritz S, Ruhe C, Jelinek L, Naber D (2009): No deficits in nonverbal memory, metamemory and internal as well as external source memory in obsessive-compulsive disorder (OCD), Behav Res Ther 47:308–315.
How Local and Global Metacognition Shape Mental Health

24. Moritz S, Klos M, von Eckstaedt FV, Jelinek L (2009): Comparable performance of patients with obsessive-compulsive disorder (OCD) and healthy controls for verbal and nonverbal memory accuracy and confidence: Time to forget the forgetfulness hypothesis of OCD? Psychiatry Res 166:247–253.
25. Tekcan AI, Topcuoglu V, Kayab B (2007): Memory and metamemory for semantic information in obsessive-compulsive disorder. Behav Res Ther 45:2164–2172.
26. McNally RJ, Kohilebeck PA (1993): Reality monitoring in obsessive-compulsive disorder. Behav Res Ther 31:249–253.
27. Cougle JR, Salkovskis PM, Wahl K (2007): Perception of memory ability and confidence in recollections in obsessive-compulsive checking. J Anxiety Disord 21:119–130.
28. Ecker W, Engelkamp J (2001): Memory and memory confidence in obsessive-compulsive disorder. Behav Res Ther 39:911–927.
29. McNally RJ, Kohlbeck PA (1993): Reality monitoring in obsessive-compulsive disorder. Behav Res Ther 31:119–129.
30. Macdonald PA, Antony MM, Maceled CM, Richter MA (1997): Memory and confidence in memory judgements among individuals with obsessive compulsive disorder and non-clinical controls. Behav Res Ther 35:497–505.
31. Tolin DF, Abramowitz JS, Brigidi BD, Amir N, Street GP, Foa EB (2001): Memory and memory confidence in obsessive-compulsive disorder. Behav Res Ther 39:911–927.
32. Roux M, Salio P, Naborscyk L, Pereira M, Roux P, Faivre N (2001): Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. Neurosci Biobehav Rev 26:329–337.
33. Moritz S, Komosa A (2012): Impaired discrimination between imagined and performed actions in schizophrenia. Psychiatry Res 195:1–8.
34. Moritz S, Woodward TS (2003): Memory confidence and false memories in schizophrenia. J Nerv Ment Dis 190:641–643.
35. Moritz S, Woodward TS, Rodriguez-Raecke R (2006): Patients with schizophrenia do not produce more false memories than controls but are more confident in them. Psychol Med 36:659–667.
36. Kwok SC, Xu X, Duan W, Wang X, Tang Y, Allé MC, Berna F (2021): Autobiographical and episodic memory deficits in schizophrenia: A narrative review and proposed agenda for research. Clin Psychol Rev 83:101956.
37. Fleming SM, Ryu J, Goffin JS, Blackmon KE (2014): Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. Brain 137:2811–2822.
38. Shekhar M, Rahnev D (2018): Distinguishing the roles of dorsolateral and anterior PPC in visual metacognition. J Neurosci 38:5078–5087.
39. Hilgenstock R, Weiss T, Witte OW (2014): You’d better think twice: Post-decision perceptual confidence. Neuroimage 99:323–331.
40. Yokoyama O, Rugg MD, Takemoto A, Uchida S, Miura N, Watanabe J, Takemoto A, Uchida S, Hester R, Nestor L, Garavan H (2009): Impaired error awareness and error detection. J Neurosci 39:913–923.
41. Wokke ME, Cleermans A, Riddervold KR (2017): Sure I’m sure: Prefrontal oscillations support metacognitive monitoring of decision making. J Neurosci 37:781–789.
42. Desender K, Van Opstal F, Hughes G, Van den Bussche E (2016): The temporal dynamics of metacognition: Dissociating task-related activity from later metacognitive processes. Neuropsychologia 82:54–64.
43. Yeung N, Summerfield C (2012): Metacognition in human decision-making: Confidence and error monitoring. Philos Trans R Soc B Biol Sci 367:1310–1321.
44. Boldt A, Yeung N (2015): Shared neural markers of decision confidence and error detection. J Neurosci 35:3479–3484.
45. Scheffers MK, Coles MGH (2000): Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. J Exp Psychol Human Percept Perform 26:141–151.
46. Jia W, Zhu H, Ni Y, Su J, Xu R, Jia H, Wan X (2020): Disruptions of frontoparietal control network and default mode network linking the metacognitive deficits with clinical symptoms in schizophrenia. Hum Brain Mapp 41:1445–1458.
47. Alkan E, Davies G, Greenwood K, Evans SLH (2020): Brain structural correlates of metacognition in first-episode psychosis. Schizophr Bull 46:552–561.
48. Hester R, Nestor L, Garavan H (2009): Impaired error awareness and anterior cingulate cortex hypopothesis in chronic cannabis users. Neuropsychopharmacology 34:2450–2458.
49. Moeller SJ, Fleming SM, Gan G, Zilverstand A, Malaker P, d’Oleire Uquillas F, et al. (2016): Metacognitive impairment in active cocaine use disorder is associated with individual differences in brain structure. Eur Neuropsychopharmacol 26:653–662.
50. Haddara N, Rahnev D (2020): The impact of feedback on perceptual decision making and metacognition: Reduction in bias but no change in sensitivity. PsyArXiv. https://doi.org/10.31234/osf.io/p8zyw.
51. Desender K, Boldt A, Yeung N (2018): Subjective confidence predicts information seeking in decision making. Psychol Sci 29:761–778.
52. Donoso M, Collins AGE, Koechlin E (2014): Human cognition: Foundations of human reasoning in the prefrontal cortex. Science 344:1481–1486.
53. Desender K, Donner TH, Verguts T (2021): Dynamic expressions of confidence within an evidence accumulation framework. Cognition 207:105422.
54. Jones RA (1977): Self-Fulfilling Prophecies: Social, Psychological, and Physiological Effects of Expectancies. Mahwah: Lawrence Erlbaum Associates.
55. Madon S, Jussim L, Eccles J (1997): In search of the powerful self-fulfilling prophecy. J Pers Soc Psychol 72:791–809.
68. Bandura A (1977): Self-efficacy: Toward a unifying theory of behavioral change. Psychol Rev 84:191–215.
69. Bandura A (1994): Self-efficacy. In: Ramachaudran VS, editor. Encyclopedia of Human Behavior. New York: Academic Press, 71–81.
70. Wells A (2011): Metacognitive Therapy for Anxiety and Depression. New York: Guilford press.
71. Orth U, Robins RW (2019): Development of self-esteem across the lifespan. In: McAdams DP, Shiner RL, Tackett JL, editors. Handbook of Personality Development. New York: Guilford press, 328–344.
72. Sowislo JF, Orth U (2013): Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. Psychol Bull 139:213–240.
73. Orth U, Robins RW, Meier LL, Conger RD (2016): Refining the vulnerability model of low self-esteem and depression: Disentangling the effects of genuine self-esteem and narcissism. J Pers Soc Psychol 110:133–149.
74. Beck AT (2002): Cognitive models of depression. In: Leahy RL, Dowd ET, editors. Clinical Advances in Cognitive Psychotherapy: Theory and Application. New York: Springer Publishing Company, 29–61.
75. Kom CW, Sharot T, Walter H, Heekeren HR, Dolan RJ (2014): Depression is related to an absence of optimistically biased belief updating about future life events. Psychol Med 44:579–592.
76. Disner SG, Beevers CG, Haigh EAP, Beck AT (2011): Neural mechanisms of the cognitive model of depression. Nat Rev Neurosci 12:467–477.
77. Segal ZV (1988): Appraisal of the self-schema construct in cognitive models of depression. Psychol Bull 103:147–162.
78. Pesiggione M, Le Bouc R, Vincikier F (2018): When decisions talk: Computational phenotyping of motivation disorders. Curr Opin Behav Sci 22:50–58.
79. Pesiggione M, Vincikier F, Bouret S, Daunizeau J, Le Bouc R (2018): Why not try harder? Computational approach to motivation deficits in neuro-psychiatric diseases. Brain 141:629–650.
80. Rouault M, Seow T, Gillan CM (2018): Psychiatric symptom dimensions are associated with dissociable shifts in metacognitive but not task performance. Biol Psychiatry 84:443–451.
81. Rouault M, Seow T, Gillan CM, Fleming SM (2018): Psychiatric symptom dimensions are associated with dissociable shifts in metacognitive but not task performance. Biol Psychiatry 84:443–451.
82. McEvoy JP (2004): The relationship between insight into psychosis and compliance with medications. In: Amador XF, David AS, editors. Insight and Psychosis: Awareness of Illness in Schizophrenia and Related Disorders. Oxford: Oxford University Press, 311–333.
83. Kiani R, Corthell L, Shadlen MN (2014): Choice certainty is informed during cooperation and competition. Neuron 91:482–493.
84. Lee ALF, de Gardelle V, Marmassian P (2021): Global visual confidence [published online ahead of print Mar 25]. Psychon Bull Rev.
85. Kim CW, Corthell L, Shadlen MN (2014): Indicators of medication compliance in Alzheimer’s disease. Int J Alzheimers Dis 20:676–688.
86. Will GJ, Rutledge RB, Moutoussis M, Dolan RJ (2017): Neural and computational processes underlying dynamic changes in self-esteem. Elife 6:e28998.
87. Cosentino S, Metcafe J, Cary MS, De Leon J, Karlawish J (2011): Memory awareness influences everyday decision making capacity about medication management in Alzheimer’s disease. Int J Alzheimers Dis 2011:483897.
88. Risko EF, Gilbert SJ (2016): Cognitive offloading. Trends Cogn Sci 20:576–588.
89. Gilbert SJ, Bird A, Carpenter JM, Fleming SM, Sachdeva C, Tsai PC (2020): Optimal use of reminders: Metacognition, effort, and cognitive offloading. J Exp Psychol Gen 149:501–517.
90. McEvoy JP (2004): The relationship between insight into psychosis and compliance with medications. In: Amador XF, David AS, editors. Insight and Psychosis: Awareness of Illness in Schizophrenia and Related Disorders. Oxford: Oxford University Press, 311–333.
91. Kampman C, Laippala P, Väänänen J, Kiviniemi P, Kilkku N, Lehtinen K (2002): Indicators of medication compliance in first-episode psychoses. Psychiatry Res 110:39–49.
92. Cosentino S, Metcafe J, Butterfeld B, Stern Y (2007): Objective metamemory testing captures awareness of deficit in Alzheimer’s disease. Cortex 43:1004–1019.
93. Chapman S, Colvin LE, Vuorre M, Cocchini G, Metcafe J, Huey ED, Cosentino S (2018): Cross domain self-monitoring in anosognosia for memory loss in Alzheimer’s disease. Cortex 101:221–233.
94. Mazancieux A, Souchay C, Casez O, Moulin CJA (2019): Metacognition and self-awareness in Multiple Sclerosis. Cortex 111:238–255.
95. Mazancieux A, Souchay C, Casez O, Moulin CJA (2019): Metacognition and self-awareness in Multiple Sclerosis. Cortex 111:238–255.
96. Zacharopoulos G, Binetti N, Walsh V, Kanai R (2014): The effect of metacognitive de...
How Local and Global Metacognition Shape Mental Health

115. Gillan CM, Kalanthroff E, Evans M, Weingarden HM, Jacoby RJ, Gerstl-Kovács M, et al. (2020): Comparison of the association between goal-directed planning and self-reported compulsivity vs obsessive-compulsive disorder diagnosis. JAMA Psychiatry 77:77–85.

116. Seow TXF, Benoit E, Dempsey C, Jennings M, Maxwell A, McDonough M, Gillan CM (2020): A dimensional investigation of error-related negativity (ERN) and self-reported psychiatric symptoms. Int J Psychophysiol 158:340–348.

117. Olvet DM, Hajcak G (2008): The error-related negativity (ERN) and metacognition: Toward an endophenotype. Clin Psychol Rev 28:1343–1354.

118. Pasion R, Barbosa F (2019): ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. Neurosci Biobehav Rev 103:133–149.

119. Dalgleish T, Black M, Johnston D, Bevan A (2020): Transdiagnostic approaches to mental health problems: Current status and future directions. J Consult Clin Psychol 88:179–195.

120. Watts AL, Poore HE, Waldman ID (2019): Riskier tasks of the validity of the bifactor model of psychopathology. Clin Psychol Sci 7:1285–1303.

121. Boronofalava MA, Choate AM, Fatimah H, Petersen KJ, Wiernik BM (2020): Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. Biol Psychiatry 88:18–27.

122. Marquand AF, Rezek I, Butelaar J, Beckmann CF (2016): Understanding heterogeneity in clinical cohorts using normative models: Beyond case-control studies. Biol Psychiatry 80:552–561.

123. Gillan CM, Seow TXF (2020): Carving out new transdiagnostic dimensions for research in mental health. Biol Psychiatry Cogn Neurosci Neuroimaging 5:932–934.

124. Sharp PB, Dolan RJ, Eldar E (2020): Disrupted state transition learning as a computational marker of compulsivity and anxious arousal. PsyArXiv. https://doi.org/10.31234/osf.io/x29jq.

125. Evensen J, Røssberg JI, Barder H, Haahr U, Hegelstad Wt, Joa I, et al. (2018): Screening heterogeneous and preliminary evidence for its efficacy. J Behav Ther Exp Psychiatry 41:207–211.

126. Normann N, van Emmerik AAP, Morina N (2014): The efficacy of metacognitive therapy for anxiety and depression: A meta-analytic review. Depress Anxiety 31:402–411.

127. Callesen P, Reeves D, Heal C, Wells A (2020): Metacognitive therapy versus cognitive behaviour therapy in adults with major depression: A parallel single-blind randomised trial. Sci Rep 11:7878.

128. Carpenter J, Sherman MT, Kievit RA, Seth AK, Lau H, Fleming SM (2019): Domain-general enhancements of metacognitive ability through adaptive training. J Exp Psychol Gen 148:51–64.

129. Hall MG, Dux PE (2020): Training attenuates the influence of sensory uncertainty on confidence estimation. Atten Percept Psychophys 82:2630–2640.

130. Moritz S, Burlon M, Woodward TS (2007): Metacognitive training in schizophrenia: A meta-analysis taking into account important moderators. Schizophr Bull 33:962–963.

131. Moritz S, Woodward TS (2005): Metacognitive Training for Patients with Schizophrenia (MCT): Feasibility and preliminary evidence for its efficacy. J Behav Ther Exp Psychiatry 34:131–139.

132. Mielg F, Demiralay C, Sure A, Moritz S, Hottenrott B, Cludius B, Jelinek L (2020): The metacognitive training for obsessive-compulsive disorder: A pilot study [published online ahead of print Nov 26]. Curr Psychol.

133. 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. Schizophr Bull 42:952–962.

134. Carpenter J, Sherman MT, Kaspit RA, Seth AK, Lau H, Fleming SM (2019): Domain-general enhancements of metacognitive ability through adaptive training. J Exp Psychol Gen 148:51–64.

135. Engeler NC, Gilbert SJ (2020): The effect of metacognitive training on confidence and strategic reminder setting. PLoS One 15:e0240858.

136. Winter L, Alam M, Heissler HE, Saryyeva A, Milakara D, Jin X, et al. (2019): Neuropsychological mechanisms of metacognitive therapy – An experimental paradigm. Front Psychol 10:660.

137. Boettcher H, Correa J, Cassiello-Robbins C, Ametaj A, Rosellini AJ, McConkey D, et al. (2020): Comparison of the association between goal-directed planning and self-reported compulsivity vs obsessive-compulsive disorder diagnosis. JAMA Psychiatry 77:77–85.

138. Barlow DH, Farchione TJ, Sauer-Zavala S, Latin HM, Ellard KK, Bullis JR, et al. (2017): Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: Therapist Guide. Oxford: Oxford University Press.

139. Cortese A, Amano K, Koizumi A, Kawato M, Lau H (2016): Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance. Nat Commun 7:13669.

140. Wells A, Fisher P, Myers S, Wheatley J, Patel T, Brewin CR (2012): Metacognitive therapy in treatment-resistant depression: A platform trial. Behav Res Ther 50:367–373.

141. Browning M, Kingston J, Dourish CT, Goodwin GM, Harmer CJ, Dawson GR (2019): Predicting treatment response to antidepressant medication using early changes in emotional processing. Eur Neuropsychopharmacol 29:66–75.

142. Lebreton M, Abitbol R, Daunizeau J, Pessiglione M (2015): Automatic evidence and strategic reminder setting. PLoS One 10:e0139341.

143. Ye Q, Zou F, Lau H, Hu Y, Kwok SC (2018): Causal evidence for mnemonic metacognition in human precuneus. J Neurosci 38:6379–6387.

144. Ye Q, Zou F, Lau H, Hu Y, Kwok SC (2018): Causal evidence for mnemonic metacognition in human precuneus. J Neurosci 38:6379–6387.

145. Taylor SF, Stern ER, Gehring WJ (2007): Neural systems for error monitoring: Recent findings and theoretical perspectives. Neurosci Biobehav Rev 31:160–172.

146. Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004): The role of the medial frontal cortex in cognitive control. Science 306:433–447.

147. Lebrat M, Abitbol R, Dawnizue J, Pesiglione M (2015): Automatic integration of evidence in the brain valuation signal. Nat Neurosci 18:1159–1167.

148. Ullsperger M, Danielmeier C, Johamm G (2014): Neurophysiology of performance monitoring and adaptive behavior. Physiol Rev 94:75–95.

149. Kiani R, Shadlen MN (2009): Representation of confidence associated with a decision by neurons in the parietal cortex. Science 324:759–764.