ABSTRACT

Objective: Patients who suffer from memory loss following an Acquired Brain Injury (ABI) may also suffer from anosognosia, or unawareness of their memory loss. How we define and measure anosognosia can have critical implications for its study and clinical assessment. Commonly used measures often lack standardization and reliability checks for responses. Moreover, these methods rely heavily on cognitive abilities (e.g., language abilities) that are often affected after brain injury. The aim of this study is to elucidate how to best conceptualize and detect anosognosia for memory loss by introducing a new method of assessment, the Visual-Analogue Test for Anosognosia for memory impairment (VATAmem). Methods: A total of 51 patients (M= 61 years; M= 13 years of education) with memory difficulties following ABI were recruited from outpatient clinics. A total of 73 informants were also recruited (M= 51 years old; M= 13 years of education). Both patients and informants evaluated the severity of patients’ everyday memory mistakes on the VATAmem, for prospective and retrospective memory deficits by using visual analogue scales, vignettes and check questions to ensure reliability. Results and conclusion: A total of 30% of the patients were deemed unaware of their memory deficits. Patients were less aware of their prospective (29%) than their retrospective memory difficulties (18%). The new method of assessment provided by the VATAmem reduced possible false positives and enhanced reliability. We conclude that careful consideration of methodology is a key step to interpreting anosognosia findings within a theoretical framework

Key words: Anosognosia, memory, assessment, acquired brain injury, unawareness

Public significance statement: This study has implications for the advancement of the field of unawareness of memory deficits (anosognosia), informing both scientific and clinical questions about how to conceptualize and measure the complex phenomenon of anosognosia of memory impairment.
1. INTRODUCTION

Anosognosia refers to unawareness of, or lack of insight into, one’s own deficits that can include motor, sensory, functional, behavioral, and/or cognitive disturbances such as memory impairment (Mograbi & Morris, 2018). Anosognosia for memory impairment can manifest following different Acquired Brain Injuries (ABI) (e.g., stroke, traumatic brain injury) or neurodegenerative diseases such as Alzheimer’s disease (AD) (Prigatano & Schacter, 1991; Agnew & Morris, 1998). The prevalence of anosognosia for memory impairment is primarily derived from studies of individuals diagnosed with AD. Such studies show a highly variable prevalence of anosognosia, with reported frequencies ranging from 25% to 80% (Starkstein & Tranel, 2012). With regard to patients with ABI, studies of patients with traumatic injuries have shown that between 30% and 40% of those with moderate to severe injuries will have some degree of unawareness of their behavioural and neuropsychological deficits, including memory loss (Fischer, Gauggel, & Trexler, 2004; O’Keeffe, Dockree, Moloney, Carton, & Robertson, 2007; Prigatano, 1996; Prigatano, Altman, & O’Brien, 1990). Studies focusing on post-stroke patients have reported that between 39% and 72% of patients present various degrees of unawareness (Anderson & Tranel, 1989). When memory impairment alone was examined, the reported prevalence of unawareness was reduced to 27% (Hartman-Maeir, Soroker, Ring, & Katz, 2002).

The assessment of anosognosia for memory impairment is traditionally based on explicit measures of awareness that assess the patient’s own evaluations of their overall memory deficits (e.g., Clare, Marcová, Verhey & Kenny, 2005). These measures typically include clinical interviews and/or structured questionnaires. In clinical interviews, the accuracy of patients’ self-evaluation
is judged by a clinician. In structured questionnaires, discrepancy scores represent the difference between a patient’s self-evaluation versus a “gold standard” (e.g., informant reports or memory performance on standardized assessments) (Clare et al., 2005; Cosentino & Stern, 2005). Although clinical interviews can provide qualitatively interesting information regarding individuals’ self-reflections, they are rarely standardized and can be prone to inter-rater bias (Clare et al., 2005). Questionnaires can provide a more systematic assessment of awareness, although they can be prone to pitfalls that can limit their validity and reliability (Cocchini, Beschin & Della Sala, 2012). For example, patients’ cognitive impairments (e.g., comprehension deficits) may affect the integrity of their self-evaluations. Moreover, some existing measures lack normative data, complicating proper interpretation of responses (see Clare et al., 2005; Cosentino & Stern, 2005 for reviews). Relatively few studies have attempted to address some of these limitations (Clare et al., 2005; Clare, Wilson, Carter, Roth, & Hodges, 2002; Cocchini et al., 2012; Cocchini, Gregg, Beschin, Dean, & Della Sala, 2010; Della Sala, Cocchini, Beschin, & Cameron, 2009; Crawford, Henry, Ward, & Blake, 2006). The aim of this study is to introduce a new measure of anosognosia for memory impairment that builds upon existing measures and attempts to overcome their common limitations.

As part of the primary goal of the current study, we report on the development of the Visual-Analogue Test for Anosognosia for memory impairment (VATAmem), a measure with reduced demands on language and memory abilities. It also comprises response reliability check questions, and items dedicated to examine information about two fundamental types of memory deficits, retrospective versus prospective. The specific questions and subscales included in the VATAmem are largely derived from an existing awareness measure, the Prospective and
Retrospective Memory Questionnaire (PRMQ) (Smith, Della Sala, Logie, & Maylor, 2000). To enhance the patient’s understanding and minimize overload of memory with relatively long instructions, each question was depicted in visual vignettes illustrating the question. To facilitate the patient’s responses, we adopted a visual analogue scale similar to two previous measures assessing anosognosia for motor deficits (VATAm) (Della Sala et al., 2009) and for language deficits (VATA-L) (Cocchini et al., 2010). The VATAmem examines explicit awareness of possible memory difficulties in a variety of everyday situations (e.g., Do you have difficulty remembering appointments?), and compares patients’ judgments with a “gold standard” (i.e., reports provided by an informant). The discrepancy between the patient’s and informant’s evaluation provides a continuous outcome measure that represents both degree and directionality of awareness (Clare et al., 2005).

A second goal of the current study is to examine potential variability in awareness across specific types of everyday memory failures, namely, prospective versus retrospective memory. Prospective memories can be defined as those pertaining to the future (e.g., remembering to carry out an action), whilst retrospective memories can be defined as those linked to the past (e.g., remembering past actions or events) (Einstein, McDaniel, Marsh, & West, 2008). Although prospective and retrospective memories are likely to share similar underlying memory networks or structures (Einstein et al., 2008; Schacter, Addis, & Buckner, 2007; Underwood, Guynn, & Cohen, 2015), prospective memories differ from retrospective memories in their inherent self-initiated processes that form the intentions to remember something in the future (Craik, 1986). Further, the properties of prospective memories vary from those of retrospective memory, with regard to the types of associations with other memories and aspects of the environment that are
required to prompt the individual to remember in the future (Marsh, Cook, & Hicks, 2006). Interestingly, individuals appear to experience these memory failures differently. Indeed, previous studies examining subjective cognitive complaints with the PRMQ have found that individuals report prospective difficulties more frequently than retrospective (Crawford et al., 2006; Crawford, Smith, Maylor, Della Sala, & Logie, 2003). Such studies, however, are based on reports by healthy older adults and may not translate to amnesic patients with variable degrees of awareness. This study will examine both retrospective and prospective awareness in a sample of patients with memory loss following ABI, adopting a new assessment measure that aims to complement existing tests and overcomes, at least in part, some of the methodological pitfalls of available tools.

2. METHODS

2.1. Participants

A total of 190 individuals with ABI were initially referred to the study screening phase by consultant neurologists or neuropsychologists from three sites, the Columbia University Medical Center Department of Neurology Stroke outpatient clinic in the U.S., the NHS St. George’s Hospital Stroke outpatient clinic in the U.K., and the Neuropsychology unit of Somma Lombardo Hospital in Italy. Of the initial group, 60 patients were considered for the study as they presented with no evidence of psychiatric illness but with evidence of memory difficulties as determined by age-corrected standardized scores of immediate and delayed story recall (Humphreys, Bickerton, Samson, & Riddoch, 2012). A further 9 patients dropped out of the study due to lack of interest or failure to follow up, leaving a final sample size of 51 patients. The final sample of 51 patients
(39% females) had a mean age of 61.40 years (SD = 14.90; range = 22-87) and 13.16 years of education (SD = 3.75; range = 4-22). Mean time since lesion onset was 2.89 months (SD = 4.85; range = 0.07 – 22). The majority had stroke (64.7% ischemic; 11.8% hemorrhagic), 17.6% had traumatic brain injury and the remaining (5.9%) had other injuries (i.e., 1 from obstructive hydrocephalus and 2 from hypoxia). Regarding localization of injury, 11 patients (22%) had right sided brain injury, 22 (43%) had left sided brain injury, and 18 (35%) had bilateral or diffuse brain injury.

For each participant, an informant was recruited to provide evaluations of the patient’s memory ability. For a subset of participants (n = 22), two informants were recruited to enable us to calculate the VATAmem cut-off score as described below in the statistical analysis and results section. This resulted in a total of 73 informants with a mean age of 50.85 (SD = 19.44; range = 18-92) years, and 13.30 years of education (SD = 3.34; range = 8 – 23). All informants were people who frequently interacted with the patient on a regular daily or weekly basis.

Fifteen patients and their informants were retested after 48 to 72 hours to examine test re-test reliability. All participants provided full consent, and procedures were approved by the Institutional Review Board at Columbia University Medical Center in the US, NHS research ethical board in the UK, and the ethical board of the Somma Lombardo Hospital in Italy.

2.2. Neuropsychological assessment

All patients underwent an initial general cognitive assessment and specific cognitive tests, described below, to evaluate long-and short-term memory, language, attention, and executive
functions. Due to different scoring systems across tests, patients’ performance in each measure was converted to standardized z-scores, which were then collapsed to represent each cognitive domain (i.e., memory, language, and executive function).

### 2.2.1. General cognitive examination

Patients were assessed with the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). The score ranges from 0–30. Higher scores represent higher cognitive functioning, and a score below 24 has been used as an indicator of general cognitive difficulties (Folstein et al., 1975; Kukull et al., 1994).

### 2.2.2. Memory

Patients were asked to complete the Rivermead Behavioural Memory Test (RBMT; Wilson, Cockburn, & Baddeley, 2003; Wilson et al., 2008). The RBMT provides an ecological measure of memory for everyday activities. It includes measures such as remembering an appointment or a route around the room. Other measures include picture and face recognition, and story recall. The RBMT-2 was used in samples recruited from the UK and US, while the RBMT-3 was used in samples recruited from the Italian site. For the RBMT-2, the total profile raw score ranges from 0 to 24. Scores between 17–21 indicate poor memory, 10–16 indicate moderate impairments, and those between 0–9 indicate severe memory problems. For the RBMT-3, the overall General Memory Index (GMI) score was considered. This ranges from 0 to 150 with a mean of 100 and a standard deviation of 15. Based on the RBMT-2 and RBMT-3 classifications,
the severity scores were determined as mild impairment (i.e., between -1 SD and -1.5 SD), moderate impairment (i.e., between -1.6 and -2 SD), and severe impairment (i.e., below -2 SD).

Patients’ short-term memory was assessed with Digit and Spatial span tests (Orsini et al., 1987; Randolph, 2012). The digit forward raw scores ranged from 0 to 16 for the English version, and from 0 to 10 for the Italian version. Raw scores of the spatial span ranged from 0 to 10 (Corsi, 1972). Visuospatial span was assessed through the Visual Pattern Test in the Italian sample (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999) where patients have to reproduce a visuospatial matrix. Raw scores range from 2 to 15.

For each patient we calculated final standardized z-scores for verbal (i.e. Digit span), and visuospatial (i.e., Spatial and visuo-spatial spans) short-term memory to determine impairment.

2.2.3. Language

Language was assessed using naming subtests from two measures. English speaking patients were assessed with the naming subtest of the BCoS Battery (Brain Behavior Analysis – Humphreys et al., 2012). Raw scores range from 0 to 14, with higher scores reflecting better naming abilities. Italian speaking patients’ naming abilities were measured with the subtest of the “Esame Neuropsicologico per l’Afasia” (Capasso & Miceli, 2008). Raw scores range from 0-10, with higher scores also representing better naming abilities. Normative data from each subtest was used to determine impaired language abilities. A within sample z-score was calculated to explore the association between the severity of language difficulties and awareness.
2.2.4. Executive functions

Measures of rule following and set switching were included for the measurement of executive function. These included: (i) the executive subtest in the BCoS Battery (Humphreys et al., 2012), consisting of a visuospatial sequencing task which displays changes in rules and the patient must detect them; and (ii) the Trail Making Test (TMT) (Reitan & Wolfson, 1985), which is composed of two trail making subtests. In Part A, the patient is asked to draw lines linking numbers in ascending order from 1 to 25 in the shortest time possible. In Part B, the patient is asked to repeat the same procedure, but alternating between letters and numbers (i.e., 1–A; 2–B; 3–C). The total raw scores of the executive BCoS Battery test (ranging from 0 to 18), and Part B of the TMT (e.g., time to complete in seconds) were combined to calculate an overall z-score of executive function. Severity scores were determined as mild impairment (i.e., between -1 SD and -1.5 SD), moderate impairment (i.e., between -1.6 and -2 SD), and severe impairment (i.e., below -2 SD).

2.2.5. Awareness of memory deficits

Self and informants’ reports of memory performance were obtained through two measures of awareness of everyday memory functioning in a counterbalanced order.

*Prospective and Retrospective Memory Questionnaire (PRMQ)*

The PRMQ (Crawford et al., 2006; Smith et al., 2000) includes a total of 16 items requiring patients and informants to rate the frequency of the patient’s everyday memory mistakes, from 5 (Very often) to 1 (Never). Raw scores are converted to true scores, as reported by Crawford et
al. (2003; 2006), and can range from 16 to 80 with lower scores representing more difficulties. Discrepancies between patients and informants can thus range from -64 to +64, with a 0 representing perfect agreement. Positive values in discrepancy scores represent an informant rating a patient as having more difficulties than he or she is endorsing, and negative values reflect reports of more difficulty by the patient. The PRMQ provides two cut-off scores to interpret the difference between an individual and their informant ratings that can be used for the assessment of anosognosia (i.e., cut-off of 7 with a significance value of $p = .05$ for the full scale and a cut-off of 9 for prospective and retrospective subscales; Crawford et al., 2006).

The Visual Analogue Test for Anosognosia for Memory impairment (VATAmem)

Preliminary version of the scale

The final version of the VATAmem was derived following a preliminary phase consisting of a series of pilot studies. These allowed selection and refinement of questions and vignettes based on feedback from a total of 9 patients with memory disorders (age $M = 53; SD = 20.50$; range 25-78; 89% male) and 40 healthy adults (age $M = 51.60; SD = 17.37$; range: 25-85; 52% male) in both languages. Items assessed were derived from the English and Italian versions of the PRMQ. Based on the outcome of these pilot studies, fifteen questions were selected for the final version as being unambiguous, culturally relevant, and relatively easy to illustrate in vignettes.
Final Scale

The VATAmem consists of 15 questions assessing everyday memory situations, 1 practice item to ensure the participant’s compliance with the test (i.e., Do you have difficulty watching TV?), and 4 check questions to control reliability of participants’ responses, as described below (see Figure 1 for an example). As in the PRMQ, the 15 memory-related items explore two different dimensions of memory: prospective and retrospective memory. Prospective memory questions refer to those activities in which an individual needs to remember an intention for a future action (e.g., remembering to call someone later as they did not answer the phone). Retrospective items examined memory for activities in which an individual needs to recall past learned information (e.g., remembering that they have already told a person a story). All items were balanced across items referring to self versus environmentally cued activities; that is, those in which an individual relies on internal cues to remember information (e.g., remembering appointments without the help of a calendar), versus remembering information when cued by something in the environment (e.g., remembering to take an umbrella with you when it’s raining). All items were also balanced across short versus long-term memory, that is, memory for information that was just learned versus information that had been learned before. Thus, following similar classification as in the PRMQ, each question represents one aspect of each of the three dimensions. For example, “Do you have problems remembering that you have already told the same story to the same person on a previous occasion?” would represent retrospective, long-term memory, and environmentally cued dimensions.
Patients were asked to rate their current ability in each task depicted by the vignettes by saying the number or pointing to a specific point on the 4–point scale (Stern et al., 1997; Della Sala et al., 2009; Cocchini et al., 2010). Informants rated the participants using the exact same items with slightly varied wording to refer to a third person.

To enable evaluation of the reliability of responses, six “check questions” were distributed evenly throughout the questionnaire. These questions allowed us to check for possible perseveration, comprehension deficits, inattention or visual spatial impairment that may prevent the respondent from attending to one side of the scale (Cocchini et al., 2010; Della Sala et al., 2009). The check questions were designed to elicit, when appropriately endorsed, scores in one extreme of the scale. Three of these check questions would have their appropriate response on the left end of the scale (0 – no problem or 1 – mild problem; see check question 2, 3 & 4) and three on the right end of the scale (3 – problem or 2 – moderate problem; see check questions 1, 5 & 6). For questions depicting tasks requiring a motor component, two versions were provided with left and right limb affected, to provide reliable check questions for people that may be experiencing weakness or paralysis in one side of the body (i.e., hemiparesis). Failure to respond correctly to any of the check questions was interpreted as a possible indicator of these various deficits. Therefore, VATAmem data were considered unreliable in participants who failed any check questions.
The 15 questions were presented in the same fixed pseudo-random order with memory dimensions (prospective/retrospective, self/environmentally cued and short/long-term) evenly distributed throughout the questionnaire. To minimise possible associated attentional disorders, such as visuospatial neglect (Cubelli, 2017; Della Sala et al., 2009), participants were shown one question at a time in a laminated A4 sheet presented in portrait orientation on the patient's ipsilesional side.

First, the practice question was presented to make sure the participant understood how to use the rating scale. Then, the examiner read the questions aloud, allowing time for the participant to read them again if they wished and to observe the vignettes. Special emphasis was made during administration to ensure that the responses reflected current abilities.

VATAmem - Total score

The VATAmem total score was calculated on the 15 memory-based questions and it ranged from 0–45. A discrepancy value was obtained subtracting the participants’ scores from informants’ scores. When two informants were available, the mean was calculated and this was compared with the patient’s score to calculate the discrepancy value. The discrepancy value ranges from -45 to 45, where a discrepancy value of 0 means perfect agreement. A positive discrepancy means that compared to the informant, the participant has overestimated his/her memory abilities; while a negative discrepancy value indicates that the participant has underestimated his/her own memory abilities. This score may provide information about possible depression or anxiety; however, we did not investigate this aspect in the current study.
VATAmem- Subscales

As it was our interest to examine variability of awareness across different types of memory failures, two main subscales were devised to measure awareness of prospective and retrospective memory loss. In Crawford and colleagues’ (2003) factorial examination of the PRMQ (2003), the authors found that although they included items reflective of various types of memories (i.e., short versus long-term memory, self versus environmentally cued memory, and prospective versus retrospective memory), only general memory, prospective memory, and retrospective memory were observed as independent factors. We thus examined two subscales to measure awareness of prospective and retrospective memory abilities. The Prospective memory subscale includes 7 questions evenly spread across the questionnaire (i.e., see questions 1, 2, 6, 7, 9, 12 & 13 of VATAmem questionnaire) with a total score ranging from 0 to 21; whilst the Retrospective memory subscale includes 8 questions also evenly distributed (i.e., see questions 3, 4, 5, 8, 10, 11, 14 & 15 of VATAmem questionnaire) with a total score ranging from 0 to 24.

2.3. Statistical analysis

As assumptions for parametric analyses were not met, Spearman correlations were conducted to examine the level of agreement between informant and patient reports of memory performance, and how informants’ reports mapped onto actual memory performance measures. The Crawford and Howell (1998) modified t-test was used to develop the VATAmem cut-off for anosognosia in the full memory scale, and for the prospective and retrospective subscales. Independent sample t-tests were conducted to examine differences in awareness of patients and
informants across prospective and retrospective scales. **Product moment Pearson correlations were used to examine the level of agreement between informants across the prospective and retrospective scales.**

As all assumptions for parametric analyses were met, product moment Pearson correlations were used to evaluate test-retest of the VATAmem overall score in patients tested on two separate occasions. Internal consistency was examined for both self and informant reports through Cronbach’s alpha and item sensitivity analysis. As all assumptions for parametric analyses were met, construct validity of the VATAmem was examined through Pearson-product correlations between self and informant reports in the VATAmem and the PRMQ. Partial and one-tailed Spearman correlations were conducted to examine the relationship between cognitive measures and anosognosia as determined by the VATAmem and the PRMQ.

### 3. RESULTS

#### 3.1. Neuropsychological assessment

Mean raw score of the sample on the MMSE was 25.94 \((SD = 2.87; \text{ range } = 19-30)\). Some patients did not complete the full battery of tests (see Table 1). In particular, one patient did not complete the RBMT but he showed evidence of long-term memory impairment on the initial Story Recall test.

As shown in Table 1, all patients showed a long-term memory impairment and two thirds of the sample had executive function deficits. Nearly half (49%) of our sample also showed language difficulties; whereas short-term memory was spared in the majority of the cases.
3.2. **Awareness of memory deficits**

*Prospective and Retrospective Memory Questionnaire (PRMQ)*

Based on the PRMQ, 54.9% \((n = 28)\) of patients were classified as unaware of their memory deficits following the Crawford et al.’s (2006) cut-off. Mean discrepancy scores were 20.2 \((SD = 9.8; \text{range} = 8 - 44)\) for patients unaware of their deficits and −4.7 \((SD = 7.5; \text{range} = −21 - 6)\) for patients aware of their deficits, indicating that patients who were aware of their deficits actually tended to overestimate compared to their informants. Within the subscales, 25 patients were classified as unaware of their retrospective memory failures versus 26 as unaware of the prospective memory failures. These patients largely overlapped with those classified as unaware by the total scale cut-off; however, two cases were classified as unaware on the prospective or on the retrospective scale but they were not deemed as unaware according to the overall scale’s cut-off.

*Visual analogue Test for Anosognosia (VATA-mem)*

Preliminary version of the scale

A series of Spearman correlational analyses for each of the 15 items was run between pairs of informants rating the same patient. In all cases, the correlation was significant (at least
\( \rho = 0.54, p = 0.01; d = 1.28 \). This result suggests that the content of the 15 items selected for the final version of the VATAmem was similarly interpreted and rated by different informants\(^1\).

Check questions

One informant and one patient had to be removed as they both failed one check question in the VATAmem. The remaining patients and informants provided the expected responses to the check questions.

Informant report and patient’s memory performance

A final total of 72 informants were included in the sample. Informant scores on the VATAmem were compared to the corresponding patients’ memory performance measured by the RBMT. A spearman correlation showed a significant association between informants’ reports and patients’ performance in standardized memory assessments (\( \rho = 0.33, p = 0.02; d = 0.70 \)), suggesting that the more severe the patient’s score on memory tasks, the worse informants reported patients’ memory to be.

Unawareness cut-off score

The unawareness cut-off was derived following a similar procedure adopted in Cocchini et al. (2010). An “informant discrepancy score” was calculated for the 22 pairs of informants that

\(^2\) Please visit https://www.dropbox.com/sh/127iz3pfng4tf7r/AAC39MM2r9STROPPrJmXupbsDa?dl=0 for access to the full scale with vignettes.
evaluated the same individual (i.e., two informants per patient). This score ranged from -45 to +45, with 0 meaning perfect agreement and ±45 complete disagreement. The mean and standard deviation of the informants’ discrepancy score was used to calculate a “discrepancy threshold”.

In order to calculate our discrepancy threshold, we used Crawford’s and Howell’s (1998) modified t-test. The mean value of the discrepancy score between the 22 pairs of informants and the standard deviation was $M = 3.75$ and $SD = 3.44$. The critical value of $t$ with d.f. = 21 in a two-tailed test was 2.080. According to the modified $t$-test by Crawford and Howell, the discrepancy value at which we would have this critical value of $t$ with a likelihood of less than 5% was 10.5.

Therefore, an overall cut-off score to indicate a significant discrepancy between the patient and informant was set to 10.5; scores above this value are suggestive of a significant lack of awareness on the part of the patient. Following Della Sala et al. (2009), additional cut off scores to signify the degree of unawareness were established. The first cut-off represented an average disagreement of 1 point in all 15 items. A discrepancy value between 10.6 and 15.0 included was then considered as indicative of mild anosognosia. The second cut-off represented an average disagreement of 2 points in all 15 questions, with discrepancy values between 15.1 and 30.0 included considered to be indicative of moderate anosognosia. Finally, a discrepancy value between 30.1 and the maximum discrepancy score of 45 was considered as indicative of severe anosognosia (see Table 2).

Two cut-off scores were derived in the same manner as above for the retrospective and prospective sub-scales considering the corresponding memory items. The mean and standard deviation of the informant discrepancy scores for the prospective subscale were $M = 1.89$ and $SD$
= 1.34, and $M = 2.23$ and $SD = 1.90$ for the retrospective subscale. The cut-off scores for the prospective and retrospective sub-scales are reported in Table 2.

**INSERT TABLE 2**

**Awareness for memory deficits**

Nearly a third of the 51 patients were classified unaware of their deficits with discrepancy values above 10.5. Different degrees of severity are reported in Table 3. **Note that this is not an epidemiological study and the percentage of unaware patients should not be generalized to the entire population of amnesic patients.**

Prospective memory questions seemed to elicit disagreement between patient and informant evaluations more often than Retrospective memory questions. A total of 15 patients were classified as unaware of their deficits in the Prospective subscale, in contrast to 9 in the Retrospective subscale. Within patients’ reports there were no differences in endorsement between prospective and retrospective scales in those unaware of their deficits ($t(14) = 1.80, p = .10$), meanwhile differences were observed in those aware of their deficits, who endorsed more problems in prospective than retrospective abilities ($t(35) = 3.08, p = 0.004; d = 0.41$). Differences though were also observed in informants, who endorsed more problems in prospective than retrospective abilities both for unaware ($t(14) = 4.04, p < 0.001; d = 1.02$) and aware patients ($t(35) = 4.62, p<0.001; d = 0.56$). **Agreement between informant reports across prospective and retrospective scales was compared on those patients who had two informants ($n = 22$). Results**
showed a good agreement between informants across both prospective ($r = 0.85, p < 0.001; d = 3.22$) and retrospective items ($r = 0.81, p < 0.001; d = 2.76$).

INSERT TABLE 3

Reliability, sensitivity & validity

Test-retest reliability

A total of 15 patients were retested on a separate occasion between 24 hours and 3 days after first assessment. A Pearson correlation analysis showed a high significant coefficient between test and retest ($r = 0.92, p < 0.001; d = 4.70$).

Internal Consistency and test sensitivity

Internal consistency was evaluated through Cronbach’s alpha. The internal consistency of the whole scale for self-evaluations was of $\alpha = 0.91$ and $\alpha = 0.90$ for informant evaluations. Internal consistency for subscales was of $\alpha = 0.88$ and $\alpha = 0.85$ for self and informants in the Prospective subscale scale and $\alpha = 0.81$ for self and $\alpha = 0.81$ for informant’s reports on the Retrospective subscale.

In terms of overall sensitivity, the VATAmem identified 15 patients as being unaware of their memory deficits while the PRMQ identified 28 patients ($\chi^2 = 12.68, p < 0.001; \varphi = 0.50$), suggesting that the VATAmem diagnostic criteria may be more conservative, but also less prone to false positives than the PRMQ. Of the 15 patients identified by the VATAmem as unaware of
their deficits, 93.3% (n = 14) were also classified by the PRMQ. Out of the 28 identified as unaware by the PRMQ 50% (n = 14) were also identified by the VATAmem. Thus 14 patients overlapped as unaware by both scales and 15 mismatched. Out of the 15, 14 were classified unaware only by the PRMQ, and 1 only by the VATAmem. To further explore the reasons underlying the mismatched cases several analyses were conducted. Specifically, we examined the subset of cases that were deemed as unaware by the PRMQ but not the VATAmem, compared to those in which both agreed upon as being unaware of their deficits. No significant differences were observed across all cognitive domains between the groups (all p > 0.05). Following these results, we explored if differences between these groups laid within the reports of either patients or informants. Results showed that informants reported similarly in the VATAmem and the PRMQ. Specifically, informants on the VATAmem reported significantly less memory difficulties in those that were deemed unaware only by the PRMQ as opposed to those who were deemed unaware by both measures (t (26) = -4.72, p < 0.001; d = 1.78). Similarly, informants on the PRMQ endorsed less memory difficulties in those deemed unaware only by the PRMQ (t (26) = 0.82, p = 0.06; d = 0.74) in line with informants reports on the VATAmem, though this difference was not significant. Further, Spearman correlations showed that informant reports mapped on similarly to memory performance on both measures (i.e., VATAmem, rho = 0.33 p = 0.02, d = 0.07; PRMQ, rho = -0.47, p < 0.001, d = 1.06). With regard to patients’ reports, results showed that patients were endorsing less memory problems on the VATAmem when they were deemed unaware by both the PRMQ and the VATAmem than when they were deemed unaware only by the PRMQ (t (26) = 2.93, p = 0.007; d = 1.11). Interestingly, patients’ reports on the PRMQ revealed no significant differences between those deemed unaware only by the PRMQ versus those by both measures (t (26) = 0.27,
Therefore, the diagnosis discrepancy between the two questionnaires for 14 patients was mainly driven by their self-evaluation being closer to informants’ evaluation in the VATAmem, but not in the PRMQ.

Finally, as with the previous VATAs, we conducted an item level analysis to examine each item’s sensitivity “correctly detected”, and specificity “correctly not detected” respective of the overall VATAmem unawareness cut off (e.g., total scale anosognosia cut-off). That is, the extent to which i) an individual item showed a positive discrepancy (i.e., evidence of over-estimation of own memory ability) between self and informant when a patient was deemed unaware based on the overall VATAmem score (Correctly detected, Hit); and ii), when a single item showed no discrepancy or negative discrepancy when the patient was classified as aware based on the overall VATAmem score (Correctly not detected, CND). As reported in Table 4, item analysis revealed that on average the items showed a relatively high sensitivity and specificity ($M = 74.8; SD = 0.04$) for detecting unawareness as defined by the overall scale. Items such as question 7 “remembering appointments”, question 9 “walking into a room” and question 11 “knowing your way around your home/ward” had the highest sensitivity and specificity.

INSERT TABLE 4

Validity

Self and informant evaluations on the VATAmem were compared to ratings provided in the PRMQ. Both self ($r = 0.64$, $p < 0.001; d = 1.67$) and informant ($r = 0.67$, $p < 0.001; d = 1.81$) evaluations were significantly associated with those reported in the PRMQ. Further, patient-
informant discrepancy scores for PRMQ and VATAmem were also correlated \( (r = 0.68, p < 0.001; d = 1.85) \) (see Figure 2).

3.3. Demographics, cognitive functions & awareness

Demographics, raw scores of global cognition, standardized scores of memory and executive function, and language scores were examined in relation to the overall VATAmem discrepancy scores. Pearson correlations revealed no association between anosognosia and age \( (r = 0.01, p = 0.97; d = 0.02) \) or education \( (r = 0.03, p = 0.85; d = 0.06) \). Independent sample t-test showed no significant differences in unawareness in relation to gender \( (t (49) = -1.81, p = 0.08, d = 0.50) \). Partial correlations adjusted for demographics showed a negative correlation between anosognosia and global cognition \( (r = -0.38, p = 0.008; d = 0.82) \). One-tailed Spearman correlations revealed a significant association between the severity of the memory impairment and unawareness \( (RBMT-2,3; \rho = 0.31, p = 0.01; d = 0.65) \), no significant association was found between the severity of executive functions difficulties and unawareness (Switching task and TMT; \( \rho = 0.22, p = 0.06; d = 0.45) \). No significant association between anosognosia and the language index was found \( (\rho = -0.12, p = 0.21; d = 0.24) \). The PRMQ was also significantly correlated with global cognition \( (r = -0.28, p = 0.03; d = 0.58) \) and memory \( (\rho = 0.38, p = 0.003; d = 0.82) \). In contrast to the VATAmem, the PRMQ was significantly correlated to language index \( (\rho = -0.26, p = 0.04; d = 0.54) \) and executive function \( (\rho = 0.25, p = 0.04; d = 0.52) \).
4. DISCUSSION

The main aim of this study was to develop a measure of explicit anosognosia for both prospective and retrospective memory impairment, capitalizing on current available tools and overcoming some of their pitfalls. As with other similar measures (VATA-L and VATAm - Cocchini et al., 2010; Della Sala et al., 2009) we aimed to provide a psychometrically sound measure that could not only detect anosognosia, but also distinguish between different levels of severity unawareness providing cut-offs for mild, moderate and severe anosognosia.

With regard to data obtained from the informants, we observed a significant association between their reports and the patients’ memory deficits on neuropsychological testing. Thus, although it has been noted that variables such as caregiver culture, burden and mood related disorders can affect informant reports of someone’s memory abilities (Prigatano, 2005; 2010), our informants, as a group, appeared to be a reasonably reliable source of information regarding patients’ level of memory functioning. However, exceptions can also occur and it is important to ensure that every informant’s rating is reliable. To this aim, the VATAmem informant version also requires informants to answer check questions to ensure reliability of their responses. One informant failed on of these check question. Interestingly, this participant endorsed the majority of the items on the VATAmem as severe, in contrast to the other informant who did not fail the check question and endorsed items of moderate and mild difficulties in line with the patient’s performance on standardized memory assessments. This result suggests that the check questions can provide a useful way to gauge the informant’s reliability, avoiding potential false positives, as it would have been for this patient if the first informant’s data were not excluded as unreliable.
Regarding its psychometric properties, the VATAmem had strong internal consistency and reliability across time, rendering it useful as a follow-up measure. In terms of its validity, the VATAmem was associated with the PRMQ, another similar measure of anosognosia, suggesting that it taps into similar self-reflective abilities captured with the PRMQ. A difference though was observed when the cut-offs for each measure were applied. With regard to rates of anosognosia according to the VATAmem, 29.4% (n=15) of the sample was classified as unaware of their deficits (11.8% of the sample classified as mildly unaware, 15.7% as moderately unaware and 2% as severely unaware). In contrast, the PRMQ identified more patients as unaware of their deficits (54.9%; n = 28). These results could reflect that the PRMQ is more sensitive to anosognosia or that the VATAmem is more specific. Post hoc analyses, between patients classified as unaware by both measures and those classed as unaware by the PRMQ only, supported this latter account. Indeed, when patients’ self-evaluations were measured with the VATAmem, their evaluations were closer to what informants believed their memory to be (in the mismatched cases; unaware by PRMQ only).

The VATAmem has been designed to account for possible associated cognitive difficulties that may interfere with assessment. This is first achieved by providing a measure of reliability (with check questions). For example, a large proportion of patients showed executive function difficulties that could result in possible perseverations for a particular response/rating. By including check questions, it ensures adequate levels of reliability in their responses accounting for lack of understanding, attentional deficits, perseverations, and/or poor compliance. Moreover, to minimize the interference of other possible cognitive deficits, the VATAmem shapes the questionnaire in a way that questions and ratings are presented in various forms to
compensate for other difficulties. Not only does the VATAmem aim to control for reliable final responses, it also aims to minimize the impact of cognitive deficits arising from ABI. For example, long written or open-ended questions may represent a challenge for patients showing memory, attention or language difficulties, and limit an accurate appraisal of their awareness. Similar to other VATAs for motor and language deficits, this scale was developed to include vignettes depicting common memory related mistakes and a visual analogue scale to reduce the demand on cognitive functioning. Indeed, we observed that the PRMQ was associated with language and executive function abilities while the VATAmem was not. The nonverbal aid of the images not only may have released, at least in part, the load of memory processing in these patients, but also provides less direct questioning. Following, when a patient is directly questioned about a deficit, psychological processes may cloud self-report. For example, different patients may have different conceptualizations of what is appropriate to share, or they might be worried about revealing too much of a deficit and what consequences might follow. Denial mechanisms at a pre-conscious level have also been suggested as underlying unawareness of memory deficits (Turnbull, Jones, & Reed-Screen, 2002; Weinstein, 1991). By providing examples of memory failures in the third person, patients might feel less threatened by the inquiry and thus provide self-reports that are a closer representation of their actual knowledge of their deficits (Clare et al., 2012). By supporting the patient with visual information, and offering them an opportunity to gain some distance from the topic discussed, the VATAmem might be accessing a more accurate measure of a patient’s true knowledge of their deficit. This might aid in minimizing the risk of false positives, which as Baier and Karnath (2005) point out is a serious caveat of traditional measures of anosognosia. In line with this idea, analyses of patients’ and informants’ reports with
regard to mismatched unaware cases by the VATAmem and the PRMQ revealed that although informants’ reports were comparable across both measures, patients’ reports differed. Indeed, when evaluated with the VATAmem patients that were deemed unaware by the PRMQ only, endorsed more difficulties than those patients deemed unaware by both measures. This difference was not observed in patients’ reports as evaluated by the PRMQ.

The VATAmem also enables measurement of awareness for prospective versus retrospective memory failures. Interestingly, we found that more patients were categorized as unaware of prospective memory deficits ($n = 15$) than retrospective ($n = 9$). As noted in previous studies, awareness for prospective and retrospective memory abilities/difficulties can differ (Mantyla, 2003; Wilkins & Baddeley, 1978), and healthy aging individuals rating themselves or others have a tendency to endorse higher levels of difficulty in prospective memory than in retrospective memory (Crawford et al., 2006; Crawford et al., 2003). In our study, we observed that patients’ reports were comparable across prospective and retrospective scales, but it was the informants who endorsed more problems in prospective than retrospective abilities. Interestingly, no differences were observed with regard to the agreement between pairs of informants across the prospective and retrospective scales. Indeed, informants had a high agreement across both scales. The differences regarding the incidence of awareness could therefore be reflective of the inherent differences these memories. Following Mantyla (2003), informants might be more sensitive to prospective memory failures given the concerning consequences of these memory lapses (e.g., missing a doctor’s appointment, forgetting to deliver an important message, etc.). These lapses might be thus more obvious and emotionally salient for informants. Another possibility is that prospective memory failures are more common than
retrospective memory failures and thus informants would note these difficulties more frequently. Finally, an interesting point to consider is the qualitative differences regarding how individuals become aware of their prospective and retrospective memory difficulties. For example, with retrospective memory failures, individuals’ are immediately confronted with their error, as they cannot remember the information they are intending to recall. Prospective memory on the other hand, error awareness is experienced at a future point in time, when the failure to remember would become apparent (e.g., missing an appointment). This quality in prospective memories might pose a further challenge for patients’ awareness, as these failures might not be as apparent for them or emotionally salient compared to their informants. These inherent differences could explain why if prospective memory failures are more common than retrospective memories failures (as reported by informants) patients’ report similarly on prospective and retrospective memories. Future studies should examine these possibilities, along with patient performance and ratings in both retrospective and prospective memories, to determine more precisely the basis of differences in awareness scores for these two types of memory.

Separate analyses were conducted to compare several cognitive measures to the discrepancy values of the VATAmem. A strong relationship between memory functioning and unawareness of memory deficits was shown, similarly to global cognition as measured by the MMSE. These results showed that overall, those who were unaware of their memory difficulties performed worse in memory and global cognitive tasks, suggesting a role of memory abilities and overall cognitive deterioration in supporting awareness in this population (see Agnew & Morris, 1998; Ansell & Bucks, 2006; Mograbi, Brown, & Morris, 2009). Finally, although previous reports
have found unawareness for memory loss to be associated with executive function (Lopez, Becker, Somsak, Dew, & DeKosky, 1994; Michon, Deweer, Pillon, Agid, & Dubois, 1994; Reed, Jagust, & Coulter, 1993), we did not replicate this finding. However, this result is not uncommon, as others have also failed to replicate this relation (Starkstein, Sabe, Chemerinski, Jason, & Leiguarda, 1996; Vogel, Hasselbalch, Gade, Ziebell, & Waldemar, 2005). The lack of consistency across studies on anosognosia has been interpreted by some as representative of the multifactorial nature of anosognosia (Cocchini et al., 2012; Cocchini, Beschin, & Sala, 2002; Fotopoulou, 2014; Marcel, Tegnér, & Nimmo-Smith, 2004; Orfei et al., 2007; Vuilleumier, 2004). That is, anosognosia is not a unitary syndrome and different subtypes of anosognosia may exist that may be linked to different associated cognitive abilities and deficits (Agnew & Morris, 1998; McGlynn & Schacter, 1989). Others have also highlighted that although executive function can contribute to unawareness, processes that are specific to self-evaluation, such as self-monitoring, are more likely to underlie unawareness of memory deficits (Cosentino, Metcalfe, Butterfield, & Stern, 2007; Rosen, 2011; Rosen et al., 2014).

This study presents some limitations that should be considered, including the lack of evaluation of mood, a relatively small sample size and a heterogeneous etiology of brain injury. Second, although the sample was heterogeneous, it was largely formed (>70%) by older patients who had suffered from a stroke. Future work should therefore examine its applicability in different clinical and demographic samples. Nonetheless, the absence of an association between VATAmem scores and demographic variables such as age, education and gender in the current sample is promising in terms of its utility for other groups, such as a younger, predominantly male sample of TBI. With regard to reliability, it should be noted that test-retest was conducted within
1-3 days. This timeframe may not reflect clinical settings (e.g., in rehabilitation) where follow up intervals are longer. Finally, visuospatial neglect was not assessed in this study and could have impacted the results. Overall, the VATAmem can provide a useful tool to measure anosognosia for memory loss. The release of cognitive load and the use of vignettes may provide a more accurate assessment of one’s awareness of memory deficits than that obtained from verbally based measures. The VATAmem can also be used to explore variability of awareness across prospective and retrospective memories as supported in this study. Since anosognosia is a multifactorial syndrome, it is likely that similar mechanisms underlie anosognosia for different deficits, whilst different mechanisms may underlie anosognosia for the same deficit. The VATAmem is the third of a series of similarly standardized measures, and thus allows standardized assessment of awareness across different deficits.

5. REFERENCES


Do you have problems remembering directions, moments after they have been given to you?

No Problem Problem
0 -------- 1 -------- 2 -------- 3
Figure 2. Scatterplot of patient-informant discrepancy scores for PRMQ and VATAmem.
Table 1. Neuropsychological assessment of overall sample of patients with ABI.

<table>
<thead>
<tr>
<th>Cognitive functions impairment</th>
<th>% of patients showing pathological performance (n/N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Term Memory (LTM) impairment</td>
<td>100% (50/50)</td>
</tr>
<tr>
<td>Mild</td>
<td>26% (13/50)</td>
</tr>
<tr>
<td>Moderate</td>
<td>38% (19/50)</td>
</tr>
<tr>
<td>Severe</td>
<td>36% (18/50)</td>
</tr>
<tr>
<td>Short Term Memory (STM) impairment</td>
<td></td>
</tr>
<tr>
<td>Verbal STM</td>
<td>2% (1/51)</td>
</tr>
<tr>
<td>Visuospatial STM</td>
<td>22% (9/41)</td>
</tr>
<tr>
<td>Executive functioning impairment</td>
<td>57% (29/51)</td>
</tr>
<tr>
<td>Mild</td>
<td>6% (3/51)</td>
</tr>
<tr>
<td>Moderate</td>
<td>6% (3/51)</td>
</tr>
<tr>
<td>Severe</td>
<td>45% (23/51)</td>
</tr>
<tr>
<td>Language functioning impairment</td>
<td>51% (25/49)</td>
</tr>
</tbody>
</table>

Summary of cognitive abilities of sample of 51 ABI patients. \( n = \) total patients with cognitive impairment; \( N = \) total patients with available data on cognitive measures. LTM: Performance on the RBMT.

Table 2. Awareness cut-off scores for total scale, prospective and retrospective subscales with degrees of severity

<table>
<thead>
<tr>
<th>Degree of unawareness</th>
<th>Mild anosognosia</th>
<th>Moderate anosognosia</th>
<th>Severe anosognosia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aware</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total scale</td>
<td>0.0-10.5</td>
<td>10.6-15.0</td>
<td>15.1-30.0</td>
</tr>
<tr>
<td>Prospective subscale</td>
<td>0.0-4.7</td>
<td>4.8-7.0</td>
<td>7.1-14.0</td>
</tr>
<tr>
<td>Retrospective subscale</td>
<td>0.0-6.4</td>
<td>6.5-8.0</td>
<td>8.1-16.0</td>
</tr>
</tbody>
</table>
Table 3. Patients classified as aware and unaware of their memory deficits

<table>
<thead>
<tr>
<th>Degree of unawareness</th>
<th>Anosognosia</th>
<th>Mild anosognosia</th>
<th>Moderate anosognosia</th>
<th>Severe anosognosia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total scale</td>
<td>15 (29.4%)</td>
<td>6 (11.8%)</td>
<td>8 (15.7%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Prospective subscale</td>
<td>15 (29.4%)</td>
<td>5 (9.8%)</td>
<td>9 (17.6%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Retrospective subscale</td>
<td>9 (17.6%)</td>
<td>1 (2%)</td>
<td>7 (13.7%)</td>
<td>1 (2%)</td>
</tr>
</tbody>
</table>

Table 4. Item level percentages of HITS (correct detected) and CND (correct non detected) for total scale and subscales for prospective and retrospective memory.

<table>
<thead>
<tr>
<th>Items</th>
<th>Total scale HITS + CND %</th>
<th>Prospective HITS + CND %</th>
<th>Retrospective HITS + CND %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Doing something</td>
<td>66.7 %</td>
<td>70.6 %</td>
<td>-</td>
</tr>
<tr>
<td>Q2. Posting a letter</td>
<td>72.5 %</td>
<td>72.5 %</td>
<td>-</td>
</tr>
<tr>
<td>Q3. Directions</td>
<td>74.5 %</td>
<td>-</td>
<td>70.6 %</td>
</tr>
<tr>
<td>Q4. Drinking coffee</td>
<td>72.5 %</td>
<td>-</td>
<td>72.5 %</td>
</tr>
<tr>
<td>Q5. Same story</td>
<td>70.6 %</td>
<td>-</td>
<td>58.8 %</td>
</tr>
<tr>
<td>Q6. Turn off the cooker</td>
<td>74.5 %</td>
<td>74.5 %</td>
<td>-</td>
</tr>
<tr>
<td>Q7. Appointment</td>
<td>80.4 %</td>
<td>76.5 %</td>
<td>-</td>
</tr>
<tr>
<td>Q8. Peoples names</td>
<td>78.4 %</td>
<td>-</td>
<td>74.5 %</td>
</tr>
<tr>
<td>Q9. Walking into a room</td>
<td>80.4 %</td>
<td>84.3 %</td>
<td>-</td>
</tr>
<tr>
<td>Q10. The time</td>
<td>72.5 %</td>
<td>-</td>
<td>76.5 %</td>
</tr>
<tr>
<td>Q11. Home/ward</td>
<td>80.4 %</td>
<td>-</td>
<td>84.3 %</td>
</tr>
<tr>
<td>Q12. Umbrella</td>
<td>72.5 %</td>
<td>72.5 %</td>
<td>-</td>
</tr>
<tr>
<td>Q13. Saying</td>
<td>74.5 %</td>
<td>74.5 %</td>
<td>-</td>
</tr>
<tr>
<td>Q14. Introduced</td>
<td>74.5 %</td>
<td>-</td>
<td>76.5 %</td>
</tr>
<tr>
<td>Q15. Names</td>
<td>76.5 %</td>
<td>-</td>
<td>84.3 %</td>
</tr>
</tbody>
</table>

Total scale mean HITS + CND % = 74.5 %; SD = 0.05%. Prospective subscale mean HITS + CND % = 73.1 %; SD = 0.08%. Retrospective subscale mean HITS + CND % = 74.8 %; SD = 0.08%.