
Article:
Fitzgerald, L.M., Arvaneh, M. orcid.org/0000-0002-5124-3497, Carton, S. et al. (3 more authors) (2022) Impaired metacognition and reduced neural signals of decision confidence in adults with traumatic brain injury. Neuropsychology, 36 (8). pp. 776-790. ISSN 0894-4105

https://doi.org/10.1037/neu0000854

Impaired Metacognition and Reduced Neural Signals of Decision Confidence in Adults with Traumatic Brain Injury

Lisa M. Fitzgerald[1,4], Mahnaz Arvaneh[1,2], Simone Carton[3], Fiadhnait O'Keeffe[3], Mark Delargy[3], Paul M. Dockree[1]

[1] Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin 2, Ireland

[2] Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street, Sheffield, UK

[3] Department of Clinical Neuropsychology, National Rehabilitation Hospital, Dun Laoghaire, Dublin, Ireland

[4] School of Psychology, Dublin City University, Dublin 9, Ireland

[5] Psychology Department, St Vincent's University Hospital, Elm Park, Dublin 4, Ireland

Author Note

Lisa M. Fitzgerald https://orcid.org/0000-0003-0551-3377

Paul M. Dockree https://orcid.org/0000-0002-6386-8160

Mahnaz Arvaneh https://orcid.org/0000-0002-5124-3497

We have no known conflict of interest to disclose. Correspondence concerning this article should be addressed to Lisa M. Fitzgerald, School of Psychology, Dublin City University, Glasnevin, Dublin 9, Ireland. Email: lisa.fitzgerald@dcu.ie. The final article will be available, upon publication, via its DOI: 10.1037/neu0000854.

Abstract

Objective: Metacognition reflects our capacity to monitor or evaluate other cognitive states as they unfold during task performance, for example, our level of confidence in the veracity of a memory. Impaired metacognition is seen in patients with traumatic brain injury (TBI) and substantially impacts their ability to manage functional difficulties during recovery. Recent evidence suggests that metacognitive representations reflect domain-specific processes (e.g., memory vs. perception) acting jointly with generic confidence signals mediated by widespread frontoparietal networks. The impact of neurological insult on metacognitive processes across different cognitive domains following TBI remains unknown. Method: To assess metacognitive accuracy, we measured decision confidence across both a perceptual and memory task in patients with TBI (*n*=27) and controls (*n* = 28). During the metacognitive tasks, continuous EEG was recorded, and event-related potentials (ERP) were analyzed. Results: First, we observed a deficit in metacognitive efficiency across both tasks suggesting that patients show a loss of perceptual and memorial evidence available for confidence judgments despite equivalent accuracy levels to controls. Second, a late positive-going ERP waveform (500-700 milliseconds) was greater in amplitude for high vs. low confidence judgements for controls across both task domains. By contrast, in patients with TBI, the same ERP waveform did not vary by confidence level suggesting a deficient or attenuated neural marker of decision confidence post-injury. Conclusions: These findings suggest that diffuse damage to putative frontoparietal regions in patients disrupts domain-general metacognitive accuracy and electrophysiological signals that accumulate evidence of decision confidence.

Keywords: Metacognition; EEG; TBI; late positive potential; decision confidence

Impact Statement

Question: What are the neurocognitive correlates that underlie impaired self-awareness and confidence (metacognition) in patients with traumatic brain injury (TBI). Findings: We report a domain general deficit in metacognitive ability in TBI patients across two cognitive domains (perception and memory) associated with reduced modulation of a centro-parietal late positivity waveform in TBI patients. Importance: These findings provide insight into how disruption of metacognitive processes may underlie inflexible or perseverative behaviors that are common features following brain injury. Next Steps: Future studies could examine these candidate cognitive and electrophysiological markers in the monitoring of targeted treatment options.

Impaired Metacognition and Reduced Neural Signals of Decision Confidence in Adults with Traumatic Brain Injury

Metacognition describes the capacity to self-evaluate or monitor another cognitive process. Efficient metacognition is crucial to shape and adjust daily decision making and optimize the accuracy of goal-directed behavior. Metacognitive deficits have been observed in patients with traumatic brain injury (TBI) across different experimental contexts including impaired retrospective confidence judgments (RCJs) in memory tasks (Chiou et al., 2020; Grossner et al., 2019; Ladowsky-Brooks, 2018), predicted confidence for prospective remembering (Ramanathan et al., 2021; O'Brien & Kennedy, 2018), and impaired error monitoring (Dockree et al., 2015; O'Keeffe et al., 2007; Shen et al., 2020; Sozda et al., 2011).

Ladowsky-Brooks (2018) has advocated for confidence measures to be included as part of neuropsychological assessment of patients with TBI. When factual questions from the information subtest of the Wechsler Adult Intelligence Scale (WAIS-III/IV) were supplemented with RCJs, patients reported higher confidence for incorrect answers, indicating the presence of metacognitive deficits. Wider recognition of metacognitive difficulties at assessment has prompted targeted rehabilitation approaches such as the application of Metacognitive Skills Training (MST) (J. Fleming et al., 2017; Ownsworth et al., 2010), which is designed to improve error detection and self-monitoring capabilities in daily activities and promote functional recovery and psychosocial integration post-injury.

From a neuro-cognitive perspective, it is debated whether metacognitive processes operate independently within specific task domains or represent global resources that are utilized across tasks. However, Morales et al. (2018) provide the first evidence demonstrating that both specific and general processes can co-exist by showing that fMRI-derived multivoxel activity patterns in anterior prefrontal cortex (aPFC) predicted domain-specific metacognition, whereas a broader network in the frontal and posterior midline regions of the brain predicted domain-general

metacognitive efficiency. These findings necessitate a more comprehensive examination of metacognitive deficits in patients with TBI assessing different domains of functioning because the most common type of damage in TBI is diffuse and occurs to the extended white matter, particularly the frontoparietal and frontal thalamic pathways implicated in a range of neuropsychological functions (Bor & Seth, 2012; Koch et al., 2016; Lau & Passingham, 2006; Rees, 2007). It remains to be seen whether patients with TBI show global metacognitive impairments across different domains; for example, perception or memory, or show dissociable domain-specific impairments.

To assess metacognitive performance in the current study, we utilize tasks that draw upon the same set of verbal material across a memory and a perceptual task to ensure that stimulus characteristics are closely matched across domains (Fitzgerald et al., 2017). To quantify metacognitive efficiency, we employed a signal detection measure, *meta-d'-d'* (Maniscalco and Lau, 2012) in order to examine metacognitive ability independent of task performance. To this end, the primary aim of this study is to accurately measure and compare metacognitive processes in patients by distilling any observed deficits of metacognition from variance in objective task performance. We sought to test the opposing hypotheses that patients with TBI would either show domain-general or domain-specific metacognitive deficits. The putative vulnerability of frontoparietal control networks in patients with TBI, and the association of these networks to domain-general metacognitive processes identified by Morales et al. (2018), may increase the probability of reduced domain-general metacognitive efficiency. However, given prior evidence for both domain-general and domain-specific patterns of activation supporting metacognition (Beck et al., 2019; Faivre et al., 2018; Morales et al., 2018; Baird et al., 2013) and evidence of uncorrelated metacognitive accuracy on our memory and perception tasks (Fitzgerald et al., 2017), there remains the possibility that dissociable domain-specific deficits could be present.

The high temporal resolution of event related potentials (ERPs), allows a precise analysis of the time course of brain activity evoked by early first-order decision processes and later second-order post-decisional processes. Later ERP components – in particular, positive parietal ERP components including the P300 (peaking at ca. 300 milliseconds (ms)) and the late positivity (LP; onset ca. 500 ms after stimulus onset) – have been shown to be modulated by post-decision confidence. For example, Boldt and Yeung (2015) have observed a systematic modulation of the P300 amplitude with graded variation in decision confidence – from '*certainly wrong*' to '*certainly correct*' **– o**n a perceptual decision task. The LP amplitude has been shown to increase with higher levels of confidence in recognizing old words (Curran, 2004). Taken together, these findings suggest that the P300 and LP can provide useful electrophysiological markers of decision confidence across different cognitive domains. Therefore, a secondary aim of this study is to investigate modulation of ERP markers (parietal P300 and LP) across domains (perception and memory) as a function of decision confidence in patients and controls. Given that a smaller sample of patients were included for ERP analysis in this study, we conducted an exploratory analysis to investigate whether late ERP markers of decision confidence are altered or attenuated in patients with TBI who show metacognitive deficits.

Finally, an assessment of functioning across attention, language, perception and memory domains and self-awareness in daily life was conducted to better understand the neuropsychological profile of patients with metacognitive deficits. The extent to which metacognitive deficits in this sample were related to more circumscribed neuropsychological deficits or to the broader construct of impaired self-awareness was also examined.

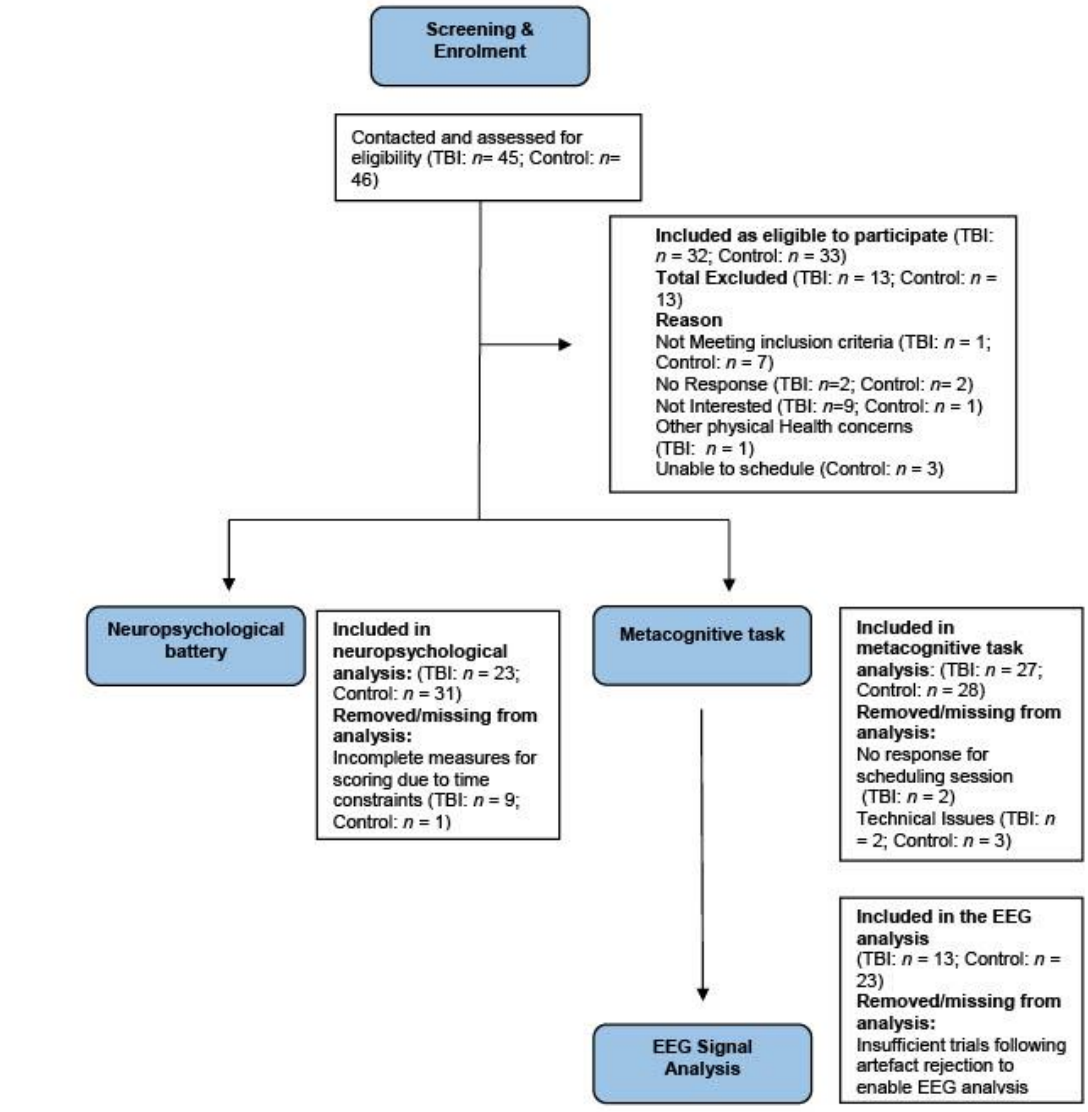
Method

Sample

Thirty-two adults with traumatic brain injury (TBI) took part in this study. Inclusion criteria for all participants included aged 18-65, corrected to normal vision and fluent in English. Exclusion criteria for all participants included a pre-trauma history of epilepsy or other neurological condition, a history of major psychiatric disorder, or a history of drug or alcohol problems. Thirty-three neurologically healthy control participants were also recruited. Recruitment of the control group consisted of a call for research participants on the Volunteer Ireland Platform, the National Men's Shed Association of Ireland, a general public seminar series open to all members of the public ('*21ST Century Brain'*) hosted in Trinity College Institute of Neuroscience Ireland and where participants with a brain injury were university students a matched number of university student participants were recruited. Controls had the additional exclusion criterion that they had no previous neuropsychiatric history (including clinically significant depression or anxiety, learning disability or alcohol or substance abuse), had never suffered a loss of consciousness from a head injury, no previous history of epilepsy or other acquired brain injury (e.g., stroke) (Figure 1). Participants who returned the self-awareness questionnaires (*n*=23 of the patients; *n*=30 of control participants) also had to have a close informant who was a first-degree family member (spouse or partner, adult child, sibling or parent; *n* =42) or close friend (*n*=10) who could verify that they knew the participant either very (*n*=43) or pretty well (*n*=9). Hospital and University Ethics Committees approved the study and all participants gave informed consent prior to participation in accordance with the guidelines of the Declaration of Helsinki.

Figure 1

Flow diagram for recruitment of patients to study



Note. *n* = number of observations; TBI = Persons with a Traumatic Brain Injury; EEG= Electroencephalography

Patients were recruited from the outpatient clinic of a National Rehabilitation Hospital. The mean (*M*) age of the TBI group was 44.06 (Standard deviation (*SD*)= 11.92). This group included 21 males and 11 females. Clinical data is displayed in Table 1. The comparison control group consisted of thirty-three individuals recruited from the general public (21 females; 12 males); the mean age of the control group was 42.69 (*SD*= 15.61). The battery of neuropsychological tests and self-awareness measures (RBANS and CFQ) was unavailable for nine patients in the TBI group and one control participant due to time constraints yielding a sample of twenty-three patients with TBI and thirty control participants for analysis of neuropsychological battery. Two patients were excluded as they did not return for their EEG/metacognitive session and one participant from the TBI group and three controls were excluded from the analysis due to technical difficulties in the task (failed to record responses) yielding a sample of twenty-eight in the TBI group and twenty-eight in the control group for analysis of metacognitive data. Groups did not differ in terms of age *t (63) =*.495*, p*=.62. However, there was a significant difference in premorbid IQ as measured by the NART *t (*50) =-4.85, *p*<.001, Patient *M*=108.10 (*SD*= 8), Control *M*=117.45 (*SD*=5.92) (see Table 2 for breakdown across samples). For EEG analysis, of the sample with data available for the metacognitive task, five control participants and fourteen patients were excluded because they had insufficient confidence trials (i.e., less than 12 high confidence and 12 low confidence trials) following artifact rejection to enable EEG analysis (Larson et al., 2010; Olvet & Hajcak, 2009) yielding a sample of *n*= 13 patients and *n*=23 control participants.

Table 1

Composition of the clinical characteristics of patients within each sample submitted for analysis

Clinical Information:	Full sample % (*n*)	Neuropsychological sample % (*n*)	Metacognitive task sample % (*n*)	EEG sample % (*n*)
Admission GCS				
13-15 Mild	15.6 (5)	13.0 (3)	18.5 (5)	30.8 (4)
9-12 Moderate	21.9 (7)	17.4 (4)	18.5 (5)	15.4 (2)
3-8 Severe	53.1 (17)	60.9 (14)	51.9 (14)	46.2 (6)
Only PTA reported	9.4 (3)	8.7 (2)	11.1 (3)	7.7 (1)
Return to Work				
No	25.0 (8)	21.7 (5)	22.2 (6)	7.7 (1)
Yes	62.5 (20)	69.6 (16)	66.7 (18)	76.9(10)
Prior unemployment	12.5 (4)	8.7 (2)	11.1 (3)	15.4 (2)
Surgical Intervention				
None	65.6 (21)	60.9 (14)	66.7 (18)	76.9 (10)
Craniotomy	18.8 (6)	21.7 (5)	18.5 (5)	7.7 (1)
ICP monitor	15.6 (5)	17.4 (4)	14.8 (4)	15.4 (2)
PTA				
<1 day	9.4 (3)	13.0 (3)	11.1 (3)	15.4 (2)
1-7 days	12.5 (4)	8.7 (2)	11.1 (3)	15.4 (2)
1-4 weeks	21.9 (7)	26.1 (6)	22.2 (6)	23.1 (3)
>4 weeks	25.0 (8)	26.1 (6)	25.9 (7)	23.1 (3)
Only GCS reported	31.3 (10)	26.1 (6)	25.9 (7)	23.1 (3)
Time to Inpatient Discharge				
<1 month	15.5 (5)	17.4 (4)	15.4 (4)	7.7 (1)
<4 months	53.1 (17)	60.9 (14)	51.9 (14)	61.5 (8)
>4 months	18.8 (6)	8.7 (2)	18.5 (5)	15.4 (2)
Outpatient record only	12.5 (4)	13.0 (3)	11.1 (3)	15.4 (2)
Type of Injury				
Fall	34.4 (11)	34.8 (8)	40.7 (11)	23.1 (3)
RTA	53.1 (17)	43.5 (10)	40.7 (11)	53.8 (7)
Assault	15.6 (5)	21.7 (5)	18.5 (5)	23.1 (3)
Time Since Injury				
1-2 years post injury	40.6 (13)	39.1 (9)	37.0 (10)	38.5 (5)
2-3 years post injury	37.5 (12)	39.1 (9)	37.0 (10)	30.8 (4)
3-5 years post injury	15.6 (5)	17.4 (4)	18. 5 (5)	15.4(2)
Greater than 5 years post injury	6.3 (2)	4.3(1)	7.4 (2)	15.4(2)
Medication				
None	53.1 (17)	56.5 (13)	55.6 (15)	38.5 (5)
Depression	18.8 (6)	17.4(4)	18.5 (5)	23.1 (3)
Muscular Pain	12.5 (4)	8.7 (2)	11.1 (3)	23.1 (3)
Other Purposes	15.6 (5)	17.4 (4)	14.8 (4)	15.4 (3)

Note. GCS= Glasgow Coma Scale; ICP monitor= Intracranial Pressure Monitor; PTA= Post Traumatic Amnesia; RTA= Road Traffic Accident

Table 2

Demographic characteristics for each sample submitted for analysis

	Neuropsychological sample (Control *n*=30;TBI *n*=23)		Metacognitive sample (Control *n*=28;TBI *n*=27)		EEG sample (Control n=23;TBI *n*=13)	
	Control	**TBI**	**Control**	**TBI**	**Control**	**TBI**
Age (Years)						
M (SD)	42.58 (16.05)	44.77 (10.87)	42.25 (15.94)	43.63 (12.71)	41.70 (15.98)	38.31 (13.16)
Gender (%)						
Male	32.2	70.8**	35.7	63.0*	30.4	69.2**
Female	67.7	29.2	64.3	37.0	69.6	30.8
NART Premorbid IQ						
M (SD)	117.45 (5.92)	108.64*** (8.21)	117.69 (4.88)	110.33*** (6.37)	118.91 (4.30)	110.0*** (7.42)
Years of Education						
M (SD)	14.52 (3.35)	15.54 (2.54)	14.68 (3.40)	14.78 (2.45)	14.52 (3.70)	14.69 (2.84)

Note. TBI= Participant with a traumatic brain injury; NART= National Adult Reading Test. * denotes significant differences between the groups at p<.05; ** denotes significant differences between the groups at p<.01; *** denotes significant differences between the groups at p <.001. Age, gender and NART premorbid IQ group differences were explored using independent samples t tests and gender differences were explored using chi square test of association.

Metacognitive tasks

Two paradigms were employed to measure metacognitive performance in the domains of decision confidence in memory and perceptual decision-making (see Figure 2). All instructions and stimuli were presented using the 'Presentation' software suite (NBS, San Francisco, CA). All stimuli were presented in a pseudo-random order and word lists were counterbalanced across participants. Stimuli appeared in white font 0.25° over a white fixation cross and on a grey background. Participants sat at a computer at a comfortable viewing distance. They were also instructed to maintain fixation at the fixation cross during task performance in order to minimize eye movements.

Memory Task

Participants were presented with 60 English words in a pseudo randomized sequential order. Words are presented on screen above a white fixation crosshair and participants were asked to memorize as many as possible from the word list presented. Words were generated using the Medical Research Council Psycholinguistic Database (Wilson, 1988). Each word was four to eight letters long, had one to three syllables, and had familiarity, concreteness and imagability ratings between 400 and 700. Following the study phase, participants completed a series of 2 choice old/new judgements. Participants were instructed to indicate if they thought the word that appeared on screen was a word from the previous memory list or a new word. If the word on screen was a word from the memory list, the participant should press the 'right arrow' key. If the word on screen was a new word not on the memory list, the participant should press the 'left arrow' key. To equate TBI and control participants for task difficulty, the two groups were given different presentation durations of the memory list words during the encoding phase. For the patients, words were randomly presented on screen for 2, 3 or 4 second durations. For the control group, words were randomly presented on screen during the learning session for 1, 2 or 3 second durations. Word lists were counterbalanced between participants. In total, each participant completed 180 memory trials (120 new words/60 old words).

Perceptual Task

The perceptual task followed the same protocol developed by Fleming (S.M. Fleming & Dolan, 2010). Following presentation of a verbal stimulus, participants were asked to decide whether a heavily masked stimulus was a word or non-word. Each of the stimuli had the same degree of masking but differed in the presentation duration of the stimulus based on their accuracy staircase. Similar to the memory task, Words were generated using the Medical Research Council Psycholinguistic Database (Wilson, 1988). Each word was four to eight letters long, had one to three syllables, and had familiarity, concreteness and imagability ratings between 400 and 700. A related set of pronounceable non-words was created by a random vowel change. Responses to indicate a Real-word were made using the 'right arrow' key and for a non-word using the 'left arrow' key. The duration of the stimulus presented on screen was titrated such that each participant's performance was maintained at a constant level. The aim of the staircase procedure was to equate the difficulty of the perceptual task between individuals. The staircase adaptively modified levels of difficulty by checking the participants' accuracy over consecutive blocks of 20 trials and adapting the stimulus duration accordingly. The first 20 stimuli of the task were presented for 40ms. The stimulus duration subsequently remained at 40ms as long as accuracy on the previous 20 trials was above 60%-70%. If accuracy exceeded 70% the stimulus duration on screen decreased to 20ms. If accuracy fell below 60% stimulus duration was set to 60ms. This evaluation and task adjustment occurred every 20 trials after. In total, each participant completed 180 trials (120 real words/60 non-words).

On each trial on both the perceptual and memory tasks, participants were presented with a 6-point Likert confidence scale. The scale ranged from 1 (low confidence) to 6 (high confidence) and participants were encouraged to use the whole scale. Responses for the confidence scale were made using the numbers 1 to 6 on the lower right hand side of the keyboard. The confidence scale accepted participants' input for 3 seconds. Participants received no feedback regarding the accuracy of their responses. Task order (perception and memory) and word lists from the MRC database were counterbalanced between participants. Before the main tasks, participants were provided with

practice blocks. Each participant was provided with a standardized set of instructions and practice protocol in 2 separate steps. First, participants were presented with example stimuli and asked to make speeded presses without confidence ratings. This section was designed to familiarize the participant with the task. The second phase consisted of 10 practice trials that simulated the main task such that participants became familiar with indicating their confidence. For the memory task there were 5 practice trials (both responses and confidence ratings) without requiring word list memorization to ensure the participants understood the correct button presses based on their decisions. In the memory task, practice blocks prompted the participants to practice indicating what button they would select for 'old' and 'new' words. It was ensured that all subjects were well practiced and fully understood the requirements of the task prior to performing the main tasks. Each participant completed 180 trials per task (120 new words/ 60 old words on the memory task and 90 real words/90 non words on the perception task).

Figure 2

Schematic Depicting the Perceptual and Memory Retrospective Confidence Tasks

Note. **A**. Post Decisional Confidence tasks. Top panel Perceptual Decision Task. Participants were required to detect whether a masked stimulus was a word or a nonword and then indicate their confidence in that choice. **B.** bottom panel Memory retrieval task. The memory task consisted of a classic verbal recognition memory paradigm. During encoding, participants viewed a word list containing 60 words. During recognition, participants were presented with each word from the full list of 180 stimuli in a random order (60 of which were presented during encoding and half of which were new) and were asked to make discrimination judgements as to whether the stimulus was old or new, and then subsequently rated their confidence in their response

Electroencephalogram Recordings

Continuous EEG data were acquired using the ActiveTwo system (BioSemi, The Netherlands) from 64 scalp electrodes, digitized at 512 Hz. EEG data were collected during all of the tasks blocks. A standard 64-channel system was used. Data were analyzed using custom scripts and EEGLAB functions (Delorme & Makeig, 2004) in MATLAB. EEG data were re-referenced offline to the average reference. The data were high-pass filtered above 0.03 Hz and low-pass filtered below 35 Hz offline using an optimum Butterworth infinite impulse response filter. The 'filtfilth' function in MATLAB was implemented to allow for a non-causal zero-phase filtering approach to eliminate any nonlinear phase distortion associated with using an infinite impulse response filter. Stimulus-locked epochs were then extracted from the continuous data and segmented into epochs of 300ms before to 900 milliseconds

(ms) after stimulus onset and baseline corrected relative to the interval -200 0 msec. Epochs were rejected if the changes in amplitude of channels CP1, CPZ, CP2, P1, Pz and P2, F1, Fz, F2, FC1, FCz, FC2 exceeded an absolute value 100μV during the epoch.

Grand average scalp topography maps were used to identify components of interest. The amplitudes of these components were measured from electrodes centered on scalp regions that presented with the maximum differences between groups. All components were visually inspected separately for the two tasks. In accordance with the spatial topography of the components in the grand average the P300 and LPC was analyzed at electrode Pz. Similarly, the widths of the latency windows used to identify component amplitudes were informed by the duration of each component in the grand average. In the averaged ERP data, the P300 was quantified in an interval from 250 to 350 ms post stimulus and the late positivity in the interval from 500-700ms post stimulus. To achieve adequate signal to noise ratio for both groups' confidence levels were collapsed into low confidence (if participants indicated a confidence of 1, 2, 3, or 4) and high confidence (if participants indicated a confidence of 5 or 6) bins for ERP analysis. These bins were used for both tasks for EEG analyses. Two-way between groups factorial ANOVAs were calculated with domain (perceptual task, memory task) and confidence (high, low) as within subjects factors and group (TBI, control) as a between subject factor. To identify sources of significant main and interaction effects, follow-up ANOVAs were calculated where appropriate. In the text the reported mean values are followed by standard error (i.e., $M \pm SE$).

Background Neuropsychological Screening

Neuropsychological measures administered to participants included: The *Repeatable Battery for the Assessment of Neuropsychological Status* (RBANS; Randolph, 1998) and the *National Adult Reading Test* (NART; Nelson, 1982; Nelson & Willison, 1991). The RBANS comprises 12 subtests measuring attention, language, visuospatial/constructional abilities, and immediate and delayed memory. It was developed for the dual purposes of identifying and characterizing abnormal cognitive

decline in the older adult/brain injured population and as a neuropsychological screening battery for younger populations. In the standardization sample (Randolph, 1998), the RBANS index scores demonstrated robust convergent validity with other neuropsychological measures including the WAIS-III (Wechsler, 1997), WMS-III, Boston naming test (BNT; Kaplan et al., 2001), judgement of line orientation (Benton et al., 1983) and verbal fluency tests. The National Adult Reading Test (NART) (Nelson, 1982; Nelson & Willison, 1991) contains 50 irregular words, which are read aloud and scored for accuracy. It provides an estimation of premorbid IQ based on the WAIS-R and has been validated in both clinical and nonclinical populations (Bright et al., 2002; Crawford et al., 2001; Watt & O'Carroll, 1999).

The *Cognitive Failures Questionnaire* (CFQ; Broadbent et al., 1982) is a 25–item scale that measures self-reported errors of perception, memory and routine action in daily life. It has been employed in a broad range of clinical and nonclinical populations and has high construct validity (e.g., Larson et al., 1997; Wallace & Vodanovich, 2003). Research indicates that the CFQ is strongly correlated with objective indices of attention but not correlated with general intelligence (Manly et al., 1999; Tipper & Baylis, 1987). Higher CFQ scores indicate poorer perceived attentional control affecting different domains (perception, memory and action) in daily life. Ratings are acquired from both the participant and a significant-other to corroborate personal reports from an external frame of reference.

Statistical Analysis

Statistical tests consisted of bivariate correlations (Pearson's *r*) and partial correlations, independent samples t-tests and mixed factor analysis of covariance (ANCOVAs) where appropriate using SPSS software (IBM; Version 27). Greenhouse-Geisser corrected degrees of freedom were used in cases of violated sphericity with corrected p-values reported. We also controlled for the effects of sex and IQ by including these variables as covariates in the analyses of all signal detection measures.

Performance was quantified as the percentage of correct responses in each task. To estimate metacognitive efficiency *meta-d'* was computed as the primary outcome measures compared across

patients with TBI and controls (Maniscalco & Lau, 2012). In a signal detection theory framework *meta-d'* is a measure of type 2 sensitivity (i.e., the degree to which a subject can discriminate correct from incorrect judgments) that is expressed in the same units as type 1 sensitivity (*d'*). A central idea is that primary task performance can influence metacognitive sensitivity, and it is informative to take this influence into account. This approach dissociates a subject- and a domain-specific metacognitive efficiency parameter (*meta-d'-d'*) from both objective task performance and subjective confidence (which both vary on a trial-by-trial basis). This model-based *meta-d-d'* model has been described and validated by Maniscalco and Lau (2012) to account for variance in primary task performance in the computation of type 2 sensitivity. Briefly, this approach exploits the link between type I and type II SDT models to express observed type II sensitivity at the level of the type I SDT model. Maximum likelihood estimation is used to determine the parameter values of the type I SDT model that provide the best fit to the observed type II data. A measure of metacognitive ability that controls for differences in type I sensitivity is then calculated by taking a numerical comparison of *meta d'* and the type I sensitivity parameter d': $M_{Difference}$ = *meta-d'-d'*. For an ideal SDT observer, *meta-d'=d'*; for suboptimal metacognitive sensitivity, *meta-d' < d'* or *meta-d' > d'*. Any instance where $M_{Difference} \neq 0$ 1 implies a deviation of type II sensitivity from expectation that is not attributable to type I performance or type II response bias (provided the standard SDT assumptions hold). Meta-*d'* is theoretically bounded at the lower end by zero, but when fit using an unbounded maximum likelihood estimation procedure, estimation error may lead to negative values. This estimation error applies to all values but becomes evident when values fall outside the theoretical range (S.M. Fleming et al., 2014; Maniscalco & Lau, 2012). Based on previous research that included negative Meta-*d'* values (Moses-Payne et al., 2021) we report analyses including negative values in the current study.

Bayes Factors (*BF*s) were conducted for the two-way group comparisons, simple effects, and for bivariate correlations to determine evidence for the presence or absence of effects. JASP software (JASP Team, 2019) was used to run Bayesian independent samples t-tests and Bayesian correlations. A

default scaling parameter of .707 was used as a prior for the neuropsychological tests and subjective reports analyses, and an informed prior was used for the analysis of Group differences in metacognitive accuracy ($M_{Difference}$). The informed prior was based on a previous study by S.M. Fleming et al. (2014) that reported a 48% decrease in metacognitive accuracy in a brain injured sample. For interpretation of *BF*s, we follow guidelines recommended by Dienes (2014). A *BF* of 0.33 or below represents considerable evidence for the null rather than alternative hypothesis. A value of 3 or above is a cut-off representing substantial evidence for the alternative hypothesis. A *BF* between 3 and 1/3 indicates that data are insensitive for differentiating the alternative from the null hypothesis.

Transparency and Openness

In this manuscript, we report how we determined our sample size, any data exclusion criteria, all manipulations, and all measures included in the study**.** We follow the Journal Article Reporting Standards (JARS; Kazak, 2018). This study's design and it's analysis were not pre-registered.

Results

Do Patients and Controls differ on first-order task accuracy?

Twenty-seven patients with TBI and twenty-eight controls were included in the final metacognitive task analysis. A 2 (TBI group, Control group) x2 (Memory domain, Perception domain) ANCOVA controlling for IQ and Sex (Figure 3) revealed that there was no main effect of Group, (F (1, 51) = .98, p= .33, η^2 = .019), or main effect of domain, (F (1, 51) = .14, p= .71, η^2 = .003) on objective performance (d') indicating no significant differences in performance across group or domain. In addition, no significant interactions of group x domain (F (1, 51) =.0001, p=.989, η^2 = .000004), group x sex, (F (1, 51) =.0003, p=.986, η^2 = .000006), or group x IQ (F (1, 51) =.270, p=.606, η^2 = .005) were observed. There were also no significant effects of Sex (F (1, 51) = .014, p= .91, η^2 = .0003) or IQ, F (1, 51) = 1.32, p= .26, η^2 = .025 as covariates.

Do Patients and Controls Differ on Levels of Metacognitive Efficiency?

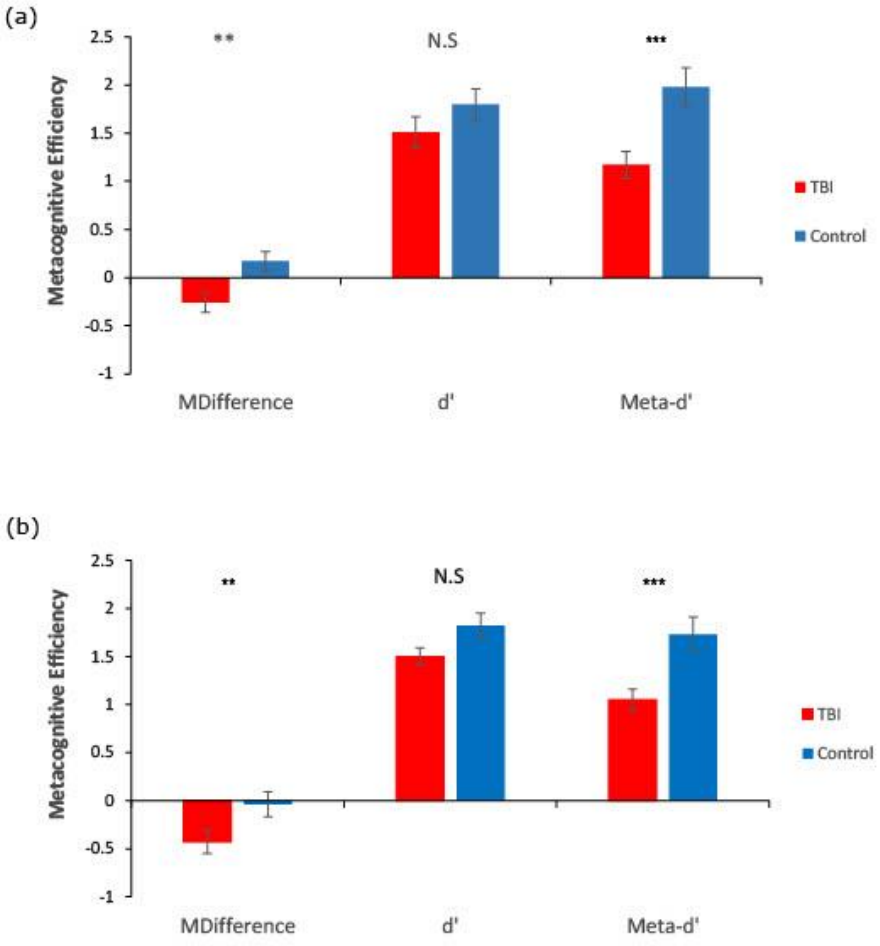
To examine metacognitive differences between the groups, we used the $M_{Difference}$ calculation (Meta d'-d'; Maniscalco & Lau, 2012). In a 2 (TBI group, Control Group) x 2 ($M_{Difference}$ memory, $M_{Difference}$ perception) ANCOVA controlling for IQ and Sex, there was no main effect of domain on metacognitive efficiency; F (1, 51) = .551, p= .46, η^2 = .011) and no significant interaction effects of group x domain, F (1, 51) = .391, p= .53, η^2 = .008, group x IQ, F (1, 51) = 1.65, p= .21, η^2 = .031 or group x sex, F (1, 51) = .400, p= .53, η^2 = .008. However, there was a significant main effect of group (F (1,51) =14.52, p=.0004, η^2 = .22) indicating the patient group had reduced efficiency relative to the control group across both domains. The effect of between-subject effect of IQ was short of significance F (1, 51) = 3.77, p= .058, η^2 = .069. Using an informed prior based on an observed reduction in metacognitive accuracy in a brain injury vs. control sample (S.M. Fleming et al., 2014), a one-tailed Bayesian independent samples t-test was conducted confirming an overall decrease

metacognitive accuracy in patients vs. controls, t (53) = 3.24, p = .001, d = 0.88, BF_{10} = 36.82, with strong evidence for a difference.

In addition, bivariate correlations revealed that the $M_{Difference}$ for perception and memory were not significantly correlated for either the control group (r=.245, p=.209, BF_{10} = .498) or the patient group (r=.199, p=.319, BF_{10} = .383) with Bayesian evidence pointing more toward evidence for the absence of an association. We also examined metacognitive differences between the groups using metacognitive sensitivity (i.e. the degree to which a subject can discriminate correct from incorrect judgements; *Meta-d'*). A similar pattern of results emerged to the estimate of metacognitive sensitivity (*Meta-d'*). In a 2 (TBI group, Control Group) x 2 (*Meta-d'* perception, *Meta-d'* memory) ANCOVA controlling for IQ and Sex, there was no main effect of domain on metacognitive sensitivity ; F(1, 51)=.005, p=.943, η^2 = .0001) and no interaction effect of group x domain, F(1, 51)=.034, p=.855, η^2 = .001, group x sex, F(1, 51)=.275, p=.602, η^2 = .005, or group x IQ, F(1, 51)=.037, p=.848, η^2 = .001 However, there was a significant main effect of group (F(1,51)=11.99, p=.001, η^2 = .19) indicating the patient group had reduced sensitivity relative to the control group across both domains. There were no significant effects of sex, F (1, 51) =.003, p=.955, η^2 = .00006, or IQ, F (1, 51) =.231, p=.633, η^2 = .005, as covariates.

Figure 3

Metacognition task performance in the a) memory and b) perception domain for both patients with TBI and controls



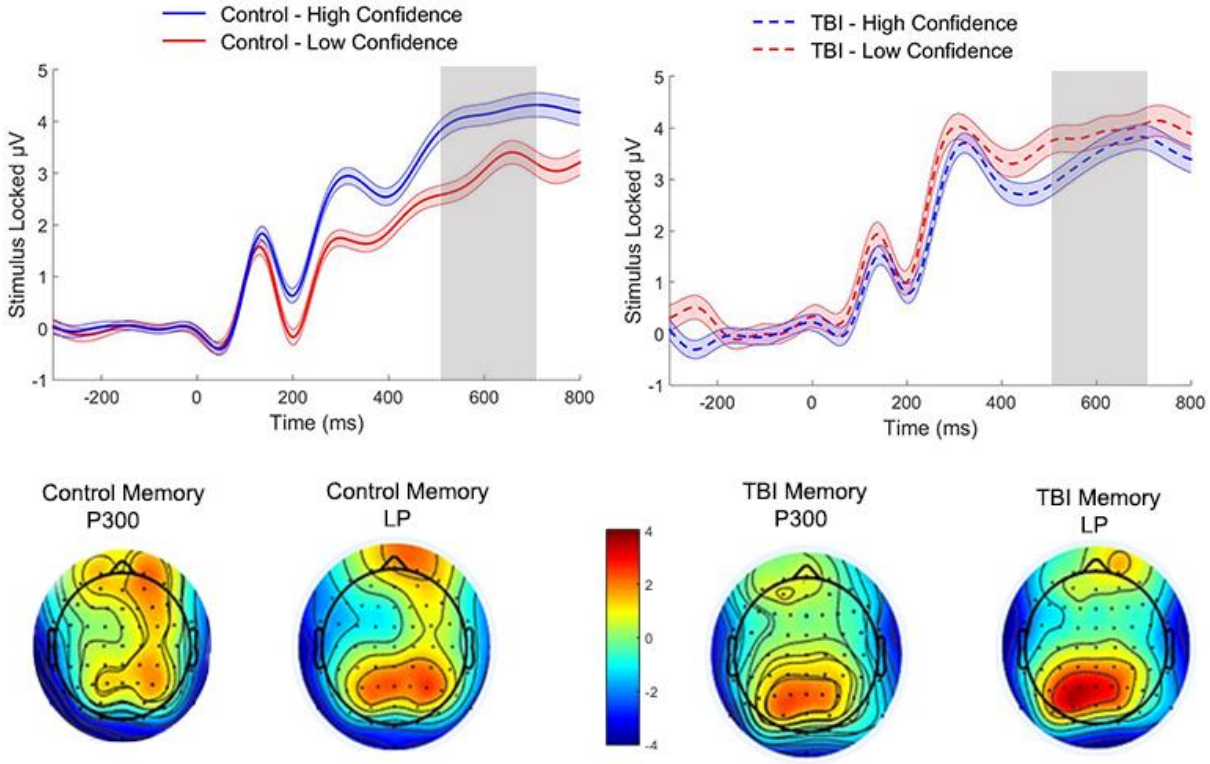
Note. ***= p,.001, **=p<.01, N.S =non-significant. $M_{Difference}$ Values close to zero indicate high metacognitive efficiency. Error bars represent standard error.

Is there a Difference in ERP markers of Decision Confidence Across Groups?

Mixed factorial ANOVAs was conducted for the P300 and LP ERP components with confidence (high, low) and domain (memory, perception) as within subject factors and group as a between subjects factor. For the P300 amplitude, no main effects were found for confidence (F (1,34) =.080, p=.779 η^2 =.002), domain (F (1,34) =.941, p=.339, η2 = .027) or group (F (1,34) =1.197, p=.282, η2 = .034). For the late positivity component, there was a strong main effect of confidence (F (1, 34) =11.24, p=.002) and an interaction effect of group x confidence (F (1, 34) =6.68, p=.014). Simple effects analysis revealed that for the control group, there was greater amplitude for high versus low confidence (F (1, 22) = 22.308, p=. 001, BF_{10} = 264.904). By contrast, TBI participants did not show a significant amplitude modulation by confidence level, and instead showed evidence in support of no difference (F (1, 12) = .279, p=. 607, BF_{10} = 0.314). The grand averaged waveforms and spatial topographies for the P300 and Late Positivity (LP) for the memory and perception task are shown in Figure 4 and Figure 5 respectively.

Figure 4

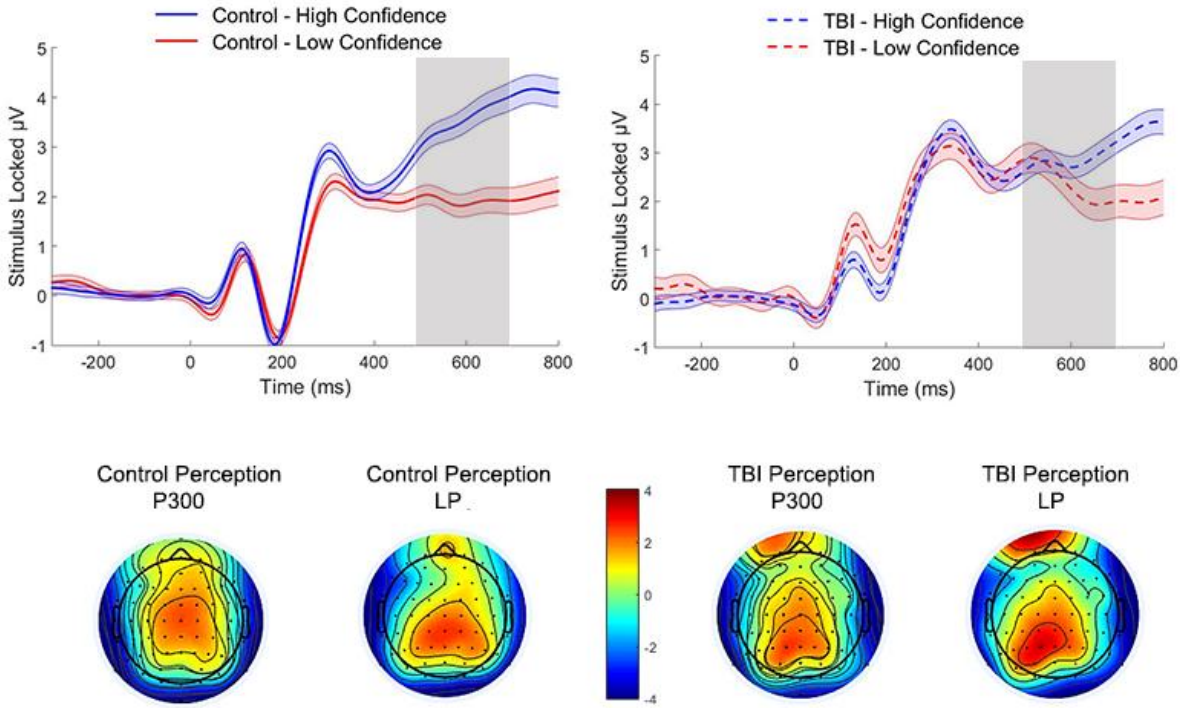
LP but not P300 predicts the accuracy of metacognitive evaluations for controls but not patients on the memory task at electrode Pz



Note. Stimulus-locked ERP for the difference between 'high confidence' and 'low confidence' from 500 to 700 ms on the memory task; the topographic plot indicates voltages as colors from blue (-4µV) to red (+4 µV) for the difference between 'high confidence' and 'low confidence' from 250 to 350ms at P300 and from 500 to 700 ms for LP; Shaded error bars = standard error of the mean (s.e.m).

Figure 5

LP but not P300 predicts the accuracy of metacognitive evaluations for controls but not patients on the perception task at electrode Pz



Note. Stimulus-locked ERP for the difference between 'high confidence' and 'low confidence' from 500 to 700 ms on the perception task; the topographic plot indicates voltages as colors from blue (-4µV) to red (+4 µV) for the difference between 'high confidence' and 'low confidence' from 250 to 350ms at P300 and from 500 to 700 ms for LP; Shaded error bars = standard error of the mean (s.e.m).

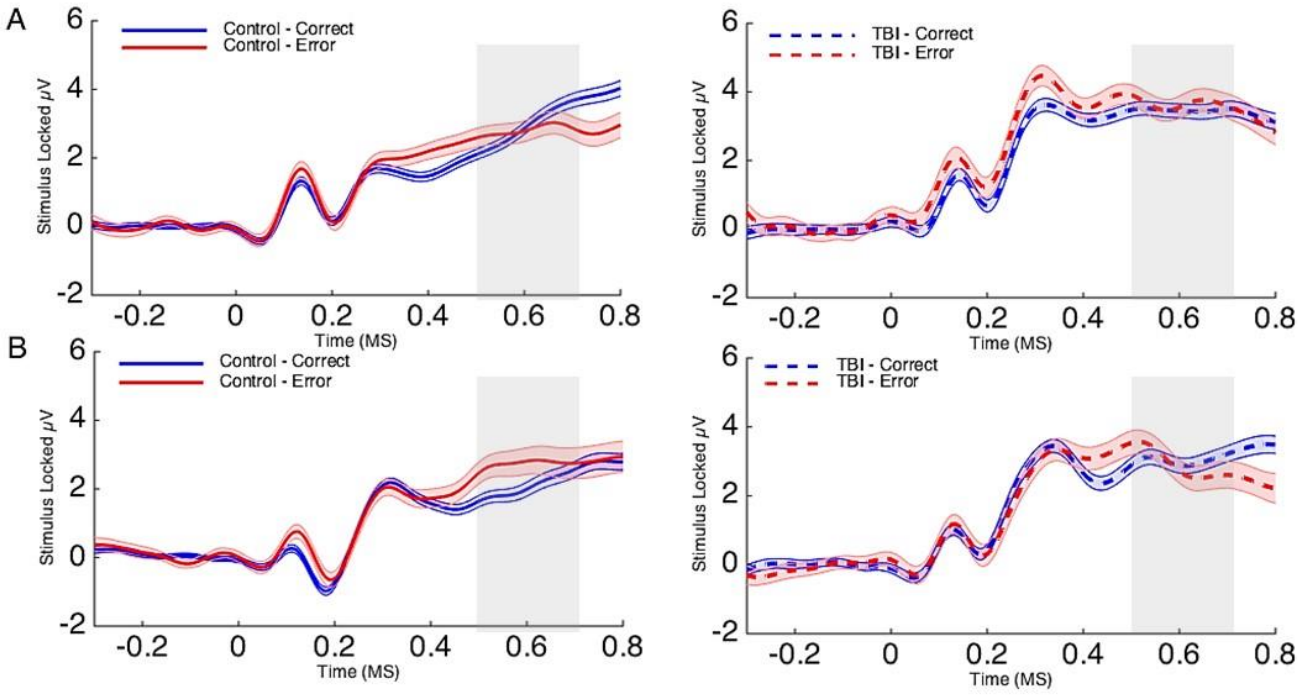
To examine the specificity of LP amplitude difference as a function of confidence level, further mixed factorial ANOVAs were conducted for the P300 and LP with accuracy (correct vs. errors) and domain (memory vs. perception) as within subject factors and group as a between subjects factor (see Figure 6). No significant differences emerged for the P300 for main effects of accuracy (F (1,31) =. 087, p =. 769, η^2 =.003), domain (F (1,31) =.343, p=.563, η^2 = .011) or Group (F (1,31) =3.308, p=.079, η^2 =.096). Additionally, no significant differences were observed for LP modulation in the objective task for accuracy (F (1,31) =.167, p=.769, η^2 =.005), domain (F (1,31) =.921, p=.345, η2 =.029) or group (F

(1,31) =.642, p=.429, η^2 =.020). These findings suggest that the observed decision confidence differences in the LP component do not appear to be explained by amplitude differences underpinned by accuracy of performance.

A one-tailed Bayesian independent samples t-test using the same informed prior (from S.M. Fleming et al., 2014) was repeated to test for metacognitive accuracy differences in the smaller sample used for ERP analysis. An overall decrease in the $M_{Difference}$ measure in the patients vs. controls, was apparent, t (34) = 1.91, p= .032, d = 0.66, BF_{10} = 3.47, with substantial evidence of a group difference and a reduced effect size relative to the full sample.

Figure 6

The P300 and Late Positivity (LP) components as a function of accuracy (correct vs. error) on the memory and perceptual tasks.



Note. (A) Stimulus aligned objective accuracy at PZ for the both TBI and HC group on the memory task. (B) Stimulus aligned objective accuracy at PZ for the both TBI and HC group on the perceptual task. There was no difference or effect of accuracy in either group. Shaded error bars = s.e.m.

How do patients with metacognitive impairment perform on a brief neuropsychological screen?

Twenty-three patients with TBI and thirty control participants were included in the final neuropsychological assessment analysis. A significant difference emerged between the groups on RBANS totals scale score (t (51) =2.25, p=. 029, d = 0.63, BF_{10} = 2.12). However, a separate analysis of each subtest revealed that the group with TBI only showed significant impairment on the attention subtest compared to control participants (t (51) = 3.76, p=. 0004, d = 1.06, BF_{10} = 82.15) but no significant differences were found on all other subtests (all p > .1, all BF_{10} < 1; see Table 3). Follow up bivariate correlations showed that scores on the attention subtest did not significantly correlate with $M_{Difference}$ for either the patients with TBI (Memory r=323, p=.17; Perception r=-.317, p=.17) or the control participants (Memory r=-.049, p=.81; Perception r=-.123, p=.55).

Do patients report more daily life problems or show impaired self-awareness relative to controls?

Patients self-reported more cognitive control failures in daily life than controls, as measured by the CFQ (t (51) = 3.47, p = .001, d = 0.98, BF_{10} = 29.62). Patients' self-reports were externally corroborated by informants of the patients who reported more cognitive failures on the CFQ-for-others compared to informants of the neurologically healthy controls (t (51) = 4.20, p = .0001, d = 1.16, BF_{10} = 205.56). Self- and informant-reports were also significantly correlated for both patients (r = .46, p = .03, BF_{10} = 2.47) and controls (r = .48, p = .007, BF_{10} = 7.29). There was no difference in the mean percentage of reported cognitive failures between informants and patients (t < 1, BF_{10} = 0.30). However, informants of the control group reported significantly fewer cognitive failures than the controls reported themselves (t (58) = 2.17, p=.03, d = 0.57, BF_{10} = 1.84). There was no significant difference between the self-other discrepancy scores for patients vs. controls (t (51) = 1.87, p = .067, BF_{10} = 1.15; Figure 7).

Table 3

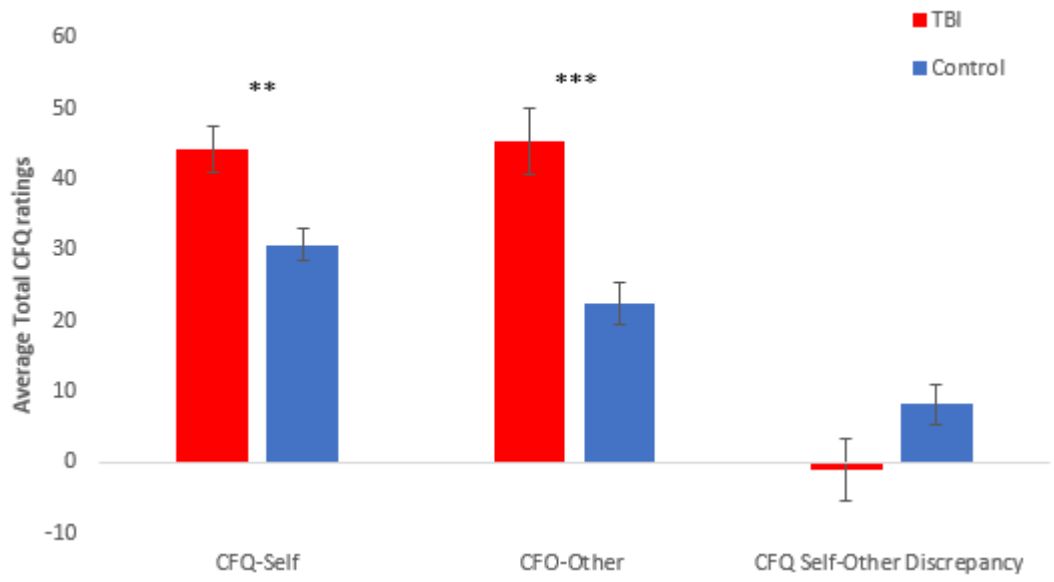
Neuropsychological performance for the mean RBANS total scores and the mean subscale scores in the TBI and control Groups

	Group	N	Mean	SD
RBANS Total Score	TBI	23	88.04	13.43*
	Control	30	97.13	15.42
RBANS Immediate Memory	TBI	23	90.44	13.97
	Control	30	97.07	17.44
RBANS Visuospatial Constructional	TBI	23	91.78	18.89
	Control	30	94.07	15.70
RBANS Language	TBI	23	93.87	12.91
	Control	30	100.00	14.31
RBANS Attention	TBI	23	91.57	18.58***
	Control	30	109.03	15.25
RBANS Delayed Memory	TBI	23	89.30	15.06
	Control	30	92.83	17.15

Note. Patients with TBI differed on the total score measure but subscale analysis found no significant differences except for the attention subscale. * p<.05 *** p<.001

Figure 7

Mean frequency of daily life cognitive failures reported by patients with TBI, controls and their informants.



Note. CFQ Self represents self-reported scores on Cognitive Failures Questionnaire. CFQ-Other represents the scores reported by the informants on the Cognitive Failures Questionnaire -Others. The grey bar represents the discrepancy between self-rated and informant-rated reports. N.B. there is parity in the mean ratings of cognitive failures across patients and their informants. By contrast, the control informant group rate controls, on average, to be marginally less prone to cognitive failures than the control group report themselves. Error bars represent standard error of the mean. ** p<.01 *** p<.001.

Discussion

The present study investigated evidence for metacognitive deficits and altered electrophysiological signatures of confidence judgments in patients with TBI vs. controls. Analysis of metacognitive performance revealed that metacognitive efficiency, defined as *meta-d' – d',* was significantly reduced across both perceptual and memory domains in patients compared to controls. This reduction in efficiency is interpreted as a greater loss of perceptual and memorial evidence available for confidence judgments in the TBI group. Importantly, deficits across both domains could not be accounted for by differences in task difficulty. By controlling for task difficulty using a longer duration study time for patients (in the memory task) and staircase procedure (in the perceptual task), first-order task performance, did not differ between the two groups.

Event related potentials were used to temporally dissociate neural activity underlying second-order decision confidence from first-order task accuracy. We hypothesized that the centro-parietal P300 and late positivity (LP) would show greater amplitude with higher levels of confidence in accordance with previous research (Boldt & Yeung,2015; Frömer et al., 2021). There was no effect of the earlier P300 amplitude but the LP waveform showed a greater mean amplitude for high vs. low confidence judgements in controls but not in patients with TBI. Neither group showed any change in LP amplitude as a function of accuracy indicating that this later amplitude modulation was specific to a metacognitive process. The absence of the LP confidence effect in patients was generalized across the perceptual and memory domains. Several investigations of the centro-parietal positivity in the context of perceptual decisions suggest it is as domain-general or supramodal signal that tracks accumulated evidence for decision choices (Faivre et al., 2018; Kelly & O'Connell, 2015). A more recent study (Herding et al. 2019) makes the distinction between an earlier parietal EEG signal (250-500ms) that reflects evidence accumulation in favor of the upcoming decision choice (first-order response), and a later parietal modulation (500-800ms) that indexed the strength of evidence for that decision and demonstrated all characteristics of decision confidence. In keeping with this distinction, it is possible that attenuation of LP amplitude in patients that we observe here, reflects a weaker

accumulation of evidence specifically contributing to second-order decision confidence, and may contribute to reduced metacognitive efficiency in this group.

The observation that variation in decision confidence was associated with a late sustained positive component supports the hypothesis that this neural signal continues to accumulate evidence after an initial first-order decision facilitating our ability to change our mind or rate our previous decisions as poor. The dynamics of post decisional processing allowing for flexibility of thought have been well documented (Pleskac & Busemeyer, 2010; Resulaj et al., 2009; Yeung & Summerfield, 2012), and disruption to these processes may underlie inflexible or perseverative behaviors that are common features after brain injury. Recent behavioral evidence (Chiou et al., 2020) also demonstrates that retrospective confidence judgments are delayed in patients with TBI highlighting that processing speed is a further variable of clinical relevance. Both reaction time and electrophysiological markers of metacognitive evaluation may prove to be useful in future brain injury research to measure confidence estimates when overt reporting of confidence is not viable or disruptive to the continuity of task performance.

Current computational perspectives (S.M. Fleming and Daw, 2017; Pouget et al., 2016; Rouault, et al., 2018) suggest both domain-general and domain-specific mechanisms may be important for supporting adaptive behavior. Indeed, Morales et al. (2018) observed the co-occurrence of activation patterns with frontoparietal networks underpinning generic confidence signals but more precise context-relevant information (e.g. perceptual or memorial content) necessitating high-order control from anterior prefrontal regions. The unavailability of structural scans from our brain injured sample precluded assessment of potential relationships between neuroanatomical damage to these dissociable networks supporting confidence judgments and the extent of metacognitive impairment seen in the TBI group. However, Hebart and colleagues found that the confidence-related late centro-parietal potential can be source localized to the inferior frontal gyrus (IFG) – a region of the frontoparietal network in which BOLD signal amplitude has also been found to vary with confidence levels (Hebart et al. 2016). Furthermore, multivoxel activity patterns were identified in bilateral IFG

regions and have been associated with across domain confidence estimates (Morales et al., 2018). Together these findings suggest that insensitivity of the LP parietal signal to confidence estimates in the current TBI sample, may reflect disruption to the IFG as a critical region relaying domain-general confidence signals within the frontoparietal network.

Although it is possible that aberrant domain-general frontoparietal signals may be an important cause of metacognitive evaluation difficulties in patients with TBI, it should be noted that metacognitive accuracy across the two task domains was uncorrelated, even though the same set of materials were used in both the perception and memory tasks to eliminate confounds related to different stimulus properties. It is therefore the case that domain independent processes are also needed, possibly through the integration of generic frontoparietal confidence signals with task-specific contextual information in local networks. The extent to which dysfunction of higher-order anterior PFC regions disrupts this integrative function could be a question for functional connectivity analysis in future studies of metacognitive impairment in patients with TBI.

Neuropsychological testing of the patients showed preservation of functioning in language, memory and visuospatial domains but a selective impairment in attention performance compared to controls. This finding is consistent with previous research showing that TBIs have difficulties with attentional control (Dockree et al., 2004; Robertson et al., 1997; Rochat et al., 2013; Zoccolotti et al., 2000) and the vulnerability of frontoparietal networks underlying attentional difficulties in patients with TBI (Hu et al., 2013; Richard et al., 2018). Moreover, there is evidence that impairments of attention may be especially consequential for monitoring of error during a task, and sustained attention capacity is reliably correlated with this critical online metacognitive ability in patients with TBI (O'Keeffe et al., 2007; Dockree et al., 2015). The ability to maintain attention in the context of routine tasks is likely to interact with other forms of metacognitive appraisal like confidence judgments. Although there was no relationship between performance on the attention subtest of the RBANS and metacognitive efficiency in our patients group, we acknowledge that the RBANS may only

capture limited aspects of attentional function. Future inquiry into how levels of vigilance over time and short-term fluctuations of attention disrupt graded judgments of decision confidence in patients with TBI is warranted.

Consistent with the presence of reduced attention performance in patients, higher subjective ratings of daily cognitive failures were reported in the TBI group and corroborated by informant ratings from the patients' significant others. Although metacognitive ability during task performance (e.g. error awareness) has been shown to be related to reduced self-awareness in patients with TBI (Dockree et al., 2015), there was no evidence of impaired self-awareness in the current TBI sample as indicated by the close similarity in self- and informant- ratings of daily cognitive errors. It is possible that preservation of language, perception and memory functioning (as measured by the RBANS) in the current patients may have insulated this group from more pervasive metacognitive knowledge deficits that underlie insight problems. Nevertheless, it is clear that this patient sample with more circumscribed neuropsychological deficits of attention and concomitant reports of cognitive failures in daily life, show impaired metacognitive processes as they unfold during task performance. Metacognitive efficiency measures together with ERP components that track metacognitive processes may serve as tractable markers to assess the effects of pharmacological and/or rehabilitation interventions for clinically significant metacognitive deficits in patients with TBI.

We acknowledge some limitations that prevented investigation of several important issues. Reliance on radiologist reports of neurological damage, allowed us to determine that frontal and/or parietal damage were common in the TBI sample, affecting all but two patients. However, MRI and/or CT scans were not available to provide more precise volumetric analysis of regions of damage. Future work employing a comprehensive approach by combining structural and functional connectivity analysis together with ERP markers of confidence estimates is needed to provide a fuller understanding impaired metacognition in patients with TBI.

Measuring confidence increases the task duration by increasing the length of time required to yield a sufficient number of trials for averaging. Fatigue can be especially difficult for TBI participants

so testing sessions were shortened, reducing the length of EEG recording and limiting the scope of the analyses employed. For example, confidence can lead to increased reaction times that can produce a more accurate focused response strategy making it difficult to measure significant numbers of errors using EEG methods that require large number of trials. ERP analysis of decision confidence in errors vs. correct trials was therefore not possible. EEG however, does circumvent some issues in the measurement of confidence as it provides a robust way to index subjective confidence.

Although this study employed a multidimensional assessment of metacognition investigating model-based metacognitive estimates, ERP markers and daily life reports of cognitive ability in the patient group, there was some missing data from patients, particularly for the ERP analysis (as noted in Table 2). It therefore remains to be seen if the deficits we observed are replicable in a larger sample of patients. The reported deficits in metacognitive efficiency in this patient group are comparable to previous work in anterior frontal lesion patients (S.M. Fleming et al., 2014), with both studies showing similar effect sizes. Nevertheless, greater statistical power in future studies is warranted, especially for ERP markers of confidence, to establish reliable effect size estimates for metacognitive deficits in brain injury populations. Although our study controlled for differences in IQ and sex, future studies should also endeavor to replicate these findings with more balanced demographic profiles. These results offer preliminary evidence for impaired metacognitive efficiency and reduced neural marker of decision confidence in patients with TBI. More generally, it must also be acknowledged that neuroplasticity and compensatory reorganization following neurological insult make it difficult to draw strong conclusions about the typical functional architecture of the metacognitive neural system (Lemaitre et al., 2018).

In conclusion, we investigated metacognitive accuracy for perceptual and memorial confidence judgements in patients with TBI and healthy controls. Altogether, our results highlight a deficit in decision confidence, independent of objective performance, in a sample of patients with TBI. In addition to poor metacognitive efficiency, reduced modulation of a late positive centro-parietal waveform was observed in patients, and we interpret this as attenuation in the strength of

accumulated evidence pertaining to the second order confidence decision. Our findings are consistent with the evidence that domain-general post-decisional processes emerging from generic frontoparietal networks are compromised following brain injury. That metacognitive accuracy is uncorrelated across domains, implies that the evaluation of different sources of evidence requires specialized domain dedicated processors (for perception and memory) to be combined with general confidence signals. Further work is necessary to understand the integration of domain-general and domain-specific processes underlying metacognition and how this is disrupted after brain injury.

Funding

This work was supported by an Irish Research Council, Government of Ireland Postgraduate Scholarship to LF (GOIPG/2013/793). PD was supported by an IRC Laureate grant: IRCLA/2017/306. The authors have no competing interests to declare.

References

Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *Journal of Neuroscience*, *33*(42), 16657-16665.

Beck, B., Peña-Vivas, V., Fleming, S., & Haggard, P. (2019). Metacognition across sensory modalities: Vision, warmth, and nociceptive pain. *Cognition*, *186*, 32-41.

Benton, A. L. (1983). *Judgment of line orientation*. Oxford University Press.

Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience*, *35*(8), 3478-3484.

Bor, D., & Seth, A. K. (2012). Consciousness and the prefrontal parietal network: insights from attention, working memory, and chunking. *Frontiers in psychology*, *3*.

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British journal of clinical psychology*, *21*(1), 1-16.

Bright, P., Jaldow, E. L. I., & Kopelman, M. D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: a comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society*, *8*(6), 847-854.

Chiou, K. S., Genova, H. M., Lengenfelder, J., & Chiaravalloti, N. D. (2020). Postdecisional Processing Delays in Metacognitive Monitoring After Traumatic Brain Injury. *J Head Trauma Rehabil, 35*(2), 152-159.

Crawford, J. R., Deary, I. J., Starr, J., & Whalley, L. J. (2001). The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval. *Psychological medicine*, *31*(3), 451-458.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, *42*(8), 1088-1106.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology*, *5*, 781.

Dockree, P. M., Tarleton, Y. M., Carton, S., & FitzGerald, M. C. (2015). Connecting self-awareness and error-awareness in patients with traumatic brain injury. *Journal of the International Neuropsychological Society*, *21*(7), 473-482.

Dockree, P. M., Kelly, S. P., Roche, R. A., Hogan, M. J., Reilly, R. B., & Robertson, I. H. (2004). Behavioural and physiological impairments of sustained attention after traumatic brain injury. *Cognitive Brain Research*, *20*(3), 403-414.

Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *Journal of Neuroscience*, *38*(2), 263-277.

Fitzgerald, L. M., Arvaneh, M., & Dockree, P. M. (2017). Domain-specific and domain-general processes underlying metacognitive judgments. *Conscious Cogn, 49*, 264-277.

Fleming, J., Ownsworth, T., Doig, E., Hutton, L., Griffin, J., Kendall, M., & Shum, D. H. (2017). The efficacy of prospective memory rehabilitation plus metacognitive skills training for adults with traumatic brain injury: study protocol for a randomized controlled trial. *Trials*, *18*(1), 1-11.

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological review*, *124*(1), 91.

Fleming, S. M., & Dolan, R. J. (2010). Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Consciousness and cognition*, *19*(1), 352-363.

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.

Frömer, R., Nassar, M. R., Bruckner, R., Stürmer, B., Sommer, W., & Yeung, N. (2021). Response-based outcome predictions and confidence regulate feedback processing and learning. *Elife*, *10*, e62825.

Grossner, E. C., Bernier, R. A., Brenner, E. K., Chiou, K. S., Hong, J., & Hillary, F. G. (2019). Enhanced default mode connectivity predicts metacognitive accuracy in traumatic brain injury. *Neuropsychology, 33*(7), 922-933.

Hebart, M. N., Schriever, Y., Donner, T. H., & Haynes, J. D. (2016). The relationship between perceptual decision variables and confidence in the human brain. *Cerebral cortex*, *26*(1), 118-130.

Herding, J., Ludwig, S., von Lautz, A., Spitzer, B., & Blankenburg, F. (2019). Centro-parietal EEG potentials index subjective evidence and confidence during perceptual decision making. *NeuroImage*, *201*, 116011.

Hu, P., Fan, J., Xu, P., Zhou, S., Zhang, L., Tian, Y., & Wang, K. (2013). Attention network impairments in patients with focal frontal or parietal lesions. *Neuroscience Letters*, *534*, 177-181.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*. Pro-ed.

Kazak, A. E. (2018). Journal article reporting standards.

Kelly, S. P., & O'Connell, R. G. (2015). The neural processes underlying perceptual decision making in humans: recent progress and future directions. *Journal of Physiology-Paris*, *109*(1-3), 27-37.

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307-321.

Ladowsky-Brooks, R. L. (2018). Evaluating semantic metamemory: Retrospective confidence judgements on the information subtest. *Appl Neuropsychol Adult, 25*(2), 137-148.

Larson, G. E., Alderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and correlates of the Cognitive Failures Questionnaire. *British Journal of Psychology*, *88*(1), 29-38.

Larson, M. J., Baldwin, S. A., Good, D. A., & Fair, J. E. (2010). Temporal stability of the error-related negativity (ERN) and post-error positivity (Pe): The role of number of trials. *Psychophysiology*, *47*(6), 1167-1171.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, *103*(49), 18763-18768.

Lemaitre, A. L., Herbet, G., Duffau, H., & Lafargue, G. (2018). Preserved metacognitive ability despite unilateral or bilateral anterior prefrontal resection. *Brain and cognition*, *120*, 48-57.

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and cognition*, *21*(1), 422-430.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661-670.

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.

Moses-Payne, M. E., Habicht, J., Bowler, A., Steinbeis, N., & Hauser, T. U. (2021). I know better! Emerging metacognition allows adolescents to ignore false advice. *Developmental Science*, *24*(5), e13101.

Nelson, H. E. (1982). *National adult reading test*. Windsor, UK: NFER-Nelson.

Nelson, H. E., & Willison, J. R. (1991). *The revised national adult reading test-test manual.* Windsor, UK: NFER-Nelson.

O'Brien, K. H., & Kennedy, M. R. (2018). Predicting Remembering: Judgments of Prospective Memory After Traumatic Brain Injury. *Journal of Speech, Language, and Hearing Research*, *61*(6), 1393-1408.

Olvet, D. M., & Hajcak, G. (2009). The stability of error-related brain activity with increasing trials. *Psychophysiology*, *46*(5), 957-961.

O'Keeffe, 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*, *13*(1), 38-49.

Ownsworth, T., Quinn, H., Fleming, J., Kendall, M., & Shum, D. (2010). Error self-regulation following traumatic brain injury: a single case study evaluation of metacognitive skills training and behavioural practice interventions. *Neuropsychological Rehabilitation*, *20*(1), 59-80.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological review*, *117*(3), 864.

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nature neuroscience*, *19*(3), 366-374.

Ramanathan, P., Liu, R., Chen, M. H., & Kennedy, M. R. (2021). Memory and Executive Functions Subserving Judgments of Learning: Cognitive Reorganization After Traumatic Brain Injury. *Neuropsychological Rehabilitation*, 1-24.

Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *Journal of clinical and experimental neuropsychology*, *20*(3), 310-319.

Rees, G. (2007). Neural correlates of the contents of visual awareness in humans. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*(1481), 877-886.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, *461*(7261), 263-266.

Richard, N. M., O'Connor, C., Dey, A., Robertson, I. H., & Levine, B. (2018). External modulation of the sustained attention network in traumatic brain injury. *Neuropsychology*, *32*(5), 541.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747-758.

Rochat, L., Beni, C., Annoni, J. M., Vuadens, P., & Van der Linden, M. (2013). How inhibition relates to impulsivity after moderate to severe traumatic brain injury. *Journal of the International Neuropsychological Society*, *19*(8), 890-898.

Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: insights from individual differences and neuroimaging. *Personality neuroscience*, *1*.

Shen, I. H., Lin, Y. J., Chen, C. L., & Liao, C. C. (2020). Neural correlates of response inhibition and error processing in individuals with mild traumatic brain injury: an event-related potential study. *Journal of neurotrauma*, *37*(1), 115-124.

Sozda, C. N., Larson, M. J., Kaufman, D. A., Schmalfuss, I. M., & Perlstein, W. M. (2011). Error-related processing following severe traumatic brain injury: an event-related functional magnetic resonance imaging (fMRI) study. *International Journal of Psychophysiology*, *82*(1), 97-106.

Tipper, S. P., & Baylis, G. C. (1987). Individual differences in selective attention: The relation of priming and interference to cognitive failure. *Personality and individual differences*, *8*(5), 667-675.

Wallace, J. C., & Vodanovich, S. J. (2003). Workplace safety performance: Conscientiousness, cognitive failure, and their interaction. *Journal of Occupational Health Psychology*, *8*(4), 316.

Watt, K. J., & O'Carroll, R. E. (1999). Evaluating methods for estimating premorbid intellectual ability in closed head injury. *Journal of Neurology, Neurosurgery & Psychiatry*, *66*(4), 474-479.

Wechsler, D. (1997). *Wechsler memory scale (WMS-III)* (Vol. 14). San Antonio, TX: Psychological corporation.

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior research methods, instruments, & computers*, *20*(1), 6-10.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Phil. Trans. R. Soc. B*, *367*(1594), 1310-1321.

Zoccolotti, P., Matano, A., Deloche, G., Cantagallo, A., Passadori, A., Leclercq, M., ... & Zimmermann, P. (2000). Patterns of attentional impairment following closed head injury: A collaborative European study. *Cortex*, *36*(1), 93-107.