

Assessing prospective and retrospective metacognitive accuracy following traumatic brain injury remotely across cognitive domains

N. J. Bourke, W. Trender, A. Hampshire, H. Lai, C. Demarchi, M. David, P. Hellyer, D. J. Sharp & D. Friedland

To cite this article: N. J. Bourke, W. Trender, A. Hampshire, H. Lai, C. Demarchi, M. David, P. Hellyer, D. J. Sharp & D. Friedland (2023) Assessing prospective and retrospective metacognitive accuracy following traumatic brain injury remotely across cognitive domains, *Neuropsychological Rehabilitation*, 33:4, 574-591, DOI: [10.1080/09602011.2022.2034650](https://doi.org/10.1080/09602011.2022.2034650)

To link to this article: <https://doi.org/10.1080/09602011.2022.2034650>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 16 Feb 2022.



[Submit your article to this journal](#)



Article views: 1168



[View related articles](#)





[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)



Assessing prospective and retrospective metacognitive accuracy following traumatic brain injury remotely across cognitive domains

N. J. Bourke , W. Trender, A. Hampshire, H. Lai , C. Demarchi, M. David, P. Hellyer, D. J. Sharp and D. Friedland

UK Dementia Research Institute: Care, Research & Technology (UK DRI: CR&T), Division of Brain Sciences, Hammersmith Hospital, Imperial College London, London, UK

ABSTRACT

The ability to monitor one's behaviour is frequently impaired following TBI, impacting on patients' rehabilitation. Inaccuracies in judgement or self-reflection of one's performance provides a useful marker of metacognition. However, metacognition is rarely measured during routine neuropsychology assessments and how it varies across cognitive domains is unclear. A cohort of participants consisting of 111 TBI patients [mean age = 45.32(14.15), female = 29] and 84 controls [mean age = 31.51(12.27), female = 43] was studied. Participants completed cognitive assessments via a bespoke digital platform on their smartphones. Included in the assessment were a prospective evaluation of memory and attention, and retrospective confidence judgements of task performance. Metacognitive accuracy was calculated from the difference between confidence judgement of task performance and actual performance. Prospective judgment of attention and memory was correlated with task performance in these domains for controls but not patients. TBI patients had lower task performance in processing speed, executive functioning and working memory compared to controls, maintaining high confidence, resulting in overestimation of cognitive performance compared to controls. Additional judgments of task performance complement neuropsychological assessments with little additional time-cost. These results have important theoretical and practical implications for evaluation of metacognitive impairment in TBI patients and neurorehabilitation.


ARTICLE HISTORY

Received 8 September 2021
Accepted 21 January 2022

KEYWORDS

Traumatic brain injury; metacognition; remote cognitive assessment; self-awareness; processing speed; memory

CONTACT N. J. Bourke  n.bourke@imperial.ac.uk

 Supplemental data for this article can be accessed <https://doi.org/10.1080/09602011.2022.2034650>.

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
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 use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background

The ability to self-reflect in order to accurately evaluate one's behaviour and thinking is frequently impaired following traumatic brain injury (TBI) (Dockree et al., 2015; Fleming & Lau, 2014; Ham et al., 2014). Clinically, this lack of self-awareness can severely impact the efficacy of rehabilitation interventions (Hart et al., 2005; Prigatano, 2005; Robertson & Schmitter-Edgecombe, 2015), and is often associated with poor functional outcomes (Kelly et al., 2014; Ownsworth & Fleming, 2005; O'Keeffe et al., 2007; Robertson & Schmitter-Edgecombe, 2015).

Metacognition is complex and can be viewed as a sub-component of self-awareness, separable from other facets such as emergent awareness. Within metacognition itself, an early definition from developmental and educational psychology states that these processes can be divided into three subcomponents of metacognitive knowledge, metacognitive regulation and metacognitive experiences (Flavell, 1979). A main focus of cognitive neuroscience has been on metacognitive knowledge; evaluating one's cognitive performance or ability in the form of cognitive judgements (Yeo et al., 2021). Recently, there has been an argument for a "metacognitive g factor", suggesting that metacognition is a domain general function, with a single overarching component "g" (Mazancieux et al., 2020). Contrary to this view, evidence from the existing literature supports that inter-subject variability in metacognition may dissociate according to the cognitive-process domains that are being self-evaluated e.g., online awareness, perception and memory (Fitzgerald et al., 2017). Patterns of neural activity during fMRI tasks add another angle to this debate, providing evidence of both domain general and domain specific metacognitive processes across perceptual and memory domains (Morales et al., 2018). In research, metacognition is frequently investigated with concurrent judgements using signal detection theory (Allen et al., 2017; Siedlecka et al., 2016). In addition, the temporal stage at which judgments are made can influence accuracy (Hacker et al., 2009). That is, whether judgments made are prospective or retrospective in relation to task performance. Retrospective judgments provide a global response following completion of a test, rather than individual items as would be with concurrent judgments, sometimes referred to relative metacognitive judgments (Kelly & Metcalfe, 2011). From the existing work, metacognition is not only likely to have generalized processes but also dissociate across cognitive abilities and can be subcategorized into different processes such as prediction/prospective judgements, online monitoring and retrospective self-reflection/judgment.

In the context of TBI, metacognitive accuracy may vary across domains in the presence of spatial variation of focal lesions, suggesting metacognition may be domain specific (Fleming et al., 2014). Given the heterogenous nature of traumatic brain injuries generalized or specific impairments may arise, highlighting the importance to investigate metacognition across domains. Impairments in

metacognition or lack of insight into disabilities are seen in cases with severe damage or focal injury to specific brain systems involving the prefrontal cortex (PFC) which is a key region for metacognition (Qiu et al., 2018). Conversely, there is a frequently cited discrepancy in emotional distress following TBI, whereby patients with milder injuries report more significant psychiatric problems than those with severe injuries (Bigler, 2001).

Metacognition has been investigated in the context of neuroimaging. The medial and lateral prefrontal cortices are activated during metamemory paradigms (Fleming & Dolan, 2012). Metamemory involves a judgment of accuracy subsequent to answering items on a memory test. Other evidence for levels of self-confidence in performance on memory from a meta-analysis indicates the role of the left dorsolateral prefrontal cortex (dlPFC), posterior medial prefrontal cortex (pmPFC) and bilateral inferior frontal gyrus/insula. Differences were seen when contrasting prospective and retrospective judgments. Retrospective judgments saw clusters in the bilateral parahippocampal cortex and the left inferior frontal gyrus, while prospective judgments had clusters in the posterior pmPFC and left dlPFC. Similar results are seen for metadecision paradigms, with activity in the medial and lateral PFC, the precuneus and the insula (Vaccaro & Fleming, 2018). Lesions in the orbitofrontal cortex (OFC) have previously been linked with impairments in judging one's own performance (Beer et al., 2006). Patients who perform poorly on objective online error monitoring paradigms in which participants have to correct errors made in the task show reduced functional connectivity of the salience network at rest (Ham et al., 2014). These results were seen independent of lesions or white matter damage suggesting metacognition deficits are likely due to a disrupted control network, in which a multitude of tracts or functional connectivity may be impaired (Dockree et al., 2015; Ham et al., 2014).

Metacognition is not routinely assessed in standard clinical neuropsychological assessments. This seems remiss given that metacognition can be significantly impaired following brain injury including traumatic brain injury and this has implications for rehabilitation, coping with everyday life and personal relationships. Discrepancy scores between caregivers and patients of symptom severity from questionnaire data have been a useful method of establishing metacognition in TBI patients (Fleming et al., 1996; O'Keefe et al., 2007). Other work has compared participants' performance with semi-quantified verbal reports of how they believe they did on the task (Hart et al., 1998).

Here we used a semi-quantitative approach to evaluate metacognition through completion of cognitive tests performed remotely via participants' smartphones; this required participants to complete a battery of tests across multiple cognitive domains known to be sensitive to TBI including processing speed, attention, working memory and emotional processing. Following completion of each task participants rated their perceived performance of the task on a sliding scale from 0-100. This reflection provides a global measure of metacognitive awareness across several cognitive domains without the

interference of task design, including assessments where task accuracy is not the primary outcome measure. Metacognitive accuracy is defined here as the difference (Δ) between retrospective confidence judgments (CJ) in performance and actual performance. In addition, patients are asked to give prospective judgements on their memory and attentional function, two commonly reported complaints following TBI. These were collected as part of a questionnaire prior to the completion of cognitive tasks.

We hypothesized that (1) TBI patients would have reduced judgment of attention and memory measured prospectively compared to controls; (2) patients would tend to overestimate their performance based on the difference between performance and retrospective confidence judgments; and (3) patterns of reduced metacognitive accuracy would have a significant association with self-reported levels of wellbeing.

Methods

Participants took part in a remote assessment via their smartphone using a bespoke App (CogniTrack), programmed in CORDOVA and deployable on practically any modern smartphone. Data were automatically synced back to a remote database on a cognitive assessment platform (Cognitron) hosted on the Amazon elastic cloud. The number of participants varied across tasks, depending on the measures they completed as part of a wider longitudinal assessment. To avoid practice effects, only the first timepoint each task was completed was taken forward for the current analysis.

Patients were recruited from a specialist outpatient TBI clinic at St Mary's Hospital, as well as from the major trauma wards at the St. Mary's and St. George's Hospitals between March 2019 to August 2019. Patients were given the option to complete elective cognitive tasks via their smartphone as part of a research programme during a routine clinical appointment. Results from these tasks were available for feedback during follow-up appointments. To aid matching socio-economic demographics as close as possible control participants were recruited by word of mouth from friends, family, or partners of patients. Additional control participants included were recruited by word of mouth through the research team. Exclusion criteria were previous neurological or psychiatric illness. Inclusion criteria were that participants required a smartphone capable of downloading and running the app. All participants gave informed consent to partake in the study through a checkbox on an information screen when they downloaded the app.

A sample of 244 individuals consisting of 142 TBI patients and 99 controls aged between 18–80 enrolled in the study and downloaded the app. Following exclusion of incomplete first assessment and missing demographic data a total of 111 TBI patients [mean age = 45.32(14.15), female = 29] and 84 controls [mean age = 31.51(12.27), female = 43] were analyzed in the current work.

Controls' highest reported completed education was University = 65; High school = 16; Unknown = 4. Patients' highest education level was University = 50; High school = 30, Primary = 3; Unknown = 28. The mean time since injury was 24.77 months, ranging from 0–473 months. Mechanism of injury included: Home/domestic = 2; Road traffic accident (RTA) = 48; Sport = 12; Workplace accident = 8; Unknown = 26; Other = 2. Self-reports of loss of consciousness (LOC): <1 min = 18; 1min–30 min = 24; 1h–24hr = 5; 24hr–7days = 1; more than 7 days = 3; unknown = 53; no LOC = 7. The sample sizes across each of the remote cognitive tasks are as follows: MC (Pat n = 101, Con n = 83); SRT (Pat n = 65, Con n = 64); CRT (Pat n = 103, Con n = 82); TD (Pat n = 99, Con n = 76); PAL; (Pat n = 104, Con n = 72); Trails (Pat n = 103, Con n = 79); EMD (Pat n = 38, Con n = 47); EMC (Pat n = 32, Con n = 51).

Before the cognitive assessment, participants completed a novel short physical and mental health questionnaire (PMHQ), which includes questions on memory, attention, and wellbeing. Self-report measures of wellbeing and cognitive function were recorded using a digital visual analogue scale ranging from 0–100. Participants completed the cognitive tasks covering domains of processing speed, memory, executive functioning, attention, and emotional processing. Prospective ratings of general memory and attention were looked at in relation to subsequent task performance in those domains.

Processing speed

Motor control (MC) requires participants to select targets (n = 30) appearing on the screen in random locations as quickly and accurately as possible. Simple reaction time (SRT) involves tapping on the screen as quickly as possible when a target appears (n = 60). Choice reaction time (CRT) requires selection of one of two arrows as quickly and accurately as possible (n = 60). If a target arrow appears pointing left, the participant needs to select left; if the target arrow points right, the participant needs to select right.

Attention

In the target detection (TD) task, a target stimulus is provided. In a field of changing shapes, the participant must identify and select the target stimulus (n = 120). Each stimulus is presented for the field of stimuli changed every 1000 ms with any given stimuli remaining on screen between 1000 ms and 3000 ms. The total number of correct selections are counted.

Executive functioning

The trail making task is a response-based task modelled from the classic pen and paper assessment of the same name. The switch cost from trail making task

(Trails) involves two levels. The first a set of sixteen numbers are dispersed across the screen. Participants are required to select each number going up in ascending order (e.g., 1-2-3-4 ...). The second level involves numbers and letters. Participants must switch between these also going up in ascending order. (e.g., 1-A-2-B-3 ...).

Working memory

Paired associates learning (PAL) is a working memory task, where participants are required to remember both an image and the location on a grid. Stimulus duration is 2,000 ms for each association. Following each correct trial an additional stimulus is added, increasing the difficulty. The task ends after 3 consecutive incorrect trials. The outcome measure is the maximum number of correct associations.

Emotional processing

Emotional discrimination (EMD) is a response-based task presenting 50 trials of two faces with emotional expressions. The task is balanced for emotional pairing containing stimuli from the Chicago face database, including varying sex and ethnicity (Ma et al., 2015). The aim is to identify if the faces are exhibiting the same or different emotions. The total number of correct responses is calculated. The emotional control task (EMC) is a variation of a Stroop paradigm, using stimuli from the Chicago face database. Two stimuli are presented, a target (small) and distractor (large) over 50 trials. In half the trials the target and distractor are incongruent emotional expressions. Overall accuracy of correct responses to the target stimuli was measured in the current study.

Statistical analysis

Data analysis was run with the statistical package (R Core Team, 2018; <http://www.R-project.org/>).

Following each task, participants were asked to provide a retrospective CJ on their task performance in relation to other people. This was given in the form of a slider ranging from 0-100, with 100 being very confident on task performance. Data from participants was filtered to take only the first time point a task was completed. Overall task performance, confidence of performance and global metacognitive accuracy were explored across a range of cognitive tasks. An index of global metacognitive accuracy was calculated using the difference between CJ and actual task performance in a given task. Task performance was normalized transformed to a percentage of maximum possible scale (POMP; 0-100) to match judgment scores. As true POMP cannot be calculated for reaction time measures with no absolute maximum or minimum, it was calculated relative to

the maximum and minimum performance of the sample population in subsequent analyses. Discrepancy scores were calculated for each task by subtracting subjects' CJ from normalized POMP scores. Scores close to 0 indicate good calibration of judging performance while values closer to +100 indicate overconfidence, and similarly, values closer to -100 indicate underconfidence in performance. Additionally, an estimate of metacognitive performance was calculated across tasks by correlating performance and confidence. This is taken as a global estimate rather than trial-by-trial judgments.

Task performance, confidence judgment and metacognitive accuracy were compared in separate analysis across groups for each cognitive task using a 2-factor ANOVA, where appropriate post-hoc pairwise comparisons were conducted with Tukey's HSD.

Results

Task performance

Overall performance across 8 cognitive tasks between TBI patients and healthy controls was investigated with an ANOVA. TBI patients had poorer performance compared to controls across a range of cognitive domains including processing speed, attention, working memory, executive functioning, but not emotional processing (Figure 1). A significant interaction between cognitive tasks and

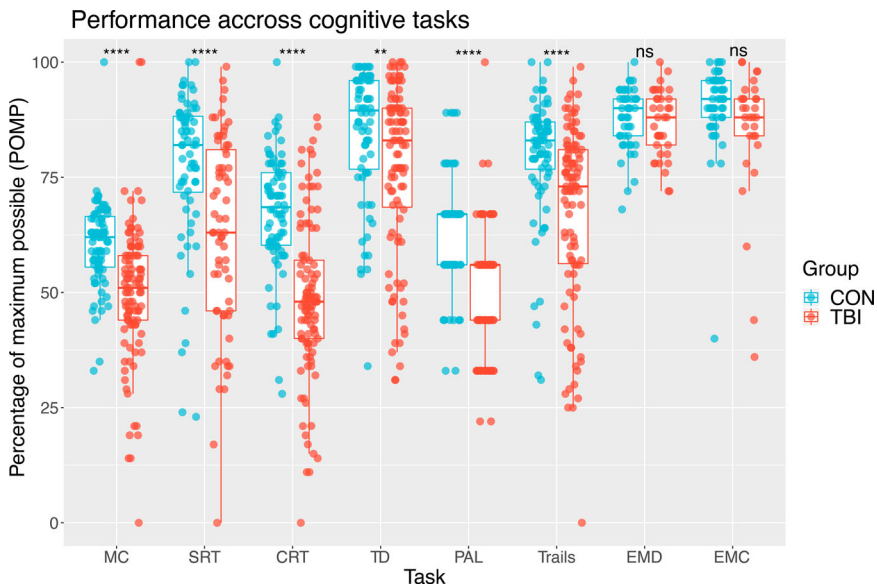


Figure 1. Normalized performance scores across cognitive tasks Higher scores on the POMP scale indicate better performance. MC = motor control; SRT = simple reaction time; CRT = choice reaction time; TD = target detection; PAL = paired associates learning; trails = trail making B-A; EMD = emotional discrimination; EMC = emotional control.

group on performance was seen $F(1,1141) = 3.69, p < 0.001$; Post-hoc pairwise estimates with Tukey correction show that patients performed worse on all cognitive tasks compared to controls apart from emotional discrimination and emotional control. MC $t(1141) = 4.67, p < 0.001$; SRT $t(1141) = 5.32, p < 0.001$; CRT $t(1141) = 8.16, p < 0.001$; TD $t(1141) = 3.32, p < 0.001$; PAL $t(1141) = 5.13, p < 0.001$; Trails $t(1141) = 5.33, p < 0.001$; EMD $t(1141) = 0.41, p = 0.67$; EMC $t(1141) = 1.6, p = 0.1$. Subsidiary analysis including age and education as nuisance regressors produced comparable group differences apart from TD ($t(1141) = 1.69, p = 0.09$).

Retrospective confidence judgment

Following completion of tasks, participants were asked to judge their performance (Figure 2). An interaction between group and cognitive domain fell just below the predetermined significance level with an ANOVA $F(1, 1014) = 2.05, p = 0.056$; A significant difference was seen overall between groups for confidence ($F(1, 1014) = 1.5, p = 0.02$); A main effect of cognitive domain was seen $F(1,1014) = 25.05, p < 0.001$. Post-hoc pairwise estimates of tasks indicate the only cognitive domain with a significant group difference was for working memory (PAL), with patients indicating lower confidence compared to controls $t(1014) = 2.5, p = 0.01$. Subsidiary analysis including age and education as covariates provide comparable results.

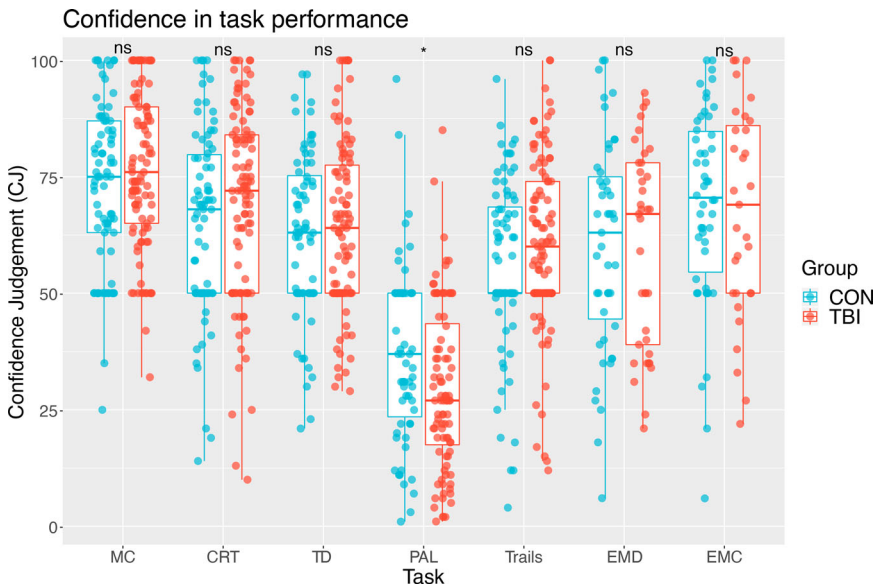


Figure 2. Rating of confidence of task performance. MC = motor control; SRT = simple reaction time; CRT = choice reaction time; TD = target detection; PAL = paired associates learning; trails = trail making B-A; EMD = emotional discrimination; EMC = emotional control.

Metacognitive accuracy (confidence – performance)

To examine metacognitive accuracy, analysis of discrepancy measures (retrospective CJ – performance) were performed, showing a significant difference in pooled metacognitive accuracy across tasks between controls (mean = $-8.41(26.85)$) and patients (mean = $2.65(30.39)$), $t(1086) = -6.37, p < 0.001$.

A two-factor ANOVA was run to examine the relationship between group and cognitive domains with metacognitive accuracy (Figure 3). There was a significant interaction, where cognitive domain moderates group effects of metacognitive accuracy $F(7, 1073) = 3.63, p < 0.001$; post-hoc comparisons of patients and controls across cognitive domains were conducted using ordinary least squares, accounting for means across factors. Results were adjusted for multiple comparisons with Tukey correction. Both controls and patients had a positive bias on motor control, believing they did better than their actual performance, with patients showing significantly higher bias compared to controls $t(1073) = -4.21, p < 0.001$. Patients also overestimate their performance more compared to controls for SRT: $t(1073) = -4.44, p < 0.001$; and CRT $t(1073) = -6.79, p < 0.001$; A negative bias was seen in target detection, PAL, trails, emotional discrimination and emotional control, with patients showing a greater negative bias in target detection $t(1073) = -2.61, p < 0.001$; and trails $t(1073) = -3.63, p < 0.001$ compared to controls. No significant differences were seen between

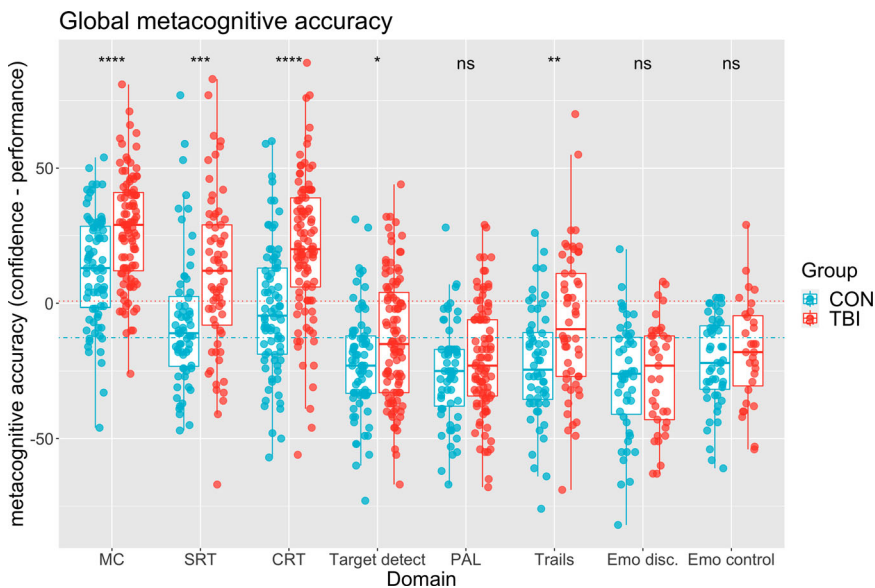


Figure 3. Metacognitive accuracy across cognitive tasks. Discrepancy was calculated using the difference between confidence judgment and task performance. Values closer to 100, indicate overestimation of performance, while values closer to -100 indicate underestimation. Dotted lines represent group pooled metacognitive accuracy. MC = motor control; SRT = simple reaction time; CRT = choice reaction time; TD = target detection; PAL = paired associates learning; trails = trail making B-A; EMD = emotional discrimination; EMC = emotional control.

groups for PAL $t(1073) = -1.3, p = 0.18$; and emotional discrimination $t(1073) = -0.21, p = 0.83$ or emotional control $t(1073) = 0.76, p = 0.44$. As with task performance, subsidiary analysis including age and education level as covariates provide comparable results apart from TD ($t(1073) = -0.862, p = 0.3889$)

Prospective judgement of memory and attention functioning

Further to metacognitive accuracy, investigating the relationship between task performance and retrospective CJs, we looked at the link between task performance and prospective reports of memory and attention functioning (Figure 4). Prior to completing the tasks, participants were asked to give a judgment of their attention and memory.

Working memory

A type III ANOVA with 2 factors, group and task, was used to investigate the relationship of working memory performance and group on prospective judgements of memory. There was a significant interaction for prospective memory judgement modulated by group and performance $F(1, 147) = 4.62, p = 0.03$; the model shows an overall significant effect for performance $F(1, 147) = 7.98, p = 0.005$; no main effect of group was seen $F(1, 147) = 13.55, p = 0.07$; post-hoc analysis of each group's simple slope showed the main effect of performance was seen in controls but not patients. The slope coefficient for controls

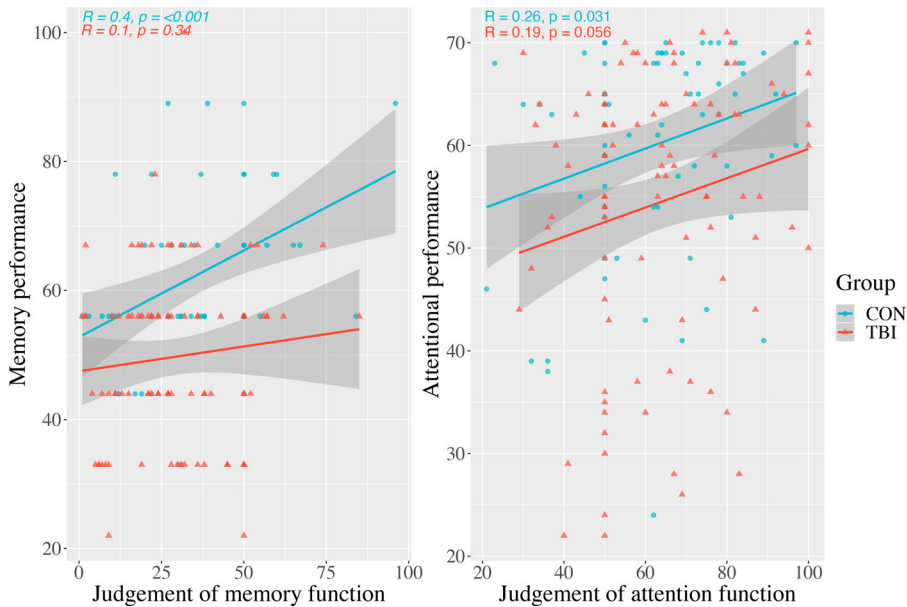


Figure 4. Relationship of general prospective ratings of memory and attention with task performance.

was 6.83, 95% CI [2.91, 10.76] and for patients was 1.37, 95% CI [-1.77, 4.51]; as the simple slope in patients crosses zero, no significant effect of performance prospective memory judgement is seen (Figure 4).

Attention

A type III ANOVA was conducted to investigate the relationship of attentional performance and group to prospective judgements of attention. There was no significant interaction for prospective judgement of attention with task performance and group on $F(1, 160) = 0.63, p = 0.42$; the model shows an overall significant effect for performance $F(1,160) = 6.91, p = 0.009$.

Relationship to reported measures of wellbeing

The relationship of mood and metacognitive accuracy pooled across cognitive tasks was investigated in relation to patient and control groups with an ANOVA (Figure 5). No significant group by mood interaction was seen $F(1, 199) = 0.009, p = 0.92$; however main effects of mood $F(1, 199) = 12.35, p < 0.001$ and group $F(1, 199) = 39.71, p < 0.001$ were seen, indicating participants that rate having a lower mood tend to underestimate their performance to a greater degree, than those rating a higher mood in both patients and controls, while patients tend to overestimate their performance compared to controls overall (Figure 5).

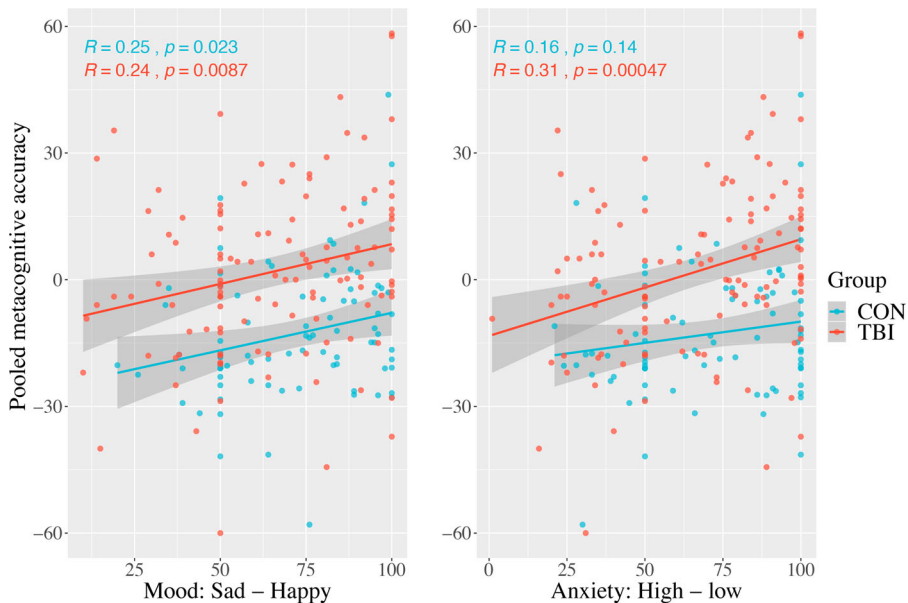


Figure 5. Pooled metacognitive accuracy and measures of mood and anxiety. Lower values on x-axis indicate low mood and higher anxiety, while larger values are associated with better mood and low anxiety.

A similar model was run for anxiety and pooled metacognitive accuracy. No significant group by anxiety interaction was seen $F(1, 199) = 1.72, p = 0.19$; However main effects of anxiety $F(1, 199) = 15.09, p < 0.001$ and group $F(1, 199) = 39.01, p < 0.001$ were seen. Those with higher anxiety tend to underestimate their performance to a greater degree, than those rating a lower anxiety.

Discussion

Here we present a remote assessment of cognitive performance and metacognition across a range of tasks in TBI patients and healthy controls. Metacognitive accuracy was assessed from the difference between task performance and retrospective confidence judgements of performance across a range of cognitive tasks. Prospective memory and attention reports were also evaluated. This measure at the start or end of cognitive assessments has little additional time cost but provides an informative semi-quantitative assessment of global metacognition that can provide novel context to raw cognitive scores.

Controls had good prospective judgement of general memory ability when contrasted with performance on a working memory paradigm, whereas TBI patients did not. Similarly, controls' judgement of general attention was correlated with a subsequent attentional processing task; this effect was not seen in TBI patients. TBI patients showed significant impairment on performance for all cognitive tasks compared to healthy controls apart from emotional discrimination. When controlling for within sample age and education, no group differences were seen for target detection. Age and education are important to consider as they fluctuate throughout the lifespan. Preliminary data from the Great British Intelligence Test (Hampshire, 2020) provide age curves in over 300,000 participants for the target detection task, indicating decreasing accuracy over the age of 50 (Supplementary). Future work will benefit from normative scoring with an independent population sample.

Overall, the range of retrospective confidence judgements did not vary significantly between groups apart from for working memory, which was appropriately lower in patients compared to controls. It is important to note in this battery the working memory task (PAL) only finished when the participant reached their individual ceiling performance. As such potential recency effects of performance on the memory task may bias overall confidence judgment.

Metacognitive accuracy, the discrepancy measure between confidence and performance across the range of cognitive tasks, indicated TBI patients overestimating their performance except in the domains of memory and emotional processing. Simpler reaction time tasks with low cognitive load had a higher bias in confidence compared to other paradigms with great cognitive load such as working memory and attention. The relationship of metacognitive accuracy pooled across all tasks and self-reported levels of mood indicated a positive

correlation with mood in both groups. This may reflect a bias when evaluating performance whereby one tends to underestimate performance with lower mood, and overestimate with higher ratings of mood.

Evidence of meta-analysis across neuroimaging studies indicate separate neural substrates for prospective and retrospective judgments (Vaccaro & Fleming, 2018). Prospective metamemory is associated with the posterior medial prefrontal cortex (PFC), left dorsolateral PFC and right insula, while retrospective metamemory is associated with bilateral parahippocampus and left inferior frontal gyrus activation. These fronto-temporal regions are particularly susceptible to damage from TBI and may explain potential inter-individual variability in deficits of one temporal orientation but not another. It is beyond the scope of this work but examining potential variation in temporal focus across domains and its relation to general and specific processes would be a good target for neuroimaging analysis in TBI cohorts.

This measure of self-reflection has evident relevance for clinical populations. The findings in this study regarding reduced metacognition in the TBI group have implications for clinical neuropsychology. These findings would indicate that metacognition should be assessed in routine clinical neuropsychological assessments involving patients with TBI and could have potential utility in a range of other neurological and psychiatric conditions. Generalizing poor metacognition to broader aspects of behaviour, patients lacking a sense of their limitations may be at increased risk of harm through the manifestation of inappropriate behaviours. Coupled with other commonly reported post-TBI behaviours such as impulsivity, a profile of likely gambling behaviour or risk-taking behaviours may emerge, compounding social, emotional and financial issues, or worse, leading to an elevated risk of physical altercations or assault (Regard et al., 2003; Turner et al., 2019; Williams et al., 2018). With increasing sample sizes, normative scores could be generated for each task paradigm. This could contribute to clinical assessments of impaired insight into disability whereby patients falling below a cut-off score (e.g., $<1SD$) on performance misattribute high confidence to their performance. The current sample did not enable further investigations with reasonable sample sizes of patients who performed poorly ($<1SD$) but with high levels of confidence, or patients with good performance ($>1SD$) with low levels of confidence. This subgrouping could be valuable in future work investigating significant discrepancies between metacognition and performance. An alternative to this frequentist approach is to use machine learning classifiers to investigate if the data can be split meaningfully into subgroups such as identifying high performers with low confidence or low performers with high confidence.

Limitations of remote cognitive assessments need to be considered compared to controlled environments for formal neuropsychology assessments. For instance, while participants are requested to set time aside to complete the tasks in a distraction-free environment, this cannot be enforced, and

some settings may have had more distractions than others. However, we have previously found good compliance among participants and observed minimal performance costs for remote vs in person online assessments. As part of ongoing development of remote cognitive testing, embedded effort tests and flags to determine if participants have left the testing platform, along with self-report effort assessment have been included for future work. Similarly, another consequence of this lack of supervision lies in the inevitability of incomplete data. Third, as in the design of any cognitive battery, an effect of task sequence is unavoidable. In this case, the emotional paradigms were the last in the set of tasks to be completed over multiple sessions. As a result, due to attrition, the number of observations for the emotional processing tasks are lower compared to that available for reaction time tasks placed at the start of the battery. We found an average attrition rate of 25% for controls and 23% for patients across 10 sessions. While this could be addressed by running randomized, or pseudo-randomized ordering of tasks, a consistent standardized order was chosen for task delivery in order to reduce order-related variance when comparing patients to controls. Additionally, there may be non-random effects present in patients who completed all tasks vs those who only completed a subset; a conceivable source of bias relates to low motivation or injury severity. Of note the design of tasks for remote assessments rely on visual perception. Further metacognitive impairments may also arise from auditory or verbal processing paradigms, reflecting challenges in daily life including understanding instructions.

The current work identifies areas of behavioural interest from remote assessments which could be the target of more detailed lab based and neuroimaging experiments. An additional online metacognition task would have added benefits alongside these confidence judgements. This could take the form of a trial-by-trial perceptual judgement, in which signal detection theory, or a paradigm such as the stop change task, could be adapted and applied. Previous work has utilized signal detection theory with modified cognitive paradigms, including retrospective confidence judgements following each trial (Fleming et al., 2010; Fleming & Dolan, 2012; Hauser et al., 2017). For the current work, this approach proved unfeasible, as it would interfere with the paradigm structure of the cognitive tasks. Metacognitive sensitivity, or bias for trial by trial accuracy, is not measured here as in (Rouault et al., 2018). Instead, a single confidence judgement is given following completion of each task paradigm giving global measures of metacognition. Limitations of “single-shot” metacognition assessments have been described previously (Fleming & Lau, 2014). Most notably, the inability to distinguish between concepts such as metacognitive sensitivity and bias. However, the presented method in this work provides useful evaluation across multiple cognitive domains, with minimal time cost or modification of existing cognitive paradigms. Its utility is most notably seen in (Kruger & Dunning, 1999). With

this in mind, a specific paradigm to measure metacognition as part of a testing battery would have added value to the current approach, fractionating components of metacognition.

This global metacognitive accuracy score would be most beneficial when considered alongside other standard questionnaires and detailed clinical information, such as the Glasgow Outcome Scale – Extended (GOSE), Post-traumatic amnesia (PTA) and other measures of functional outcome—or perhaps a customized digital testing paradigm. Initial results here highlight a link between metacognitive accuracy and mood, which could be investigated further. Such metacognitive measures may provide better insights into patient outcomes beyond what is conventionally assessed with cognitive tests. Further investigation of the neural correlates and relationship to functional outcomes is also required.

Conclusion

Metacognitive accuracy is easily measured alongside remote cognitive assessments. Short questions regarding performance provides a global measure of metacognition that is sensitive to deficits in the TBI population. Presented alongside raw cognitive results, this extra dimension supplements our understanding of behavioural and functional outcomes post-TBI.

In addition to supporting this theoretical perspective, there are practical clinical implications. The finding that TBI patients have limited prospective insight into their memory performance supports the importance of administering standardized clinical neuropsychological assessments as routine. These results clearly demonstrate that simply asking a patient how they feel their memory is, whilst important, is unlikely to provide accurate information leading to patients not being offered necessary neurorehabilitation. The present results suggest that measures of metacognition have additive value alongside cognitive measures in a population of TBI. Further work needs to be completed investigating the neural correlates and relationship of metacognition to functional outcomes. The findings in this study indicate that the assessment of metacognition in standard neuropsychological assessments requires further development and investigation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Author N.J.B. is funded by the Imperial presidents PhD scholarship. The development of the online testing platform was supported by NIHR i4i (invention for innovation) Track-Cog-TBI: Computer-based TRACKing and Training COGnition after traumatic brain injury (TBI) Award ID:II-LB-0715-20006.

ORCID

N. J. Bourke  <http://orcid.org/0000-0001-6478-1430>

H. Lai  <http://orcid.org/0000-0002-1600-8440>

References

- 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>
- Beer, J. S., John, O. P., Scabini, D., & Knight, R. T. (2006). Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion-cognition interactions. *Journal of Cognitive Neuroscience*, *18*(6), 871–879. <https://doi.org/10.1162/jocn.2006.18.6.871>
- Bigler, E. D. (2001). The lesion(s) in traumatic brain injury: Implications for clinical neuropsychology. *Archives of Clinical Neuropsychology*, *16*(2), 95–131. [https://doi.org/10.1016/S0887-6177\(00\)00095-0](https://doi.org/10.1016/S0887-6177(00)00095-0)
- Dockree, P. M., O'Connell, R. G., & Robertson, I. H. (2015). Connecting clinical and experimental investigations of awareness in traumatic brain injury. *Handbook of Clinical Neurology*, *128*, 511–524. <https://doi.org/10.1016/B978-0-444-63521-1.00032-7>
- 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), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fleming, J., Strong, J., & Ashton, R. (1996). Self-awareness of deficits in adults with traumatic brain injury: How best to measure? *Brain Injury*, *10*(1), 1–15. <https://doi.org/10.1080/026990596124674>
- 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., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, <https://doi.org/10.3389/fnhum.2014.00443>
- 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>
- Hacker, D. J., Dunlosky, J., & Graesser, A. C. (2009). *Handbook of metacognition in education*. Routledge.
- Ham, T. E., Bonnelle, V., Hellyer, P., Jilka, S., Robertson, I. H., Leech, R., & Sharp, D. J. (2014). The neural basis of impaired self-awareness after traumatic brain injury. *Brain: A Journal of Neurology*, *137*(Pt 2), 586–597. <https://doi.org/10.1093/brain/awt350>
- Hampshire, A. (2020). Great British Intelligence Test protocol. *Research Square*. <https://doi.org/10.21203/rs.3.pex-1085/v1>
- Hart, T., Giovannetti, T., Montgomery, M. W., & Schwartz, M. F. (1998). Awareness of errors in naturalistic action after traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *13*(5), 16–28. <https://doi.org/10.1097/00001199-199810000-00004>

- Hart, T., Whyte, J., Polansky, M., Kersey-Matusiak, G., & Fidler-Sheppard, R. (2005). Community outcomes following traumatic brain injury: Impact of race and preinjury status. *Journal of Head Trauma Rehabilitation*, 20(2), 158–172. <https://doi.org/10.1097/00001199-200503000-00004>
- Hauser, T. U., Allen, M., Purg, N., Moutoussis, M., Rees, G., & Dolan, R. J. (2017). Noradrenaline blockade specifically enhances metacognitive performance. *ELife*, 6, e24901. <https://doi.org/10.7554/eLife.24901>
- Kelly, K. J., & Metcalfe, J. (2011). Metacognition of emotional face recognition. *Emotion*, 11(4), 896–906. <https://doi.org/10.1037/a0023746>
- Kelly, M., McDonald, S., & Kellett, D. (2014). Development of a novel task for investigating decision making in a social context following traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 36(9), 897–913. <https://doi.org/10.1080/13803395.2014.955784>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037//0022-3514.77.6.1121>
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4), 1122–1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Mazancieux, A., Fleming, S. M., Souchay, C., & Moulin, C. J. A. (2020). Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *Journal of Experimental Psychology: General*, <https://doi.org/10.1037/xge0000746>
- Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific Patterns of activity supporting metacognition in Human prefrontal cortex. *Journal of Neuroscience*, 38(14), 3534–3546. <https://doi.org/10.1523/JNEUROSCI.2360-17.2018>
- O'Keefe, F., Dockree, P., Moloney, P., Carton, S., & Robertson, I. H. (2007). Awareness of deficits in traumatic brain injury: A multidimensional approach to assessing metacognitive knowledge and online-awareness. *Journal of the International Neuropsychological Society: JINS*, 13(1), 38–49. <https://doi.org/10.1017/S1355617707070075>
- Owensworth, T., & Fleming, J. (2005). The relative importance of metacognitive skills, emotional status, and executive function in psychosocial adjustment following acquired brain injury. *The Journal of Head Trauma Rehabilitation*, 20(4), 315–332. <https://doi.org/10.1097/00001199-200507000-00004>
- Prigatano, G. P. (2005). Disturbances of self-awareness and rehabilitation of patients With traumatic brain injury: A 20-year perspective. *The Journal of Head Trauma Rehabilitation*, 20(1), 19–29. <https://doi.org/10.1097/00001199-200501000-00004>
- Qiu, L., Su, J., Ni, Y., Bai, Y., Zhang, X., Li, X., & Wan, X. (2018). The neural system of metacognition accompanying decision-making in the prefrontal cortex. *PLOS Biology*, 16(4), e2004037. <https://doi.org/10.1371/journal.pbio.2004037>
- R Core Team (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Regard, M., Knoch, D., Güntling, E., & Landis, T. (2003). Brain damage and addictive behavior: A neuropsychological and electroencephalogram investigation with pathologic gamblers. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, 16(1), 47–53. <https://doi.org/10.1097/00146965-200303000-00006>
- Robertson, K., & Schmitter-Edgecombe, M. (2015). Self-awareness and traumatic brain injury outcome. *Brain Injury*, 29(7–8), 848–858. <https://doi.org/10.3109/02699052.2015.1005135>
- Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: Insights from individual differences and neuroimaging. *Personality Neuroscience*, 1. <https://doi.org/10.1017/pen.2018.16>

- Siedlecka, M., Paulewicz, B., & Wierchoń, M. (2016). But I was so sure! metacognitive judgments Are less accurate given prospectively than retrospectively. *Frontiers in Psychology, 7*. <https://doi.org/10.3389/fpsyg.2016.00218>
- Turner, N. E., McDonald, A. J., Ialomiteanu, A. R., Mann, R. E., McCready, J., Millstone, D., Hamilton, H., Elton-Marshall, T., Rehm, J., Kurdyak, P., Ilie, G., Wickens, C. M., Le, T. L., van der Maas, M., Faregh, N., Cook, S., Bondy, S., Sanchez, S., & Cusimano, M. D. (2019). Moderate to severe gambling problems and traumatic brain injury: A population-based study. *Psychiatry Research, 272*, 692–697. <https://doi.org/10.1016/j.psychres.2018.12.170>
- 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*, <https://doi.org/10.1177/2398212818810591>
- Williams, W. H., Chitsabesan, P., Fazel, S., McMillan, T., Hughes, N., Parsonage, M., & Tonks, J. (2018). Traumatic brain injury: A potential cause of violent crime? *The Lancet. Psychiatry, 5* (10), 836–844. [https://doi.org/10.1016/S2215-0366\(18\)30062-2](https://doi.org/10.1016/S2215-0366(18)30062-2)
- Yeo, Y. X., Pestell, C. F., Bucks, R. S., Allanson, F., & Weinborn, M. (2021). Metacognitive knowledge and functional outcomes in adults with acquired brain injury: A meta-analysis. *Neuropsychological Rehabilitation, 31*(3), 453–478. <https://doi.org/10.1080/09602011.2019.1704421>