

Assessment of Self-Awareness of Cognitive Function: Correlations of Self-Ratings with Actual Performance Ranks for Tests of Processing Speed, Memory and Executive Function in Non-Clinical Samples

Johannes Rothlind^{1,*}, Paul Dukarm², Matthew Kraybill³

¹San Francisco VA Medical Center, Psychology Services, San Francisco, CA, USA

²Alaska Neuropsychological and Behavioral Health, LLC, Neuropsychology, Fairbanks, AK, USA

³Ross University School of Medicine, Behavioral Science, Roseau, Dominica

*Corresponding author at: San Francisco VA Medical Center, Psychological Services (116B), 4150 Clement Street, San Francisco, CA 94121, USA.

E-mail address: johannes.rothlind@va.gov (J. Rothlind).

Editorial Decision 20 November 2016; Accepted 30 November 2016

Abstract

Objective: For individuals with neurologic disorders, self-awareness of cognitive impairment is associated with improved treatment course and clinical outcome. However, methods for assessment of levels of self-awareness are limited, and most require collateral information, which may not be readily available. Although distortions in self-awareness are most often associated with low cognitive ability, the frequently mixed pattern of cognitive strengths and deficits in individuals with neurologic disorders complicates assessment. The present study explores relationships between actual test performance and self-ratings, utilizing a brief probe administered during testing. The “common-metric” approach solicits self-appraisal ratings in percentile equivalents and capitalizes on available normative data for specific standardized neuropsychological tests to allow direct comparisons.

Method: A convenience sample of 199 adults recruited from community sources participated in this study, including healthy adults and neuropsychologically “at-risk” volunteers who were HIV positive and/or endorsing heavy current alcohol consumption. Immediately following completion of standardized neuropsychological tests, participants estimated their own percentile ranking.

Results: Across study groups, participant’s estimates of their own percentile rank were modestly correlated with actual performance ranking. Highest correlations were obtained for tests of learning, memory and conceptual reasoning, and executive function, with smaller correlations for simple tests of motor and psychomotor speed.

Conclusions: The study reveals normal biases affecting the self-appraisal during standardized testing, and suggests that a common-metric approach for assessing self-appraisal may play a role in establishing clinical thresholds and identifying and quantifying reductions in insight in persons with neuropsychological deficits

Keywords: Assessment; Meta cognition

Introduction

Self-awareness of cognitive ability is often compromised in neuropsychiatric disorders (McGlynn & Schacter, 1989; LaBuda & Lichtenberg, 1999; Rosen et al., 2010). These deficits in self-awareness often manifest in underestimation of cognitive impairment (e.g., following traumatic brain injury and in dementing disorders), but in some instances, the bias may be in the direction of exaggerated negative self-ratings, such as in the setting of depressive symptoms (Blackstone et al., 2012; Chin, Oh, Seo, & Na, 2014; Lerner et al., 2014; Zlatar, Moore, Palmer, Thompson, & Jeste, 2014). Following traumatic brain injury, impaired self-awareness of cognitive abilities (SA-C) has been linked to poor adherence to therapy, less appropriate goal setting in rehabilitation, and lower clinician rating of employability following rehabilitation (Lam, McMahon, Priddy, & Gehred-Schultz, 1988; Sherer et al., 2003; Fischer, Trexler, & Gauggel, 2004). In severe mental illness, deficits in SA-C have been

shown to be associated with lower treatment adherence (Smith et al., 1999) and impaired vocational performance (Lysaker, Bell, Milstein, Bryson, & Beam-Goulet, 1994). In dementing illnesses, impaired SA-C has been associated with negative outcomes including failure to modify driving behavior (Cotrell & Wild, 1999), difficulties with everyday decision-making regarding medication management (Cosentino, Metcalfe, Cary, De Leon, & Karlawish, 2011), and greater caregiver burden and distress (Selzer, Vasterling, Yoder, & Thompson, 1997; Vogel, Waldorff, & Waldermar, 2010).

In light of these findings, neuropsychologists need techniques for assessment of patient's insight into their level of disability (SA-C). Unfortunately, there are few well-established standardized methods for evaluating this meta-cognitive ability (reviewed in Prigatano, 2010). Most clinical research on self-appraisal accuracy in neurologic diseases has focused on discrepancies between ratings of behavioral functioning obtained from patient and an informant (e.g., caregiver), or patient and clinician (Fleming, Strong, & Ashton, 1996; Clare, 2004; Ecklund-Johnson & Torres, 2005; Pannu & Kaszniak, 2005; Prigatano, 2005; Vanderploeg, Belanger, Duchnick, & Curtiss, 2007). However, such an approach is time consuming, input from knowledgeable informants is not always easily available in clinical settings, and informant data may be vulnerable to reporting biases and other sources of inaccuracy (Dalla-Barba, Parlato, Iavarone, & Boller, 1995; Duke, Seltzer, Seltzer, & Vasterling, 2002). Moreover, clinicians and other informants are not privy to the subjective emotional and motivational states that influence patients' behaviors in a specific situation, which potentially confounds their interpretation of the patient's ability and functioning. The questions asked of patients and other informants do not typically focus on a clearly defined event or behavior, which means that the two raters may focus on very different circumstances in making their ratings, and may thus both be accurate even if numerically quite discrepant. Similarly, when a clinician's judgments are used as a reference, these typically involve synthesis of information drawn from multiple sources or a narrow sample of a particular kind of behavior or functional capacity based only on observations within the clinic setting.

Alternative approaches have involved reference to test performance as the standard, but this approach is rarely used (Ecklund-Johnson & Torres, 2005). Although it could be argued that examinees naive to a particular neuropsychological test would not have adequate points of reference for giving SA-C rankings from this perspective, social comparison judgements are made rapidly in many situations and contexts over the life span, and clinically unimpaired adults must routinely access memories of past experiences (both episodic and semantic and autobiographic memory) to guide their own self-ratings in new situations involving a social comparison frame of reference. Anderson and Tranel (1989) were among the first to compare responses on an awareness interview with performance on cognitive tests in a mixed sample of neurologically impaired patients, using deviation from established test norms to define impairment in SA-C. They found that SA errors were associated with unilateral right-hemisphere damage, lower verbal IQ, and greater temporal disorientation. In another study involving participants with moderate and severe head injuries, Allen and Ruff (1990) obtained individuals' subjective ratings of functioning across various cognitive domains and contrasted *z*-scores on these ratings with scores on standardized neuropsychological measures evaluating the same domains. Participants with more severe head injuries were less accurate in their SA-C, particularly in the domains of attention and sensorimotor functioning. The authors suggest that accurate SA-C following acquired brain injury requires an adjustment in self-perception, and that more severely brain-damaged individuals have difficulty in modifying or updating their perspective regarding their own abilities.

Similar methods have also been employed in studies examining SA-C in dementia. In a study by McGlynn and Kazniak (1991), participants with Alzheimer's Dementia (AD) were provided information regarding the average scores expected from healthy, age-matched peers, and were then asked to predict what score they would obtain on various neuropsychological tests. In comparison with caregivers, AD patients showed markedly inflated predictions of their performances, most notable in the domain of delayed recall, despite generally accurate prediction of their relative's performance on the same tasks. In another study, individuals with dementia were asked immediately after completing cognitive tasks how well they thought they had performed compared with other people of their age (e.g., below average, average, or above average; Graham, Kunik, Doody, & Snow, 2005). SA-C accuracy was operationalized by Graham and colleagues as the discrepancy between self-ratings and demographically adjusted scores, and was shown to be inversely related to dementia severity.

In a more recent study, a social comparison framework has been utilized to objectively measure aspects of SA-C in a mixed sample of patients with dementia by asking participants to estimate percentile equivalent for their own performance (Williamson et al., 2010). These rankings were completed both before and immediately after completing standardized tests. The ranking estimates were later contrasted with the actual percentile ranking for the same test to yield a discrepancy score quantifying SA-C accuracy. The study revealed markedly lower SA-C accuracy in individuals with AD or frontotemporal dementia when judging their own neuropsychological test performance using this method. Furthermore, ratings obtained using this social comparison method showed robust correlation with more traditional indicators of diminished awareness derived through calculation of patient-caregiver discrepancies (Williamson et al., 2010).

Questions remain about (a) the specificity of SA-C for different task domains (e.g., memory vs. executive functioning, motor speed, and visuospatial problem solving) and (b) the extent to which measuring this domain is constrained by

factors such as the psychometrics of different tests, floor and ceiling effects in different patient groups and healthy controls, and by natural biases affecting all performance ratings (e.g., representativeness heuristic). It is not uncommon for individuals with disorders of brain functioning to display spared cognitive abilities in some domains, and neurologically healthy adults frequently perform in the impaired range on a few tests when large batteries of tests are utilized (Schretlen, Testa, Winicki, Pearson, & Gordon, 2008). Efforts to develop valid approach to clinical measurement of SA-C must therefore also resolve questions regarding possible biases in self-appraisal of residual cognitive *strengths* in the context of neurologic disease, and the degree to which individuals *without* clinical disorders of insight and self-awareness are biased in appraising their cognitive strengths and weaknesses also remains to be evaluated.

In summary, further research is needed to develop brief standardized methods of assessing SA-C that can be employed in situations where it is important to evaluate this aspect of meta-cognitive ability. Research is needed to refine methods and to understand factors that mediate SA-C accuracy across the wider spectrum of cognitive strengths and vulnerabilities observed in association with healthy cognitive functioning and neurologic disease.

Objectives of the Current Study

The goal of the current study was to examine correlates of SA-C in a mixed sample of adults, including healthy controls and other non-clinical samples without known or suspected impairments of insight or self-awareness. We examined SA-C contrasting estimates of study participants performance on a range of specific neuropsychological tests with norms-based ranking. Following completion of different neuropsychological tests, the examinee was asked to compare their own performance on a specific neuropsychological test to the distribution of scores they would expect of demographically matched peers on that same test. The examinee provided SA-C rankings immediately after completion of specific tests, and the estimate was generated for performance on the same test for which actual normative information is available, using the same metric (percentile equivalents) referred to by the clinician in evaluating the examinee's performance. The examinee's subjective estimate was then compared with the percentile equivalent obtained from published normative data.

The primary aim of the study was to identify tests that hold promise for use of the SA-C discrepancy method to identify abnormal biases and distortions in self-appraisal. Our criterion for potentially valuable measure were that the test showed at least a modest correlation between actual and estimated performance across the range of functioning, so that it was expected for high scorers to rate themselves as higher on the test than low scorers across groups of individuals not known to have major disturbance of self-awareness. We investigated SA-C for commonly used neuropsychological measures assessing learning and memory, visuospatial problem solving, motor speed, psychomotor speed, and executive function.

We hypothesized that in a community sample of individuals displaying a broad range of scores across a battery of tests but not currently suffering from major neurologic or psychiatric disability, the primary correlate of SA-C ranking would be performance in the immediately preceding task for which the ratings were obtained. Secondary aims of the study were to explore the relationship of SA-C rankings for different tests to demographic and clinical factors (e.g., education, health status, alcohol consumption, and mood).

Materials and Methods

Participants

Study participants included 199 adults consecutively enrolled in an NIAAA-sponsored longitudinal study of the effects of alcohol use in HIV disease (Rothlind et al., 2005). The 199 participants are a subset of the larger community sample ($N = 268$) recruited to participate in the study for whom SA data were obtained using our method; the SA-C procedure was added to the study protocol after the study was already underway. Among the participants who completed the SA-C ratings, 63 were HIV– but reporting heavy drinking (HIV–/HD), 46 were HIV+ and reporting light or no drinking (HIV+/LD), 35 were HIV+ and reporting heavy drinking (HIV+/HD), and 55 were HIV– and reporting light or no drinking (HIV–/LD; see Rothlind et al., 2005 for more on the operational definition of LD and HD in this sample). The mean age of the study sample was 42.6 ± 8.0 , and mean years of education was 14.8 ± 2.3 . Estimated verbal IQ based on AMNART score and Education (Grober, Sliwinski, & Korey 1991) was 114.9 ± 9.7 . The average reported number of drinks during the past week ranged from 2.4 ± 3 in the LD groups to 39.3 ± 30.6 in the two HD groups. Mean score on the Beck Depression Inventory (BDI) (Beck & Steer 1987) was 10.1 ± 8.4 in the combined sample. The participants in the current study did not differ significantly on any demographic parameters from the remaining 69 participants reported previously (Rothlind et al., 2005).

The present sample constituted a convenience sample for the purposes of the present investigation. However, the availability of discrete subgroups offered an opportunity to examine the stability/generalizability of the main study findings across multiple independent samples, and to also examine SA-C in a group at increased risk of neurocognitive impairment (HIV positive and heavy alcohol use) in addition to a more highly educated control sample with no identified risk factors for neurocognitive impairment. This study was carried out using a protocol approved by the Institutional Review Boards (IRB) of the San Francisco Department of Veterans Affairs Medical Center and the University of California, San Francisco. All participants gave written informed consent.

Measures

All study participants completed a large battery of neuropsychological tests (see Rothlind et al., 2005), including eight measures that formed the basis for the SA-C ratings. The tests utilized in the present study are listed below in order of administration: *Rey–Osterrieth Complex Figure* (ROCF; Rey, 1941; Osterrieth, 1944), norms from Denman (1984); *California Verbal Learning Test* (CVLT; Delis, Kramer, Ober, & Kaplan, 1987); *Trail Making Test* (TMT; Reitan, 1958), Heaton, Grant, and Matthews (1991) norms; *Symbol Digit Modalities Test Oral and Written versions* (SD-O and SD-W; Smith, 1973); *Grooved Pegboard* (GP; Matthews & Klove, 1964), Heaton and colleagues (1991) norms; *Brief Visuospatial Memory Test-Revised* (BVMTR; Benedict, 1997); *Controlled Oral Word Association Test* (COWAT; Benton, Hamsher, & Sivan, 1994); norms from Ruff, Light, Parker, & Levin (1996); *Short Category Test* (SCT; Wetzel & Boll, 1987). To limit the time and burden of this ancillary investigation on study participants while sampling a wide range of tasks and abilities, we restricted our queries regarding SA-C to the above measures.

Procedures

Self-Appraisal of Neuropsychological Performance

In our “common-task, common-metric” (CM) approach for assessing SA-C, participants were asked to estimate their own ranking on selected tests immediately *after* completion of those tests. More specifically, they were asked to estimate what

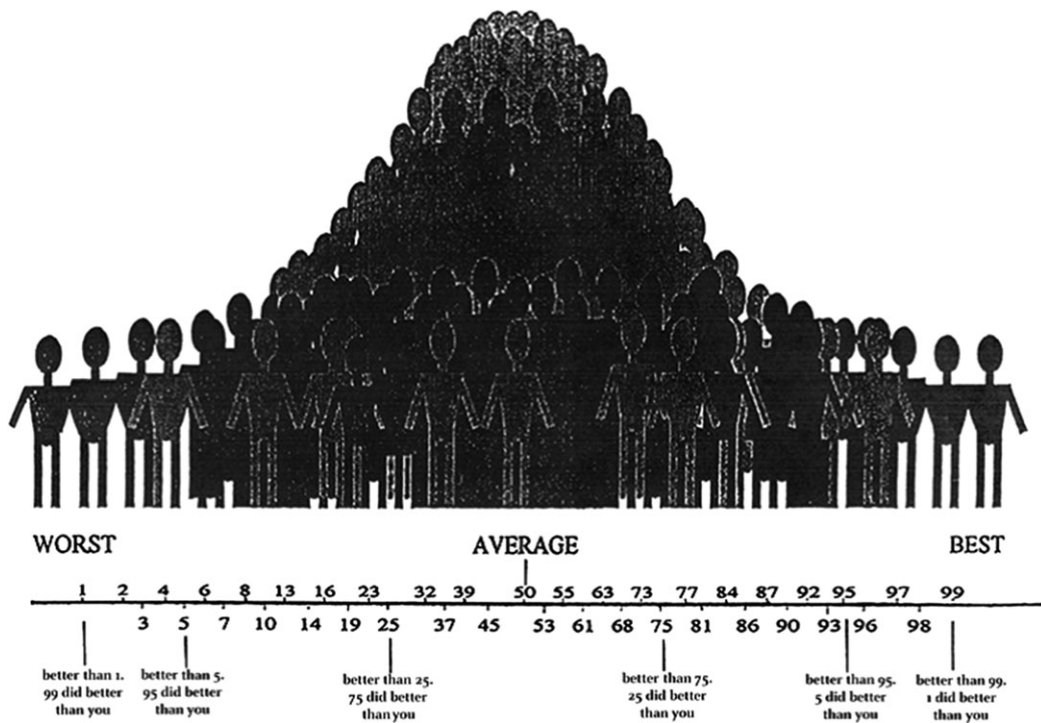


Fig. 1. Visual depiction of normal distribution of test scores used in orienting study participants to the SA-C task for this study. SA-C, self-awareness of cognitive abilities.

percentile score they believe that they would receive for the test if compared with a randomly selected healthy demographically matched peer group. A graphic depiction of the normal distribution (Fig. 1) served as a visual aid and was referred to explicitly during the instructions. The instructions asked study participants to rate themselves in comparison with peers matched according to the stratification applied in the published normative samples that were used as a reference (e.g., age only for SCT, SD-O and SD-W, BVMT-R, ROCF, and CVLT), or age and education for other measures. Participants were asked to estimate their score as percentile equivalents, and were encouraged to guess even when they experienced uncertainty about how well they did compared with healthy peers. A standardized SA-C script was used to orient study participants to the SA-C task. The script instructed the examinee to rate themselves “compared to other people your age (and education) that took that same test.” Instructions included a general reminder concerning the normal distribution of test scores: “Many people get similar scores and would therefore be in the average range (examiner points to center of the bell curve graphic). Fewer people get scores that are much higher or much lower than average.” No other specific guidance or feedback was given during the course of testing that could be used to refine SA-C ratings for measures administered later in the evaluation. After this brief initial explanation, individual study participants typically were able to make SA-C ratings in a matter of seconds following completion of each of the neuropsychological test procedures.

Statistical Analyses

ANOVA was carried out by comparing the four subgroups on age and education, and ANCOVA was used in preliminary analyses to compare the subgroups with regard to raw neuropsychological test scores (adjusting for age and estimated baseline verbal IQ). Pearson’s correlations were computed to examine the strength of the association between actual test scores (converted to demographically adjusted percentile equivalents) and SA-C rankings (percentile equivalents). To lower the risk of Type II error, threshold for statistical significance was set at $p < .01$ for these and subsequent analyses. Variables showing the smallest correlations between SA-C and actual percentile rankings ($r < .35$) were thereby excluded from subsequent analyses.

Stepwise linear regression analyses were utilized to examine the strength of the the association between demographic variables (age and education), health/clinical factors (HIV, Alcohol, and BDI mood ratings), actual test performance, and SA-C self-ratings. All analyses were carried out using SPSS-12.0.

Results

Preliminary Analyses of Group Differences on Demographic and Clinical Variables and Neuropsychological Test Performance

Participants in the HIV–/LD subgroup had more education than each of the other groups ($p < .05$). Similar group differences were also observed in Verbal IQ. There was a trend for the HIV+ study participants to be older than other study participants ($p = .055$). With regard to neuropsychological test performance, the 199 study participants completing the SA-C ratings did not differ from the remaining research sample described in our previous publication (Rothlind et al., 2005; $p > .05$). See Fig. 2 for an overall summary of neuropsychological test scores for the combined sample for which self-appraisal ratings were obtained.

Similar to the pattern observed in our previously published study (Rothlind et al., 2005), after adjusting for age and baseline verbal IQ (estimated on basis of education and AMNART score), only a few group differences in neuropsychological test performance remained statistically significant. The HIV+/HD subgroup showed the greatest number of statistically significant group differences in comparison with the remaining subgroups (Rothlind et al., 2005). HIV+ serology was associated with lower scores on the CVLT, Trail Making B, and the Symbol Digit Modalities test ($p < .05$). Study participants consuming alcohol heavily earned lower scores than the light drinking group on the Trail Making Test A, Symbol Digit-Written, and Non-dominant Grooved Pegboard ($p < .05$).

Correlations Between SA-C Rankings and Test Performance

As noted earlier, group differences were generally small or non-significant after adjusting for baseline demographic differences. Moreover, the primary goal of this study was not to focus on the influence of the grouping variables on the dependent measure (SA-C), but rather to examine the more general pattern of relationships between actual performance and self-appraisal. To that end, the main statistical analyses (i.e., correlations and regression modeling) were initially applied to the combined sample.

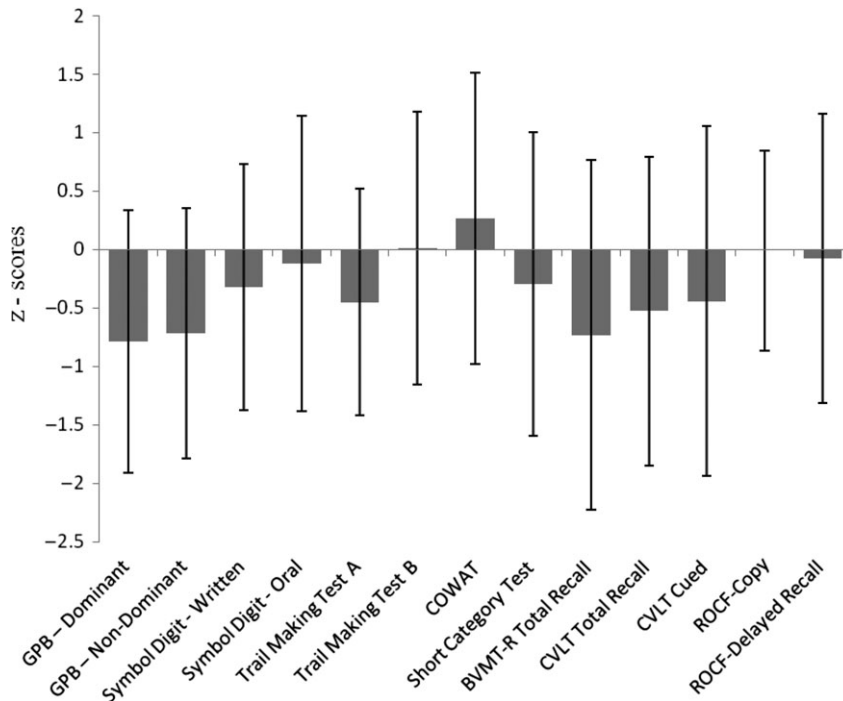


Fig. 2. Demographically adjusted test scores (means, SD) for the total sample ($N = 199$).

Notes: GPB = Grooved Pegboard; COWAT = Controlled Oral Word Association Test (CFL); BVMT-R = Brief Visuospatial Memory Test-Revised; CVLT = California Verbal Learning Test; ROCF = Rey–Osterrieth Complex Figure.

Table 1. Correlation coefficients (Pearson’s r) between test performance and self-appraisal (percentile equivalents) for the combined sample ($n = 199$)

Self-appraisal											
Performance	GPB-ND	SD-W	SD-O	TMT-A	TMT-B	COWAT	SCT	BVMT-R	CVLT-Total	CVLT-Cued	ROCF-D
GPB-ND	.24	ns	.17	ns	ns	ns	ns	ns	ns	ns	ns
SD-W	ns	.28	.27	ns	.23	ns	ns	.25	.23	ns	.22
SD-O	ns	.27	.40	ns	.23	ns	.21	.32	.26	.26	.30
TMT-A	ns	ns	ns	ns	ns	ns	ns	.25	ns	ns	.18
TMT-B	ns	.26	.27	ns	.37	ns	.30	.25	.26	.24	ns
COWAT	ns	.19	.23	ns	.21	.40	ns	ns	ns	ns	ns
SCT	ns	.26	.23	ns	.21	ns	.57	.34	.20	.21	.26
BVMT-R	ns	.19	.23	ns	ns	ns	.26	.55	.31	.31	.36
CVLT-Total	ns	ns	ns	ns	ns	ns	.20	.28	.43	.42	.28
CVLT-Cued	ns	ns	ns	ns	.21	ns	.23	.22	.32	.46	.26
ROCF-D	ns	.23	.23	ns	ns	ns	.30	.42	.28	.31	.56

Note: Correlation coefficients reported if $p < .01$; highest correlations are in bold for each task. GPB-ND = Grooved Pegboard non-dominant hand; SD-W = Symbol Digit-Written; SD-O = Symbol Digit-Oral; TMT-A = Trail Making Test Part A; TMT-B = Trail Making Test Part B; COWAT = Controlled Oral Word Association Test; SCT = Short Category Test; BVMT-R = Brief Visuospatial Memory Test-Revised; CVLT = California Verbal Learning Test; ROCF-D = Rey–Osterrieth Complex Figure-Delayed Recall.

Statistically significant positive correlations were observed between SA-C percentile ranking estimates and actual percentile rankings on specific tests in the combined sample (see Table 1). Even without direct knowledge of the test norms, the performance of other study participants, or the performance of their peers on the specific tests in question, individuals who had better task performance also ranked their performance higher in comparison with healthy peers, and individuals who performed poorly evaluated their performance less favorably on average.

Performance on measures of visuospatial problem solving, executive functioning, and learning and memory showed the strongest correlations with SA-C. Smaller correlations were also noted between SA-C rankings and performance on some

tests of speeded information processing, including TMT Part B and SD-O. Correlations between self-appraisal and speed of performance on other simple timed tests (e.g., Trail Making Test A and Grooved Pegboard-Dominant) were small, and in several instances did not reach statistical significance. SA-C for Rey-O Copy trials was likewise not significant, likely a reflection of the ceiling effect on this task, with most study participants performing at close to the highest level (correlations not shown).

In general, SA-C rankings were conservative, suggesting that individuals may employ a representativeness heuristic in ranking their own performance, with low performers ranking themselves as more proficient in comparison with peers than was actually the case, but high performers also ranking themselves as less proficient than was actually true, so that in each instance, the SA-C estimate suggested that the examinee perceived their own performance as more like the average of their

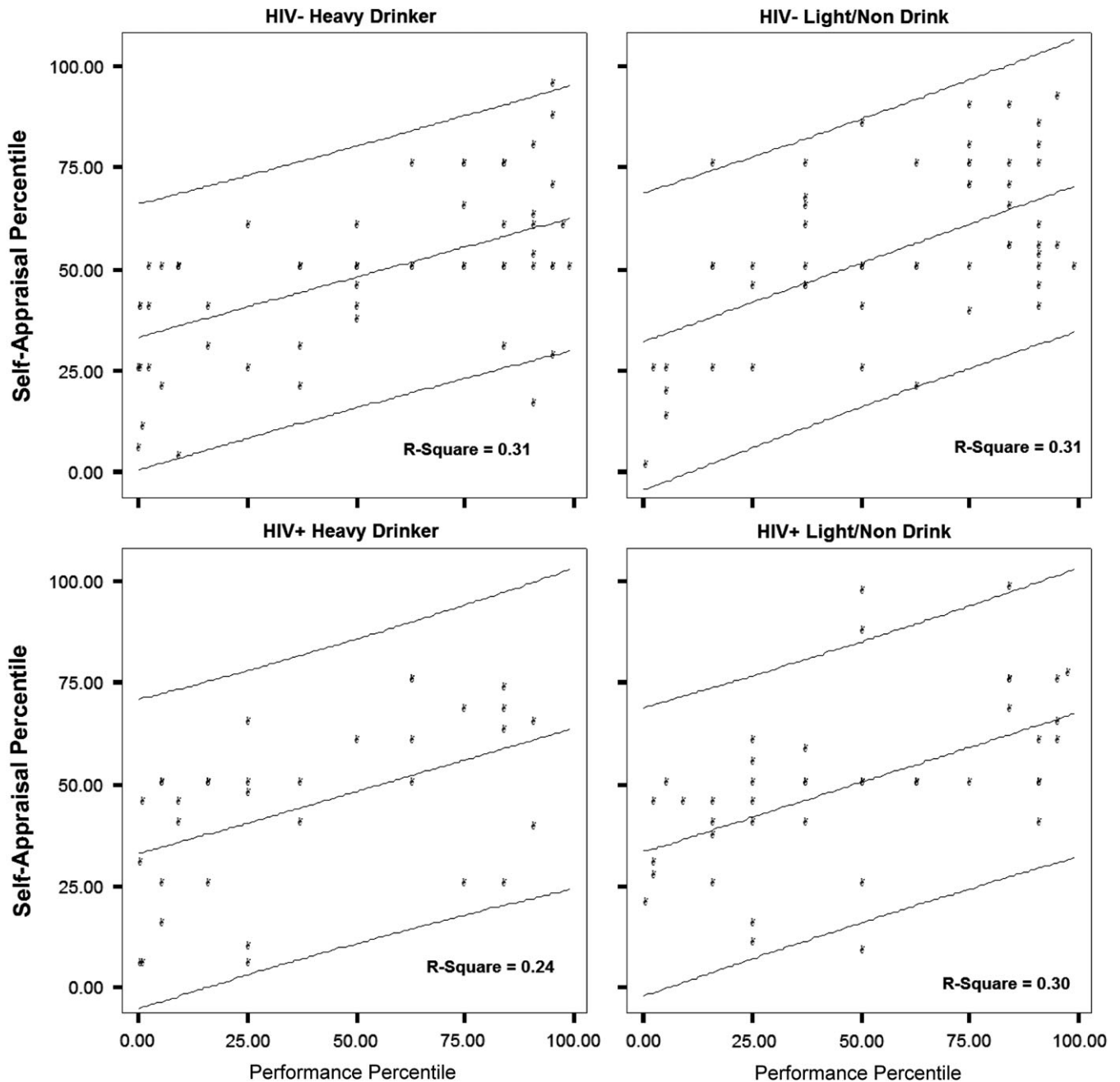


Fig. 3. Rey–Osterrieth Complex Figure-Delayed Reproduction test performance and self-appraisal (percentile equivalents) with regression line and 90% individual prediction intervals.

peers than was actually the case. The four panels shown in Fig. 3 document this trend for The Rey–Osterrieth Complex Figure-Delayed Recall for the four independent samples completing the study. Similar trends were noted across several other tests within the battery.

Demographic and Physical and Mental Health/Adjustment as Predictors of SA-C Ratings

To evaluate factors that may have contributed unique variance to SA-C rankings during neuropsychological testing in the present study, we also carried out separate, stepwise linear regression analyses for each of the neuropsychological tests. In each analysis, individual SA-C ranking for a specific test served as the dependent variables, and the actual test scores (percentile ranks) and demographic and physical and mental health variables were entered as predictors. Table 2 shows the results (best fitting model) for each of these regression analyses. As expected, performance on the specific (common) task under consideration generally showed the largest association with SA-C ranking. For a few of the SA-C rankings, particularly for tests where actual performance and SA-C did not correlate strongly, small proportions of additional variance were accounted for by the participants’ performance on other specific tests of executive function and memory retrieval. After adjusting for BVMT-R performance, a higher SA-C ranking estimate made for that test was associated with higher actual Rey-O Delay percentile rank (see Table 2).

Age was associated with higher appraisals of performance on the COWAT and on the Rey–Osterrieth Figure Copy trial, but with lower appraisals of CVLT-Cued Delayed Recall. However, in each instance, age accounted for only a small percentage (<3%) of the total variance in SA-C rankings. Education was associated with higher self-appraisal on the CVLT learning and Cued Delayed Recall trials, but like age, it also accounted for only a small percentage of the variance in SA-C for these tasks.

Neither HIV nor alcohol status accounted for unique variance in SA-C for any of the tests, with the exception of the Trail Making Test A. For the latter measure, the only one for which actual test performance was not a significant predictor of SA-C rank, HIV+ serology was associated with lower SA-C ranking (R^2 increment = .028, $p = .019$). More generally, the pattern of association between performance and SA-C ranking did not vary with group membership. Scores on the BDI did not contribute significantly to the prediction of specific SA-C judgments in the present study.

Table 2. Results of stepwise regression analysis for prediction of self-appraisal (SA-C)

Analysis	Self-awareness criterion	Predictor	β	SE	Std. β	R^2	$R^2 \Delta$	F Δ	Sig. F Δ
1	BVMT-R	BVMT-R	0.335	0.037	.544	.296	.296	80.23	<.001
		Rey-Delay	0.101	0.042	.17	.316	.021	5.72	.018
2	Rey-Delay	Rey-Delay	0.339	0.037	.550	.303	.303	82.92	<.001
3	CVLT-Total	CVLT-Total	0.233	0.037	.415	.172	.172	39.64	<.001
		Education	1.45	0.533	.183	.203	.031	7.36	.007
4	CVLT-Cued	CVLT-Cued	0.289	0.040	.471	.222	.222	52.99	<.001
		BVMT-R	0.137	0.044	.211	.260	.038	9.57	.002
		Age	−0.425	0.162	−.166	.287	.027	6.93	.009
		Education	1.18	0.582	.131	.302	.016	4.11	.044
5	SCT	SCT	0.370	0.040	.564	.318	.318	85.86	<.001
		CVLT-Cued	0.098	0.040	.149	.339	.021	5.91	.016
		Trail B	0.096	0.046	.133	.355	.016	4.42	.037
6	COWAT	COWAT	0.314	0.049	.421	.177	.177	40.90	<.001
		Age	0.312	0.155	.132	.194	.017	4.06	.045
7	Trail A	SD-Oral	0.112	0.042	.191	.036	.036	7.21	.008
		HIV	6.57	2.78	.171	.064	.028	5.61	.019
		Trail A	0.091	0.045	.126	.331	.015	4.12	.044
8	Trail B	Trail B	0.241	0.042	.381	.145	.145	32.51	<.001
		Trail A	−0.104	0.051	−.144	.164	.019	4.23	.041
9	SD-Oral	SD-Oral	0.232	0.037	.412	.170	.170	38.99	<.001
		AMNART	0.294	0.132	.159	.191	.021	4.98	.027
10	SD-Written	SD-Written	0.179	0.043	.291	.085	.085	17.67	<.001
		SCT	0.094	0.037	.180	.114	.029	6.28	.013
11	Rey-Copy	SikCT	0.180	0.044	.291	.085	.085	16.87	<.001
		Rey-Copy	0.162	0.070	.175	.111	.027	5.43	.021
		Age	0.411	0.180	.159	.136	.025	5.19	.024

Note: BVMT-R = Brief Visuospatial Memory Test, Revised, Trials 1–3; Rey-Delay = Rey–Osterrieth Complex Figure-Delayed Recall; CVLT = California Verbal Learning Test; SCT = Short Category Test; COWAT = Controlled Oral Word Association Test (CFL); SD-Oral = Symbol Digit-Oral; SD-Written = Symbol Digit-Written.

Discussion

The present study examined a “common-task, common-metric” (CM) approach for measuring SA-C in a non-clinical sample displaying broad variation in levels of neuropsychological test performance. We identified numerous standardized neuropsychological tests on which performance is significantly correlated with self-appraisal. The correlations between self-ratings and test performance were generally modest. The relationships suggest that individuals varying in cognitive ability but free of frank neurologic or major mood disorder show general awareness of their level of competence in performing commonly used standardized neuropsychological tests. This held true despite the novelty of the testing situation and the lack of information directly available to study participants about the performance of demographically comparable peers on the tests.

Tests that showed highest correlations with SA included those with clear parameters for task success (e.g., Rey-O Delay), and repeated exposure to the desired response over time, either through repetition of the standard for successful performance (e.g., repeated learning trials for CVLT and BVMT-R), or via corrective feedback after each response during the test (e.g., SCT). The correlations between SA accuracy and performance were smaller for tasks on which individuals differ primarily in terms of speed rather than in accuracy and where self-appraisals with regard to accuracy could confound ratings with regard to speed (e.g., Trail A and Grooved Pegboard). Although instructions regarding what parameters to focus on in making SA-C rank estimates were tailored for each test to mirror variables used to derive the normatively adjusted score, it is possible that the retrospective nature of these appraisals may have led some examinees to focus more on the overall success in completing the task rather than the speed (efficiency) of performance, resulting in SA-C ranking “noise” that lowered the correlations for these measures.

Although self-monitoring of performance on specific tests appears to be the main predictor of SA-C rankings for those tests, for a few SA-C rankings performance on other tests drawing upon similar cognitive functions also made a small contribution to predictions (e.g., small incremental prediction of BMVT performance SA deriving from actual success in performing Rey-O delayed recall). This suggests that study participants may have made some reference to recent subjectively experienced success or failure on tasks assessing similar domains (Rey-O was administered before the BVMT in the present study), or perhaps relied on more global self-perceptions (e.g., about relative strengths in verbal vs. visual processing modalities). More generally, an examination of the correlations between performance and rankings obtained at different times during the evaluation session did not suggest that the order of test administration accounts for differences in degrees of association between test performance and self-appraisal.

Demographic factors and health and adjustment variables played a very limited role in SA-C after adjusting for test performance. Age and education appeared to make only small contributions to the prediction of SA-C in the CM paradigm utilized in this study. Individuals with higher education tended to rank themselves more highly on a challenging test of list learning. General classification categories used to group individuals according to higher or lower “risk” for neurocognitive morbidity (e.g., HIV sero-status and alcohol consumption pattern) likewise accounted for very little unique variance, and the similarity of the correlations across the four independent study samples suggests that the variables mediating self-perception supersede these variables. Similarly, mood ratings did not correlate with SA-C in the present study.

Exclusion criteria limited participation in the current study to those without frank dementia, more severe TBI, or major mood disorders. Similar research with clinical samples exhibiting a wider range of cognitive impairments, and wider range of mood or other psychiatric disturbance will be valuable to further contrast trends in the high-functioning community sample of the present study with patterns seen in populations experiencing functional impairment and/or reduced insight.

With regard to the potentially biasing effects of mood disorder, some studies suggest that depression contributes to a more negative self-appraisal bias (e.g., Blackstone et al., 2012; Chin et al., 2014). However, there are models of the cognitive bias in depression which suggest that this condition may be characterized by a failure to engage in positive appraisal biases common to most individuals (so-called depressive realism, reviewed by Moore and Fresco, 2012). The literature summarized by these authors supports the possibility that individuals with depression may actually be more accurate in self-ratings in situations where they perform below the mean of healthy peers. We are not aware of any study that has utilized a common-metric approach to investigating this question in persons with more prominent mood disturbance, and further study will be valuable in elucidating the role the CM approach may play in the evaluation of persons suffering from depression.

Limitations and Directions for Further Research

The present findings suggest a CM approach to the evaluation of SA-C that has promise as a brief, quantitative tool that can be deployed in research and clinical settings where standardized neuropsychological testing is carried out. The implementation of the CM approach following completion of standardized tests allows it to be used as a part of clinical neuropsychological evaluation and consultation without requiring the examiner to deviate from standard test instructions. The study

identified several commonly used psychometric tests of learning, memory, and executive function on which examinees without neuropsychiatric morbidity show the capacity to estimate their level of functioning compared with peers across the spectrum of ability levels.

We utilized the stepwise approach to regression modeling because we wanted to explore the possibility that variables other than actual test performance accounting for significant variance in self-appraisal, even though we did not have strong hypotheses and are not aware of literature to suggest that this would be the case within the parameters of the current study. We recognize that the use of stepwise regression modeling procedures risks capitalizing on chance fluctuations in the data. However, for every one of the eleven SA-C values analyzed, the associated cognitive test score (or score on another related task) was the strongest predictor. The demographic and other variables had previously been shown to have only modest correlation with neurocognitive performance (Rothlind et al., 2005), and they therefore had the potential to account for unique variance in SA rankings even if entered later into the model. However, in contrast to test performance, which accounted for as much as 31% of the variance in SA rankings (SCT), the demographic and other variables accounted for less than 5% of the variance when they did enter into the models. These findings suggest that the selection of stepwise regression is not likely to account for the main findings of the study.

The current findings suggest that examinees may have less appreciation of their performance in comparison with peers for simple timed measures of motor and processing speed. Response time receives heavy emphasis in neuropsychological measurement, particularly when assessing individuals with certain disorders thought to involve frontal-subcortical dysfunction (e.g., Parkinson's disease, Huntington's disease, and some cases of cerebrovascular dementia). Thus, it is noteworthy that self-ratings showed the least correlation for tasks focused on speed and efficiency in test performance. It may be that a different instructional set that more explicitly encourages examinee to focus on this aspect of their own performance will shed additional light on individual differences in SA with regard to timing and efficiency of performance. This could have implications for research on self-awareness in the above clinical disorders. Alternatively, the findings may point to a particularly important role of neuropsychologists in providing valid information on speed of simple mental operations in relation to diagnostic and other clinical questions. The association between self-ratings and actual performance on other commonly used neuropsychological tests warrants further investigation.

Current findings suggest that it may be possible to establish cut-points defining when individuals with neurologic or neuropsychiatric illness display poor SA-C on specific measures. However, wide variability is noted in self-appraisal rankings made by participants in the present study, and a brief review suggests that only those self-ratings that are highly discrepant from actual levels of performance will clearly exceed cut-points established for any specific test based on a conservative 90% confidence interval. For the few measures examined, cut-points for abnormal self-ranking were similar across the four independent samples in our study, but the findings also suggest that these thresholds will vary considerably across the spectrum of test performance. As an example, the current findings suggest that for the Delayed Recall trial of the Rey–Osterreith Figure, it would be very unusual for an individual who performs at the second percentile (a value commonly considered to suggest impaired functioning) to rank themselves in the upper quartile compared with healthy peers (see Fig. 3), and only scores of this discrepancy would be considered evidence of faulty self-appraisal within this domain for an individual with very low scores. Different thresholds would apply for persons scoring in the average range. Of note, a self-rating in the low average or average range would not be considered clear evidence of abnormal self-appraisal based on the current findings, as this is not an uncommon self-ranking for many individuals who perform at a low level. Examination of biases over multiple measures may prove to be more valuable in clinical practice, but further research is needed to establish cut-points for specific clinical groups and for scores falling into different levels on the normal distribution.

In previous research involving the first author (Williamson et al., 2010), frank deficits in SA-C have been documented in fronto-temporal dementia and Alzheimer's disease using our CM approach in conjunction with other validated neuropsychological tasks. This and other recent studies offer validation of the CM approach by showing that distortions in self-appraisal determined with this approach correlate with self-appraisal deficits documented using other well-established methods of assessing deficits in self-awareness, and with other important clinical variables, even after adjusting for performance (Krueger et al., 2011).

In the present study, both high and low performers tended to rank themselves as closer to the mean than was actually true using normative data as a reference. Further investigation is needed to explore potential differences in accuracy between those who perform at the highest levels and those who perform poorly. Whether individuals who perform unusually low are disproportionately biased in their SA-C is controversial, but the merit of the CM approach in neurologic populations would be documented more clearly if it is shown that inaccuracy in SA-C is not merely an artifact of performing differently from the norm in absolute terms.

Further research is also needed to compare the CM approach to other indicators of self-awareness of functioning, and to examine the potential relationships between SA-C rankings obtained using a CM approach and other specific indicators of psychosocial adjustment and everyday functioning (e.g., employability, caregiving requirements, independence, and safety) and to measures of self-concept and psychological adjustment obtained through traditional psychological assessment techniques.

In considering potential implications of SA-C for clinical work, it may be important for clinicians to note that although the CM approach could provide valuable data, qualitative factors may be equally important to consider during feedback and consultation. For example, a client may rank their performance accurately (highly) on a measure on which they have performed well, but appraise it very harshly (e.g., “That was horrible”). Conversely, there may be opportunities for valuable feedback when patients vastly over-estimate their own performance rank in important domains, or appear indifferent to very low performance despite general awareness of their low ranking in comparison with peers.

Finally, the current findings suggest that quantitative measurement of SA-C using a CM approach maybe of interest to researchers probing the neural substrates of meta-cognitive abilities (Kedia, Lindner, Mussweiler, Ihssen, & Linden, 2013). Recent research utilizing our CM approach in a mixed cohort of individuals suffering from dementia has implicated gray-matter volume in the right ventromedial prefrontal cortex in SA-C accuracy (Rosen et al., 2010). In other research, making SA-C judgments has been shown to activate midline structures comprising portions of the “default-mode” network (Prigatano, 2010), although different cortical areas appear to become involved depending on the kind of social comparisons being made (Kedia et al., 2013). Further research involving structural and functional neuroimaging methods may be valuable to explore how a CM approach to SA-C relates to activation of distributed cortical structures that comprise the default-mode network and other areas thought to be involved in semantic memory, theory of mind and other aspects of social cognition, and in judgments and decision making related to everyday social communication and problem-solving.

Funding

This research was supported by NIAAA PO1 11493 (M.W. Weiner, PI).

Conflict of Interest

None declared.

Acknowledgements

The authors are grateful to Dr Fred Loya for helpful comments on a previous draft of this manuscript. We are grateful to Dr Aaron Kaplan for assisting in the development of the visual stimuli used to anchor rankings during the study, as depicted in Fig. 1.

References

- Allen, C. C., & Ruff, R. M. (1990). Self-rating versus neuropsychological performance of moderate versus severe head-injured patients. *Brain Injury*, 4, 7–17.
- Anderson, S., & Tranel, D. (1989). Awareness of disease states following cerebral infarction, dementia, and head trauma: Standardized assessment. *The Clinical Neuropsychologist*, 3, 327–339.
- Beck, A. T., & Steer, R. A. (1987). *Beck Depression Inventory: Manual*. San Antonio, TX: Psychological Corp.
- Benedict, H. B. (1997). *Brief Visual Memory Test - Revised Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Benton, L., Hamsher, K., & Sivan, A. B. (1994). *Multilingual aphasia examination: Manual of Instructions*. Iowa City: AJA Associates Inc.
- Blackstone, K., Moore, D. J., Heaton, R. K., Franklin, D. R., Woods, P., Clifford, D. B. et al. (2012). Diagnosing symptomatic HIV-associated neurocognitive disorders: Self-report versus performance-based assessment of everyday functioning. *Journal of the International Neuropsychological Society*, 18, 79–88.
- Chin, J., Oh, K. J., Seo, S. W., & Na, D. L. (2014). Are depressive symptomatology and self-focused attention associated with subjective memory impairment in older adults? *International Psychogeriatrics*, 26, 573–580.
- Clare, L. (2004). Awareness in early-stage Alzheimer’s disease: A review of methods and evidence. *British Journal of Clinical Psychology*, 43, 177–196.
- Cosentino, S., Metcalfe, J., Cary, M. S., De Leon, J., Karlawish, J. (2011). Memory awareness influences everyday decision making capacity about medication management in Alzheimer’s disease. *International Journal of Alzheimer’s Disease*, 483897.
- Cotrell, V., & Wild, K. (1999). Longitudinal study of self-imposed driving restrictions and deficit awareness in patients with Alzheimer disease. *Alzheimer’s Disease and Associated Disorders*, 13, 151–156.
- Dalla Barba, G., Parlato, V., Iavarone, A., & Boller, F. (1995). Anosognosia, intrusions and ‘frontal’ functions in Alzheimer’s disease and depression. *Neuropsychologia*, 33, 247–259.
- Delis, D. C., Kramer, J., Ober, B., & Kaplan, E. (1987). *The California Verbal Learning Test: Administration and interpretation*. San Antonio, TX: Psychological Corporation.
- Denman, S. B. (1984). *Manual for the Denman Neuropsychology Memory Scale*. Charleston, SC: Privately published.
- Duke, L. M., Seltzer, B., Seltzer, J. E., & Vasterling, J. J. (2002). Cognitive components of deficit awareness in Alzheimer’s disease. *Neuropsychology*, 16, 359–369.

- Ecklund-Johnson, E., & Torres, I. (2005). Unawareness of deficits in Alzheimer's disease and other dementias: Operational definitions and empirical findings. *Neuropsychological Review, 15*, 147–166.
- Fischer, S., Trexler, L. E., & Gauggel, S. (2004). Awareness of activity limitations and prediction of performance in patients with brain injuries and orthopedic disorders. *Journal of the International Neuropsychological Society, 10*, 190–199.
- Fleming, J. M., Strong, J., & Ashton, R. (1996). Self-awareness of deficits in adults with traumatic brain injury: how best to measure? *Brain Injury, 10* (1), 1–15.
- Graham, D. P., Kunik, M. E., Doody, R., & Snow, A. L. (2005). Self-reported awareness of performance in dementia. *Cognitive Brain Research, 25*, 144–152.
- Grober, E., Sliwinski, M., & Korey, S. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology, 1*, 933–949.
- Heaton, R., Grant, I., & Matthews, C. (1991). *Comprehensive norms for an expanded Halstead-Reitan Battery*. Odessa, FL: Psychological Assessment Resources.
- Kedia, G., Lindner, M., Mussweiler, T., Ihssen, N., & Linden, D. E. J. (2013). Brain networks of social comparison. *NeuroReport, 2*, 259–264. (e-publication).
- Krueger, C. E., Rosen, H. J., Taylor, H. G., Espy, K. A., Schatz, J., Rey-Casserly, C., et al. (2011). Know thyself: Real world behavioral correlates of self-appraisal accuracy. *The Clinical Neuropsychologist, 25*, 741–756.
- LaBuda, J., & Lichtenberg, P. (1999). The role of cognition, depression, and awareness of deficit in predicting geriatric rehabilitation patients' IADL performance. *The Clinical Neuropsychologist, 13*, 258–267.
- Lam, C. S., MCMahon, B. T., Priddy, D. A., & Gehred-Schultz, A. (1988). Deficit awareness and treatment performance among traumatic head injury adults. *Brain Injury, 2*, 235–242.
- Lehrner, J., Moser, D., Klug, S., Gleiss, A. M., Auff, E., Dal-Bianco, P., et al. (2014). Subjective memory complaints, depressive symptoms and cognition in patients attending a memory outpatient clinic. *International Psychogeriatrics, 26*, 463–473.
- Lysaker, P., Bell, M., Milstein, R., Bryson, G., & Beam-Goulet, J. (1994). Insight and psychosocial treatment compliance in schizophrenia. *Psychiatry, 57*, 307–315.
- Matthews, C., & Klove, H. (1964). *Instruction manual for the Adult Neuropsychological Test Battery*. Madison: University of Wisconsin Medical School.
- McGlynn, S. M., & Kaszniak, A. W. (1991). When metacognition fails: Impaired awareness of deficit in Alzheimer's disease. *Journal of Cognitive Neuroscience, 3*, 183–187.
- McGlynn, S. M., & Schacter, D. L. (1989). Unawareness of deficits in neuropsychological syndromes. *Journal of Clinical and Experimental Neuropsychology, 11*, 143–205.
- Moore, M. T., & Fresco, D. M. (2012). Depressive realism: A meta-analytic review. *Clinical Psychology Review, 32*, 496–509.
- Osterrieth, P. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie, 30*, 206–256.
- Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychological Review, 15*, 105–130.
- Prigatano, G. P. (2005). Disturbances of self-awareness and rehabilitation of patients with traumatic brain injury: A 20-year perspective. *Journal of Head Trauma and Rehabilitation, 20*, 19–29.
- Prigatano, G. P. (2010). *The study of anosognosia*. New York: Oxford University Press.
- Reitan, R. M. (1958). Validity of the Trail Making test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*, 271–276.
- Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathietraumatique. *Archives of Psychology, 28*, 286–340.
- Rosen, H. J., Alcantar, O., Rothlind, J., Sturm, V., Kramer, J. H., Weiner, M., et al. (2010). Neuroanatomical correlates of cognitive self-appraisal in neurodegenerative disease. *Neuroimage, 49*, 3358–3364.
- Rothlind, J. C., Greenfield, T. M., Bruce, A. V., Meyerhoff, D. J., Flenniken, D. L., Lindgren, J. A., et al. (2005). Heavy alcohol consumption in individuals with HIV infection: Effects on neuropsychological performance. *Journal of the International Neuropsychological Society, 11*, 70–83.
- Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1996). Benton Controlled Oral Word Association Test: Reliability and updated norms. *Archives of Clinical Neuropsychology, 1*, 329–338.
- Schretlen, D. J., Testa, M., Winicki, J. M., Pearson, G. D., & Gordon, B. (2008). Frequency and bases of abnormal performance by healthy adults on neuropsychological testing. *Journal of the International Neuropsychological Society, 14*, 436–445.
- Seltzer, B., Vasterling, J. J., Yoder, J., & Thompson, K. A. (1997). Awareness of deficit in Alzheimer's disease: Relation to caregiver burden. *Gerontologist, 37*, 20–24.
- Sherer, M., Hart, T., Nick, T. G., Whyte, J., Thompson, R. N., & Yablon, S. (2003). Early impaired self-awareness after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 84*, 168–176.
- Smith, A. (1973). *Symbol Digit Modalities Test*. Los Angeles: Western Psychological Services.
- Smith, T. E., Hull, J., Goodman, M., Hedayat-Harris, A., Willson, D., Israel, L. M., et al. (1999). The relative influences of symptoms, insight, and neurocognition on social adjustment in schizophrenia and schizoaffective disorder. *Journal of Nervous and Mental Disease, 187*, 102–108.
- Vanderploeg, R. D., Belanger, H. G., Duchnick, J. D., & Curtiss, G. (2007). Awareness problems following moderate to severe traumatic brain injury: Prevalence, assessment methods, and injury correlates. *Journal of Rehabilitation Research and Development, 44*, 937–950.
- Vogel, A., Waldorff, F. B., & Waldermar, G. (2010). Impaired awareness of deficits and neuropsychiatric symptoms in early Alzheimer's disease: The Danish Alzheimer Intervention Study (DAISY). *Journal of Neuropsychiatry and Clinical Neuroscience, 22*, 93–99.
- Wetzel, L., & Boll, T. J. (1987). *Short Category Test, Booklet Format*. Los Angeles: Western Psychological Services.
- Williamson, C., Alcantar, O., Rothlind, J., Cahn-Weiner, D., Miller, B. L., & Rosen, H. J. (2010). Standardized measurement of self-awareness deficits in FTD and AD. *Journal of Neurology, Neurosurgery, and Psychiatry, 81*, 140–145.
- Zlatar, Z. Z., Moore, R. C., Palmer, B. W., Thompson, W. K., & Jeste, D. V. (2014). Cognitive complaints correlate with depression rather than concurrent objective cognitive impairment in the successful aging evaluation baseline sample. *Journal of Geriatric Psychiatry and Neurology, 27*, 181–187.