Journal of the International Neuropsychological Society (2017), **23**, 778–790.
Copyright © INS. Published by Cambridge University Press, 2017.
doi:10.1017/S1355617717001060

Neuropsychological Assessment: Past and Future

INS is approved by the American Psychological Association to sponsor Continuing Education for psychologists. INS maintains responsibility for this program and its content.

Kaitlin B. Casaletto,[1] AND Robert K. Heaton[2]

[1]Department of Neurology, University of California, San Francisco, California
[2]Department of Psychiatry, University of California, San Diego, California

(RECEIVED March 16, 2017; FINAL REVISION August 29, 2017; ACCEPTED September 1, 2017)

Abstract

Neuropsychological assessment tools are the staple of our field. The development of standardized metrics sensitive to brain-behavior relationships has shaped the neuropsychological questions we can ask, our understanding of discrete brain functions, and has informed the detection and treatment of neurological disorders. We identify key turning points and innovations in neuropsychological assessment over the past 40–50 years that highlight how the tools used in common practice today came to be. Also selected for emphasis are several exciting lines of research and novel approaches that are underway to further probe and characterize brain functions to enhance diagnostic and treatment outcomes. We provide a brief historical review of different clinical neuropsychological assessment approaches (Lurian, Flexible and Fixed Batteries, Boston Process Approach) and critical developments that have influenced their interpretation (normative standards, cultural considerations, longitudinal change, common metric batteries, and translational assessment constructs). Lastly, we discuss growing trends in assessment including technological advances, efforts to integrate neuropsychology across disciplines (e.g., primary care), and changes in neuropsychological assessment infrastructure. Neuropsychological assessment has undergone massive growth in the past several decades. Nonetheless, there remain many unanswered questions and future challenges to better support measurement tools and translate assessment findings into meaningful recommendations and treatments. As technology and our understanding of brain function advance, efforts to support infrastructure for both traditional and novel assessment approaches and integration of complementary brain assessment tools from other disciplines will be integral to inform brain health treatments and promote the growth of our field. (*JINS*, 2017, *23*, 778–790)

Keywords: Psychometrics, Cognitive testing, Cross-cultural, Computerized assessment, Normative standards

INTRODUCTION

The goal of this review is to highlight key developments and turning points in neuropsychological assessment over the past several decades. We will touch on a breadth of assessment-related topics to provide an overview of the evolution of neuropsychological evaluation, leading to consideration of current practices and areas for future research. One limitation is that the content here is selective and admittedly draws heavily from practices based in the United States; while a majority of published neuropsychological tests and test norms have been developed within the United States and primarily for English-speaking populations, such cultures actually contain a disproportionately small fraction of the world's population. Still, the accomplishments and progress of neuropsychological assessment to date are remarkable, and we hope to inspire continued growth through this review. By its nature, neuropsychology is a distinctly transdisciplinary service. Neuropsychologists are in the unique position to be able to interpret brain measurement across modalities in a clinically meaningful way (e.g., neuroimaging, proteomics, genetics, technology-based service delivery). Leveraging the systematic and foundational approaches of our field, neuropsychologists are encouraged to continue to expand their repertoire and impact by creatively integrating brain assessment techniques across biological, cognitive science, and technology-based platforms as the field continues to evolve and define itself.

CLINICAL NEUROPSYCHOLOGICAL ASSESSMENT

Although the roles of neuropsychologists have evolved over time, the primary purposes for clinical neuropsychological assessment have remained fairly constant, to: (1) detect neurological dysfunction and guide differential diagnosis, (2) characterize changes in cognitive strengths and weaknesses over time, and (3) guide recommendations regarding

Correspondence and reprint requests to: Kaitlin B. Casaletto, Memory and Aging Center, Department of Neurology, University of California, San Francisco, 675 Nelson Rising Lane, Suite 190, San Francisco, CA 94158. E-mail: kaitlin.casaletto@ucsf.edu

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

everyday life and treatment planning. With the advent of increasingly sensitive and multimodal neurologic biomarker data, neuropsychological assessment shifted from its original role in "finding the lesion" to in-depth characterization of the patterns arising from disruptions in brain-behavior relationships. Although perhaps taken for granted given the body of gold-standard measures available at this stage in the field, advancement of cognitive assessment tools that are (1) sensitive to brain–behavior relationships of interest, (2) developed alongside other brain measurement tools, and (3) accessible and feasible across diverse settings continue to be needed areas of study. Ultimately, the neuropsychological questions we ask are only limited by the tools with which we have to probe the brain.

NEUROPSYCHOLOGICAL ASSESSMENT APPROACHES: A BRIEF MODERN HISTORY

Lurian Approach

Alexander Luria's seminal work following World War II in Russia represents an important early time-point in neuropsychological assessment. Luria's innovative ideas that the brain underlies the ability to carry out goal-directed behaviors and is shaped by environmental and cultural contexts (e.g., language) led to his systematic characterization of functional brain systems (Luria, 1966). Although largely qualitative, Luria's primary goal was to describe the cerebral bases that support corresponding functional systems, emphasizing the importance of understanding the multiple components that may comprise even simple neurobehavioral functions (Luria, 1966). Luria stimulated appreciation of brain specialization and moved away from the one-size-fits-all diagnosis of brain disease; however, his techniques were highly flexible and nonstandardized, making it difficult to reliably reproduce across examiners and patients.

Recognizing this problem, his student, Anne-Lise Christensen, published a more structured version of the Lurian approach, combining both qualitative and quantitative aspects of his battery (Draper, 1976). Ultimately, Charles Golden further standardized and combined the works of both Luria and Christensen in his development of the Luria-Nebraska Neuropsychological Battery (LNNB; Golden, Purisch, & Hammeke, 1979), scaling down almost 2000 original measurement items to 269 items covering 14 scales (e.g., motor, rhythm, memory, intelligence; Purisch, 2001).

In the United States, other neuropsychological assessment standardizations were also developing. In the 1950s, Arthur Benton was among the first thought leaders to criticize neurology for the lack of validated tools to measure common neurological syndromes (e.g., aphasia, agnosia). Through his pursuit of systematic neurobehavioral assessment, Benton developed several individual measures that are still widely used today (e.g., Benton Visual Rention Test) and shaped the Iowa-Benton school of neuropsychology (Tranel, 2009). Additionally, through these test developments, Benton began raising awareness of the apparent import of demographic factors (e.g., age and education) on test performances.

Flexible Battery

One assessment style that naturally extended from clinicians' early instincts to describe observed behaviors was the flexible integration of standardized measures. In the pure flexible battery approach, neuropsychologists administer only those measures directly related to the patient's presenting symptoms (e.g., multiple memory measures to characterize memory symptoms). Consequently, the administered battery inherently changes with a patient's referral question, while the examiner pursues the aim of being both as efficient and sensitive as possible to the presenting problem (Benton, 1994; Lezak, 1976).

Fixed Battery Approach

Concurrently, an entirely distinct assessment approach was developing that used a common, relatively comprehensive set of measures administered to all patients regardless of presenting symptoms, neuromedical history, or clinically apparent syndrome. Although lengthier and requiring greater burden and expense, fixed batteries maintain a constant testing condition, allowing for direct comparison of measures across disparate neurological disorders. Given their highly standard nature, research regarding fixed batteries is more easily facilitated, has garnered more comprehensive normative data, better understanding of their psychometric properties, and a broader empirical base from which to draw interpretations.

Two major leaders in the history of neuropsychological assessment pioneered the fixed battery approach, Ward Halstead and Ralph Reitan, his student (Reitan & Davidson, 1974). As a physiological psychologist in the mid-twentieth century, Halstead espoused a philosophy of neuropsychological assessment that was akin to a series of scientific experiments. He believed that there was no science in the individual event, and emphasized the need to develop exact, systematic procedures with sufficient comparison cases to interpret an individual test score (Reitan, 1994). Reitan similarly held strong beliefs that empiricism should be central in neuropsychological assessment (Grant & Heaton, 2015).

Building off of Halstead's early work, Halstead and Reitan closely observed individual patients in clinic and at home, helping to shape and refine measures that captured the observed brain–behavior relationships that could be generalized more broadly. Backed by more than two decades of blind interpretation of patient results and studies of groups with diverse neurologic disorders, Reitan ultimately published the Halstead-Reitan Battery (HRB) in 1985. The goal of the HRB was to be a systematic, quantitative means to measure the presence, location, extent, and nature of neurological disease equitably across clinical syndromes (Reitan, 1985). With its replicable procedures and highly quantitative results, the HRB intended to shift the practice of neuropsychology from an "art to a science" (Reitan, 1985) and indeed has been one of the most researched assessment batteries in neuropsychology (Kreutzer, DeLuca, Caplan, 2011).

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

As part of the trend toward highly quantitative assessment in the 1970s, some early approaches attempted to develop computerized algorithms for neuropsychological interpretation: the Key Approach by Russell, Neuringer, and Goldstein (1970), Finkelstein's BRAIN (1977), and Adams' ability-based algorithm (Adams, 1975). An overarching goal of these automated programs was to leverage the predictive power of computers to more accurately localize, categorize, and potentially infer etiology of brain damage such that "a technician or clerk, without knowledge of neurology, neuropsychology, or psychometrics" could use these tools (Deuel, 1971, p. 95). Although forward-thinking, the accuracy of such automated systems was inconsistent. The programs reliably identified presence of brain injury, but inferences regarding localization and etiology were imprecise and led to the conclusion that contextual information (i.e., clinical history) was critical to neuropsychological interpretation (Adams, Kvale, & Keegan, 1984).

Boston Process Approach

Another important turning point in neuropsychological practices was the development of what is now referred to as the "Boston Process Approach" led by Edith Kaplan (Kaplan, 1988). Although the Boston approach can be used with either flexible or fixed batteries, it is most commonly associated with and perhaps conceptually suitable to more flexible evaluations. The Boston approach places emphasis on *how* the patient comes to an answer (e.g., types of errors committed) rather than reliance on a single objective score. In this approach, testing the limits of a patient's cognitive abilities to elicit behaviors that may not traditionally present during standardized testing is emphasized.

Through the influence of Kaplan's training, Dean Delis advanced the standardization of the Boston approach, along with Kaplan and Joel Kramer, into what have come to be some of the most widely used neuropsychological assessment tools today. The California Verbal Learning Test (CVLT, Delis, Kramer, Kaplan, & Ober, 1987; and CVLT-second edition, Delis, Kramer, Kaplan, & Ober, 2000) and the Wechsler Adult Intelligence Scale-Revised as a Neuropsychological Instrument (Kaplan, Fein, Morris, Delis, 1991), followed by the Delis Kaplan Executive Functions System (DKEFS; Delis, Kaplan, Kramer, 2001) were borne out of a need to provide statistical parameters quantifying cognitive strategies that deviated from expectations (e.g., CVLT primacy effects; DKEFS error analysis).

Flexible Evaluation Approach

To avoid misconceptions, the pure *flexible battery* differs from the *flexible evaluation* approach that is used by over three-quarters of contemporary clinical neuropsychologists (Larrabee, 2008). Current common practice represents aspects of both flexible and fixed approaches, such that a fairly standardized, fixed set of measures is given to most patients with some flexibility to add or subtract measures given the specific referral question (Bigler, 2007).

NORMATIVE STANDARDS

As these fundamental assessment tools developed, there also grew increasing recognition that nonneurological, premorbid factors were significantly impacting test scores. The concept that cultural environment shapes brain development dates back at least to Luria (Luria, 1966). Mounting evidence in the 1970-80s then quantified that >40% of the variance in many test scores was accounted for by a single demographic factor (e.g., age, education), while even greater amounts of normal variance could be adjusted for when combing demographic factors together (Heaton, Grant, Matthews, 1991). Yet, development of a systematic way to adjust for such premorbid factors lagged.

While initial normative efforts focused on the effects of age alone (e.g., Wechsler measures), Robert Heaton, one of Reitan's students, recognized this gap and pioneered the development of normative standards adjusting for multiple background demographic factors. Publication of the "Heaton Norms" for the expanded HRB in 1991 signaled the most comprehensive standards at the time, adjusting for effects of age, sex, and education (Heaton et al., 1991). His subsequent revision in 2004 incorporated the effects of race, integrating the increasing literature demonstrating its influence on test performance (Heaton, Miller, Taylor, & Grant, 2004).

His advocacy to apply race as a proxy for important socio-cultural factors that impact brain development (e.g., education quality, socioeconomic status, medical care access, testing acculturation) represents a pivotal point in assessment approaches that at times is still received as controversial today (see detailed discussion by Manly, 2008). Heaton was not alone in addressing the need: in parallel, the Mayo Clinic, led by Robert Ivnik and Glenn Smith, was carrying out the Mayo Older Americans Normative Studies (MOANS), which aimed to improve the utility of commonly used neuropsychological tools for older adults (Ivnik et al., 1992a, 1992b). The MOANS project began publishing normative standards for various measures in 1992, and continued to update these standards through the early 2000s, including adjustments for racial minorities and careful consideration of what is considered neurologically normal in the elderly (e.g., Harris, Ivnik, & Smith, 2002; Lucas et al., 2005; Machulda et al., 2008; Smith, Wong, Ivnik, & Malec, 1997).

Other novel normative advancements were the development of co-normed batteries (e.g., WAIS-IV, WMS-IV, CVLT-II) and "robust norms" (Zhu & Tulsky, 2000). Co-normed batteries allow for test score interpretation on a common metric that is adjusted for the same demographic factors from the same normative cohort. Thus, they provide an equitable standard against which performances on all tests can be compared and interpreted. Robust normative standards simply refer to the use of as neurologically normal individuals as possible to provide standards of comparison. This may be particularly relevant in adult standards in which

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

underlying neurodegenerative pathology may accumulate decades before clinical manifestation (Holtzer et al., 2008). Inclusion of preclinical but asymptomatic individuals in a normative cohort may increase representativeness, but it also increases variability and decreases the sensitivity to acquired injury or disease. Nevertheless, there are many practical difficulties that limit feasibility in identifying highly neurologically normal individuals for inclusion in large-scale normative studies (e.g., expense in accessing brain scans and comprehensive family histories).

Despite major growth in normative development in the past two decades, many other interesting and potentially important questions about how to operationalize "normal" brain functioning remain unanswered. For example, do age effects on cognition differ across presumably "normal" groups based upon educational backgrounds, cognitive reserve, access to quality healthcare, nutrition, lifestyle factors, or environmental risks? Do education effects differ on the basis of how resourced the person's educational system has been, how education was valued in the home, and what other stresses in the environment may have interfered with educational pursuits?

For example, Jennifer Manly's innovative work parsing out contributing factors that account for racial differences on neuropsychological testing demonstrates that *quality of education* may be a critical underlying differentiator in these testing disparities (Manly et al., 2002). Other questions include, will gender effects on test performance differ based upon the society's traditional gender roles, gender-based opportunities and expectations? Also, if such effects of background factors change over time, how will this inform the need for updated normative standards, or perhaps even generation-based norms that reflect different effects of age, education, gender, and race/ethnicity? Answers to these questions now, and how they may change across time with shifting cultural values (e.g., gender/minority equality programs), may importantly impact our expectations and interpretation of neuropsychological test performances in the future.

CULTURAL AND INTERNATIONAL CONSIDERATIONS

Relatedly, during the past several decades, there has been a dramatic shift in the demographics of the United States, especially including individuals who speak languages other than English and who were born, grew up, and were educated (in part or totally) in cultures outside of the United States. Converging evidence demonstrates that test performance is significantly influenced by the values, customs, experiences, and cognitive styles that differ from the majority culture on which tests were originally developed and standardized (e.g., Arnold, Montgomery, Castaneda, & Longoria, 1994). Although their overlap and individual contributions are less clear, a host of culture-specific factors are undoubtedly associated with test scores across populations even within a single country (e.g., United States): acculturation, language, education quality and literacy, poverty and low socioeconomic status, familiarity with the evaluation process, and communication style (Fuiji, 2017).

As a rule of thumb, the larger the discrepancy between the individual being assessed and the majority culture in which the measure was developed, the higher the possibility that the test score may not reflect the construct it is posited to measure. Even when a measure is developed in the patient's native language and for examinees with the same racial/ethnic background, variability within a culture may still significantly influence test scores. For example, a recent publication demonstrated significant effects of frequency of everyday Spanish language use and location of education and birth among healthy U.S. Hispanic adults, even on a battery developed and normed in a national Hispanic sample (Flores et al., 2017).

These issues are further underscored and potentially amplified when neuropsychological measures developed in American culture are applied in non-U.S. contexts. The importance of the basic human ability being assessed may differ across cultures and be dependent on the values and everyday requirements of a given culture. For example, in the majority U.S. culture, speeded information processing is valued and educational experiences frequently reinforce that faster performances indicate better results. However, in Hispanic culture, speed and quality are oftentimes contradictory goals, such that slow and careful processing is thought to lead to the best results (Ardila, Rosselli, Matute, & Guajardo, 2005).

Some authors have suggested that what are needed in developing countries are tests that are more relevant to the backgrounds of those populations. However, it is unknown whether changes in the tests themselves would make them more valid in other settings, or whether all that is needed are different normative standards. Nonetheless, it may be that some tests simply are not useful across settings whereas others are, and we do not yet know what factors may drive these differences. Similarly, especially in the large and diverse developing world, it is unclear whether normative standards in one country or setting can be generalizable to other settings and, if they can, what common factors account for that generalizability. These are all questions to which neuropsychology researchers still need to seek answers. The need for valid neuropsychological practice and research is worldwide and it is simply not feasible to develop population-specific test norms, or new tests for every population on earth.

Currently, measures developed in the United States and Europe are commonly applied in other cultural contexts, although there are several notable efforts for culture-specific test and normative developments. As one example, Egypt has been at the forefront developing standardized Arabic evaluations, beginning as early as the 1920s (led by El-Kabbani) and extending into commonly used batteries developed in other cultures (e.g., Halstead-Retain and Luria-Nebraska; Al-Joudi, 2015). A recent review by Fasfous, Al-Joudi, Puente, and Perez-Garcia (2017) identified that,

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

although 117 individual neuropsychological measures were used in studies examining Arabic populations, only 53 were normed in the culture and evidenced appropriate cultural adaptation and validation per standard guidelines. Notably, almost all of the identified measures were tools initially developed in the United States and Europe: Verbal Fluency, Wechsler Memory and Intelligence batteries, Trail Making Test, Wisconsin Card Sorting Test, Ravens Matrices, and Bayley Scale of Infant Development-second edition were among the most extensively used cognitive measures in contemporary Arabic countries (Fasfous et al., 2017).

Other efforts to develop normative standards outside of American and European cultures suggest comparability of the construct measured when normative standardizations and underlying issues of cultural validity are appropriately addressed (e.g., *Zambia*, Hestad et al., 2016, Kabula et al., 2017; *China*, Gupta et al., 2014, Heaton et al., 2008, Shi et al., 2015; *Brazil*, de Almeida, et al., 2013; *India*, Ghate et al., 2015, Kamat et al., 2012, 2017, Malda, van de Vijver, Srinivasan, Transler, & Sukumar, 2010; *Czech Republic*, Bezdicek et al., 2012, 2014; *South Africa*, Nell, 1999; Nell, Myers, Colvin, & Rees, 1994; *Cameroon*, Ruffieux, et al., 2010; *South Korea*, Ko, Rosen, Simpson, & Brown, 2014), although there are some exceptions (e.g., lexical fluency and block design in Cameroon, Ruffieux, et al., 2010).

A common theme in reports of these cross-cultural test adaptations is the nonequivalence of the testing experience from Western to non-Western environments (e.g., Cassitto, Camerino, Hanninen, & Anger, 1990; Valciukas, Levin, Nicholson, & Selikoff, 1986). A host of factors, including repetition of test instructions, in-depth orientation to the testing environment, performance motivation, and profound effects of vocation and socioeconomic status have been identified as potentially contributing to test performance (Weinstein, Fucetola, & Mollica, 2001). As the world becomes increasingly interconnected and the need to assess and treat culturally different individuals grows, our field and its journals have an ethical responsibility to increase their priorities for such cross-cultural research endeavors.

SCIENTIFIC ADVANCEMENTS IN NEUROPSYCHOLOGICAL ASSESSMENT

Techniques for Defining Longitudinal Change

Neuropsychological batteries were originally constructed with the goal of identifying brain dysfunction, but a prominent emerging role of the neuropsychologist is to monitor syndrome progression or recovery *via* repeated evaluations. This is particularly true in the rehabilitation setting. As such, quantifying what constitutes significant change on a test battery is another relatively recent advance in our empirical assessment approach. Among the first and still widely used methods is the Reliable Change Index (RCI; Jacobson & Truax, 1991). The RCI was developed by Jacobson and Truax from the personality literature to establish whether the difference between an initial and follow-up test score exceeds what can be attributed to chance variation *versus* a clinically significant change. The RCI is based upon the reliability of the individual measure, which can be used to estimate the standard error of the difference between scores on the test in question.

Several years later, Chelune, Naugle, Luders, Selak, and Awad (1993) developed change modeling that additionally adjusted for practice effects, and McSweeney, Naugle, Chelune, and Luders (1993) proposed a linear multivariate regression approach adjusting for demographic effects and baseline performances to additionally control for regression to the mean. In 1996, similar regression models developed by Sawrie et al. and Hermann et al. included test–retest intervals (Standardized Regression Based norms) to quantify cognitive changes following epilepsy surgeries (Hermann et al., 1996; Sawrie, Chelune, Naugle, & Luders, 1996). These methods along with more complex multiple regression-based modeling including baseline scores, other potentially contributing factors (e.g., retest interval, demographics, nonlinear effects), and their interactions were then directly compared by Temkin, Heaton, and colleagues in neurologically normal and clinical groups (Heaton et al., 2001; Temkin, Heaton, Grant, & Dikmen, 1999).

Of interest, the authors found that the simple RCI was the least accurate method while the other approaches appeared comparably accurate. However, regardless of approach, baseline performance on individual tests had the largest impact on change classification, especially in clinical cohorts, and should be accounted for when developing change modeling (see Cysique et al., 2011 for further support and importance of "neuropsychological competence" predicting longitudinal cognitive change).

These studies laid the groundwork for the Serial Assessment report from the Wechsler Advanced Clinical Solutions software and the development of other calculators and published regression equations aiding in our quantification of cognitive "impairment" and change scores. For example, Crawford and colleagues leveraged existing datasets to develop enriched regression-based predictions of cognitive functioning for the individual case, which may also be applied longitudinally (Crawford & Garthwaite, 2007; Crawford, Garthwaite, Denham, & Chelune, 2012). Such programs that use the multitude of existing datasets to answer ongoing neuropsychological questions are innovative and will help optimize the growth of our field.

Common Metric Assessment Batteries

In an effort to leverage data and combine scientific efforts, development of standardized methods to evaluate cognition comparably across studies has been a major recent goal of the U.S. National Institutes of Health (NIH). Beyond answering innumerable scientific questions with greater power, such tools are particularly well-positioned for multisite epidemiological studies and clinical trial monitoring. With these goals in mind, several novel common metric batteries were

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

developed that are population-specific (Measurement and Treatment Research to Improve Cognition in Schizophrenia, MATRICS; Nuechterlein et al., 2008), construct-specific (Executive Abilities: Measure and Instruments for Neurobehavioral Evaluation and Research, EXAMINER; Kramer et al., 2014), and address broader, lifespan neurobehavioral assessment questions (NIH Toolbox for Assessment of Neurological and Behavioral Function; Gershon et al., 2013).

While these batteries represent massive accomplishments in the standardized assessment of neurologic function, and potential turning points in our resulting future understanding of brain function, they still maintain the common pitfalls of our traditional set of tools. Namely, although large infrastructure and broad expertise were dedicated to the theoretical and empirical development of these measures, efforts toward gathering sufficient normative information across populations, languages, and especially minority cultures, is ongoing and difficult, and generalizability of results with each battery remains to be demonstrated.

Novel and Translational Constructs

Although the traditional set of cognitive domains are an important foundation in the neuropsychological toolset (e.g., episodic memory, language, visuospatial skills), several domains drawing on the cognitive psychology and neurosciences literatures are highly complementary and warrant mention. These constructs have significantly furthered our capture of brain–behavior relationships and their assessments have demonstrated strong psychometric properties. Application of such measures depends entirely on the investigator or clinician's aims (e.g., if aiming to measure interpersonal function, social cognition is highly appropriate), although we encourage readers to consider such constructs for inclusion in their neuropsychological toolkit. We note only several exemplars here and encourage reading across related fields for further inspiration (e.g., Carter & Barch, 2007; Merikle, Smilek, & Eastwood, 2001).

Action Fluency

Although not a novel cognitive construct, action fluency (ability to generate verbs, for example, "things that people do") was developed as a counterpart to the traditional lexical and semantic fluency tasks. Action fluency taps into a more selective frontal-striatal network compared with the more temporal networks involved in noun generation (Piatt, Fields, Paolo, & Troster, 1999). Clinical lesion studies demonstrate a double-dissociation such that patients with frontal *versus* temporal lobe injury demonstrate disproportionate difficulties generating verbs than nouns, and *vice versa* (Damasio & Tranel, 1993). Similarly, more recent work in Parkinson's and HIV diseases (diseases with known selective involvement of frontal-striatal networks) show disproportionate disparities on verb compared to lexical or semantic generation tasks that is importantly predictive of daily functioning outcomes (Piatt, Fields, Paolo, Koller, & Troster, 1999; Woods, Carey, Troster, Grant, & Centre, 2005; Woods, Scott, et al., 2005; Woods et al., 2006).

Prospective Memory

Prospective memory (PM) is the ability to remember an intention in pursuit of a future goal. PM necessitates both executive (e.g., planning, monitoring, set-shifting) and episodic memory (e.g., recalled intention) abilities, drawing on frontal-temporal systems and their related networks (Ellis & Kvavilashvili, 2000). Although initial studies date back to Elizabeth Loftus in 1971, the past decade has seen a surge of renewed interest in the neuropsychological study of PM (Raskin, 2004). Prospective remembering declines across a range of neurologic conditions (e.g., HIV, Carey et al., 2006; schizophrenia, Twamley, Woods, Dawson, Narvaez, & Jeste, 2007; Parkinson's disease, Kliegel, Altgassen, Hering, & Rose, 2011; even normal aging, Einstein & Mcdaniel, 1990) and is consistently associated with declines in everyday functioning, even more so than traditional episodic memory measures (Einstein & Mcdaniel, 1990; Kliegel et al., 2011; Woods et al., 2007, 2008, 2009; Zogg, Woods, Sauceda, Wiebe, & Simoni, 2012).

Social Cognition

Social cognition was originally identified through the social psychology literature and is a multifaceted construct referring to the ability to perceive, interpret, and generate responses to the intentions and behaviors of others (see McDonald, 2017, for a comprehensive review; McDonald, Flanagan, & Rollins, 2011; McDonald, Flanagan, Rollins, & Kinch, 2003). Examples of social cognition tasks include identifying emotions, theory of mind (ability to understand and attribute others' mental states), and knowledge of social pragmatics (Green, Olivier, Crawley, Penn, & Silverstein, 2005).

Although its early and most in-depth neuropsychological study is largely in the context of schizophrenia-related adaptive functioning (e.g., NIMH MATRICS), there has been an explosion of studies in the last several years recognizing the primary social cognition impairment across neurologic conditions (e.g., behavioral variant of frontotemporal dementia, Shany-Ur & Rankin, 2011; Autism spectrum disorder, Pinkham, Hopfinger, Pelphrey, Piven, & Penn, 2008; Alzheimer's disease, Bediou et al., 2009; substance use disorders, Homer et al., 2008). Importantly, social cognitive measures are more strongly associated with community functioning outcomes than traditionally assessed constructs (Pijnenborg et al., 2009).

Incorporating social cognition measures into the standard neuropsychological toolbox may, therefore, afford highly relevant information for guiding daily functioning recommendations. Several examples of well-validated measures that neuropsychologists may consider include the NIMH MATRICS Social Cognition subdomain (Green et al., 2004), Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998), Ekman and Friesen Battery of Emotion Processing

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

(Ekman & Friesen, 1976), and Benton Facial Recognition Test (Benton, Hamsher, Varney, Spreen, 1983).

Everyday Functioning

A common complaint regarding neuropsychological evaluations is their apparent lack of relevance to the real-life problems that the patient may be experiencing. Expansion of the standardized neuropsychological toolset to include assessments of everyday abilities that are also sensitive to neurologic decline will enrich the ecological validity of evaluations. The Behavioral Assessment of the Dysexecutive Syndrome (BADS) was among the first batteries developed with this in mind (Wilson, Alderman, Burgess, Emslie, and Evans, 1996). For example, one of the BADS subtests, the Six Elements Test, is a widely used measure of cognitive multitasking that elicits planning and strategic thinking in a relatively unstructured context.

Other direct measures of daily living skills have grown in popularity in research settings, although their application in the clinic appears less common. A review by Moore, Palmer, Patterson, and Jeste (2007) identified 31 published performance-based measures of various functional skills, ranging from medication management and cooking to dressing and safety. We also encourage review of other comprehensive texts, such as self-report everyday measures as reviewed in Robyn Tate's *A compendium of Test, Scales, and Questionnaires* (2010) and *The Neuropsychology of Everyday Functioning* edited by Marcotte and Grant (2009), which provides in-depth performance-based measurement approaches of daily functioning across diseases (e.g., driving simulators).

Additionally, given the ubiquity of the Internet, tools assessing the ability to navigate the World Wide Web are highly relevant, easily accessible, and gaining momentum. For example, Woods and colleagues (2016) recently examined the validity of the Simulated Market Task (S-MarT) and Web-based Evaluation of Banking Skills (WEBS) as examiner-controlled Web sites of common household (i.e., shopping and banking) and health-related (i.e., pharmacy refills and healthcare communications) abilities. These tasks differentiated between HIV + individuals with and without mild cognitive impairment independent of previous Internet experience, and were moderately related to performances on standard neuropsychological measures (Woods et al., 2016; Woods, et al., in press).

Although the reliability, internal consistency, and concurrent validity of these real-world measures are consistently high, surprisingly few studies have directly examined their predictive validities. Given that this is one of their primary goals, future work supporting the ability of objective functional measures to relate to real world status is needed, although admittedly difficult to achieve (e.g., operationalization of real world status is inherently complex). In an effort to not simply become another neuropsychological test, performance-based instrumental activities of daily living measures must also balance ecological validity (e.g., less structured environment for task completion) and ability to reliably administer the task. Empirical translations of neuropsychological assessments that more precisely predict and guide patient recommendations for the real world continue to be an area of ongoing growth for the field.

TRENDS AND FUTURE DIRECTIONS

Technological Advancements: Computerized Neuropsychological Assessment

As technology becomes ever more accessible, the availability of computer- and tablet-based neuropsychological assessments has rapidly risen over the past decade. Recent U.S. NIH assessment initiatives investing in the development of computerized tools (e.g., NIH EXAMINER, Kramer et al., 2014; NIH Toolbox, Gershon et al., 2013) mirror and further highlight this trend. Although we are only able to briefly touch upon computerized assessment techniques here, we refer readers to the 2012 *American Academy of Clinical Neuropsychology* and *National Academy of Neuropsychology* position paper discussing relevant issues ranging from ethics and privacy to psychometrics, device marketing, and automated reporting services (Bauer et al., 2012).

Of historical relevance, the MicroCog Assessment of Cognitive Functioning in 1993 represents one of the first computerized cognitive assessment batteries commercially available *via* Pearson (see review, Elwood, 2001). Relatedly, the Automated Neuropsychological Assessment Metrics^{R} (ANAM) is unique in that it was initially developed by the U.S. Department of Defense and became commercially available through Vista Sciences at the University of Oklahoma (Reeves, Winter, Bleiberg, & Kane, 2007). There are now decades of clinical and laboratory data generated on the ANAM^{R}, including >300 peer-reviewed articles, detailing its development and use (for comprehensive review, see the McCaffrey (2007) special issue of *Archives of Clinical Neuropsychology*).

While several other relatively older computerized batteries (e.g., Cambridge Neuropsychological Test Automated Battery, Sahakian & Owen, 1992; Cutler et al., 1993; Lenehan, Summers, Saunders, Summers, & Vickers, 2016; and CogState, Fratti, Bowden, & Cook, 2016) have developed a large user-base and substantial validation data, novel assessment platforms continue to be developed. Online, tablet, and smart phone-based test batteries, wearable devices, as well as virtual reality-based cognitive assessment represent only some of the variety of emerging modalities (Brouliette et al., 2013; Parsey & Schmitter-Edgecombe, 2013; Parsons, 2015). Using these platforms, one innovative application is frequent, at-home serial cognitive assessments over multiple days or weeks. Such sequential ambulatory testing approaches have demonstrated strong reliability, construct validity with gold-standard measures, and enhanced ecologic validity (Sliwinski et al., 2016). Combining both traditional static, comprehensive testing with ambulatory, brief tests may represent a

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

complementary brain measurement approach that is more adept at detecting clinical changes at earlier stages of disease or injury, akin to other medical fields (e.g., cardiology one-time stress test *vs*. real-time Holter monitor).

Computers can do many things well and automatically, supporting their integration into neuropsychological practices (e.g., follow standardized procedures, time responses, and quickly and accurately score results). These assessment techniques have the potential to provide highly standardized and ecologically valid windows into real-time cognitive and behavioral changes, and could greatly improve accessibility (and costs) for underserved people who may have difficulty traveling to laboratory or healthcare settings. Yet, integration of such technologies is complex, not simply plug-and-play, and there are still major hurdles that must be addressed to ensure competent and valid application of online assessment (e.g., Internet connectivity, or examinee difficulties with comprehension of instructions, inconsistent effort, or distractions). A major consideration is that technology-based assessment tools inherently rely on continually changing technologies; while integration of the latest technology is appealing, any alterations to a validated measure will change its psychometric properties potentially rendering the new version as invalid or at the very least, highly difficult to interpret without a costly process of re-norming. If technology changes are unavoidable after test norming, careful piloting and documentation of any effects on test results are highly recommended and should occur before adopting or deploying the new technology. Technology will undoubtedly play a major role in the future of neuropsychological assessment by enhancing outreach, and potentially standardization and facilitation of diagnosis and early treatment. The development and implementation of such initiatives necessitates involvement of the neuropsychologist with our highly trained psychometric skillsets at the forefront.

Integration into Primary Care and Brain Health Assessment Initiatives

Although our understanding of the (oftentimes modifiable) factors associated with brain health has grown immensely in the past two decades, the systematic dissemination of this information to the public (including other medical fields) has lagged significantly behind. Emerging initiatives are underway to implement screening tools and visualize neurobehavioral data in real-world practices outside of the neuropsychologist's office. For example, University of California, San Francisco and Quest Diagnostics™ have partnered to develop the *Dementia Care Pathway*, a suite of technology-based tools to support navigation, brain health assessment (tablet-based cognitive screener, TabCAT; Possin et al., manuscript submitted for publication), automated analysis, and multimedia educational materials to nonspecialist providers (Rankin et al., manuscript submitted for publication).

Primary care is at the crux of health prevention for other body systems and is ideally positioned for inclusion of the brain. With the development of sensitive, standardized, easy-to-administer and portable screening devices, objective monitoring of brain status will be possible. Continued efforts to translate the important advances our field into real-world diagnostics and treatments are major areas for anticipated growth.

Neuropsychological Assessment Infrastructure

Another significant, under-addressed need in neuropsychology is normative infrastructure. Despite our progress, too often, we do not know how unusual or abnormal a test score or pattern of scores is for a specific group or individual. Of course, all tests must be interpreted in relation to "normal expectations" for the examinee, yet many tests do not have available norms that are appropriate for many types of examinees. Large-scale normative studies would of course help, but these are expensive and funding agencies typically are not supportive of grant funding for such work.

When a governmental funding agency (e.g., U.S. NIH) does fund norming of a nonproprietary set of cognitive tests (e.g., NIH Toolbox), this typically is a one-time effort and the norms may become out of date due to changes in a national census composition. By contrast, proprietary tests may get revised fairly often due to financial motives, and whatever may have been learned about associations between those tests and others (with their different norms) may be rendered tenuous or invalid. This issue is particularly evident in technology-based assessment development, as noted earlier. This of course is a moving target, but our field and its journals should prioritize work in this area, even if it does not have the cachet of a new, experimental approach to assessing a particular disease or brain system. At the very least, we recommend that new studies that use healthy control groups consider and report whether the groups' results are in the expected range on available norms.

CONCLUSIONS

Neuropsychology has deservedly prided itself as being more empirically based than other domains of professional psychology. But there is still much to learn. Many types of neuropsychological test interpretations are still performed by clinical judgment unaided by empirical guidelines, and many of the questions without empirical guidelines are actually empirical questions that have just not been asked and answered, to date. The evolution from patient observation to highly standardized measures largely relying on only a pencil and paper to obtain a reliable and objective window into brain functioning is resourceful and remarkable, particularly given the relative nascence of our field. As neuropsychology continues to grow, assessment advancements will center on increased ability to detect brain changes at the earliest possible points, likely in conjunction with multimodal markers of neurological function (e.g., neuroimaging and biofluid markers), and dissemination of our understanding of brain–behavior relationships to promote public brain health more broadly.

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

ACKNOWLEDGMENTS

This project was supported by a Larry L. Hillblom Fellowship Award (2017-A-004-FEL) and the HIV Neurobehavioral Research Center (NIMH P30MH062512). The authors have no conflicts of interest to disclose.

REFERENCES

Adams, K.M., Kvale, V.I., & Keegan, J.F. (1984). Relative Accuracy of 3 Automated Systems for Neuropsychological Interpretation. *Journal of Clinical Neuropsychology*, *6*(4), 413–431. doi: 10.1080/01688638408401232

Adams, K.M. (1975). *Automated clinical interpretation of the neuropsychological test battery: An ability based approach*. Detroit, MI: Wayne State University, University Microfilms.

Al-Joudi, H. (2015). Availability of Arabic language tests in the Middle East and North Africa. In J.H. University (Ed.), *INS NET*: JINS.

Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology*, *28*(1), 539–560. doi: 10.1207/s15326942dn2801_5

Arnold, B.R., Montgomery, G.T., Castaneda, I., & Longoria, R. (1994). Acculturation and performance of Hispanics on selected Halstead-Reitan neuropsychological tests. *Assessment*, *1*(3), 239–248.

Bauer, R.M., Iverson, G.L., Cernich, A.N., Binder, L.M., Ruff, R.M., & Naugle, R.I. (2012). Computerized neuropsychological assessment devices: Joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *The Clinical Neuropsychologist*, *26*(2), 177–196. doi: 10.1080/13854046.2012.663001

Bediou, B., Ryff, I., Mercier, B., Milliery, M., Henaff, M.A., D'Amato, T., … Krolak-Salmon, P. (2009). Impaired social cognition in mild Alzheimer disease. *Journal of Geriatric Psychiatry and Neurology*, *22*(2), 130–140. doi: 10.1177/0891988709332939

Benton, A.L. (1994). Neuropsychological assessment. *Annual Review of Psychology*, *45*, 1–23. doi: 10.1146/Annurev.Ps.45.020194.000245

Benton, A.L., Hamsher, K.S., Varney, N., & Spreen, O. (1983). *Contributions to neuropsychological assessment: A clinical manual*. Oxford, UK: Oxford University Press.

Bezdicek, O., Motak, L., Axelrod, B.N., Preiss, M., Nikolai, T., Vyhnalek, M., … Ruzicka, E. (2012). Czech version of the Trail Making Test: Normative data and clinical utility. *Archives of Clinical Neuropsychology*, *27*(8), 906–914. doi: 10.1093/arclin/acs084

Bezdicek, O., Stepankova, H., Motak, L., Axelrod, B.N., Woodard, J.L., Preiss, M., … Poreh, A. (2014). Czech version of Rey Auditory Verbal Learning test: Normative data. *Aging Neuropsychology and Cognition*, *21*(6), 693–721. doi: 10.1080/13825585.2013.865699

Bigler, E.D. (2007). A motion to exclude and the 'fixed' versus 'flexible' battery in 'forensic' neuropsychology: Challenges to the practice of clinical neuropsychology. *Archives of Clinical Neuropsychology*, *22*(1), 45–51. doi: 10.1016/j.acn.2006.06.019

Brouliette, R.M., Foil, H., Fontenot, S., Correro, A., Allen, R., Martin, C.K., … Keller, J.N. (2013). Feasibilty, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. *PLoS One*, *8*(6), e65925.

Carey, C.L., Woods, S.P., Rippeth, J.D., Heaton, R.K., Grant, I., & HIV Neurobehavioral Research Center (HNRC) Group. (2006). Prospective memory in HIV-1 infection. *Journal of Clinical and Experimental Neuropsychology*, *28*(4), 536–548.

Carter, C.S., & Barch, D.M. (2007). Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: The CNTRICS initiative. *Schizophrenia Bulletin*, *33*(5), 1131–1137.

Cassitto, M.G., Camerino, D., Hanninen, H., & Anger, W.K. (1990). International collaboration to evaluate the Who Neurobehavioral Core Test Battery. In B.L. Johnson, W.K. Anger, A. Durao & C. Xintaras (Eds.), *Advances in neurobehavioral toxicology: Applications in environmental and occupational health* (pp. 203–223). Chelsea, MI: Lewis.

Chelune, G.J., Naugle, R.I., Luders, H., Selak, J., & Awad, I.A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, *7*, 41–52.

Crawford, J.R., & Garthwaite, P.H. (2007). Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology*, *21*(5), 611–620. doi: 10.1037/0894-4105.21.5.611

Crawford, J.R., Garthwaite, P.H., Denham, A.K., & Chelune, G.J. (2012). Using regression equations built from summary data in the psychological assessment of the individual case: Extension to multiple regression. *Psychological Assessment*, *24*(4), 801–814. doi: 10.1037/a0027699

Cutler, N.R., Shrotriya, R.C., Sramek, J.J., Veroff, A.E., Seifert, R.D., Reich, L.A., & Hironaka, D.Y. (1993). The use of the Computerized Neuropsychological Test Battery (CNTB) in an efficacy and safety trial of BMY 21,502 in Alzheimer's disease. *Annals of the New York Academy of Sciences*, *695*, 332–336.

Cysique, L.A., Franklin, D. Jr., Abramson, I., Ellis, R.J., Letendre, S., Collier, A., … Simpson, D. (2011). Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. *Journal of Clinical and Experimental Neuropsychology*, *33*(5), 505–522.

Damasio, A.R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(11), 4957–4960. doi: 10.1073/Pnas.90.11.4957

de Almeida, S.M., Ribeiro, C.E., de Pereira, A.P., Badiee, J., Cherner, M., Smith, D., … Heaton, R.K. (2013). Neurocognitive impairment in HIV-1 clade C-versus B-infected individuals in Southern Brazil. *Journal of Neurovirology*, *19*(6), 550–556.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test: Adult version. Manual*. San Antonio, TX: Psychological Corportation.

Delis, D., Kramer, J.H., Kaplan, E., & Ober, B. (2000). *The California Verbal Learning Test–Second edition*. San Antonio, TX: The Psychological Corporation.

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis-Kaplan Executive Function System*. San Antonio, TX: The Psychological Corporation.

Draper, I.T. (1976). Luria's neuropsychological investigation. *Journal of Neurology, Neurosurgery, and Psychiatry*, *39*(4), 409–410.

Deuel, R.K. (1971). Assessment of brain damage: A neuropsychological key approach. *Archives of Neurology*, *25*(1), 95–95.

Einstein, G.O., & Mcdaniel, M.A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology. Learning Memory and Cognition*, *16*(4), 717–726. doi: 10.1037/0278-7393.16.4.717

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

Ekman, P., & Friesen, W.V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ellis, J., & Kvavilashvili, L. (2000). Prospective memory in 2000: Past, present, and future directions. *Applied Cognitive Psychology*, *14*, S1–S9. doi: 10.1002/Acp.767.Abs

Elwood, R.W. (2001). MicroCog: Assessment of cognitive functioning. *Neuropsychology Review*, *11*(2), 89–100. doi: 10.1023/A:1016671201211

Fasfous, A.F., Al-Joudi, H.F., Puente, A.E., & Perez-Garcia, M. (2017). Neuropsychological measures in the Arab World: A systematic review. *Neuropsychology Review*, *27*, 158–173. doi: 10.1007/s11065-017-9347-3

Finkelstein, J.N. (1977). *BRAIN: A computer program for interpretation of the Halstead-Reitan Neuropsychological Test Battery*. New York, NY: Columbia University, University Microfilms.

Flores, I., Casaletto, K.B., Marquine, M.J., Umlauf, A., Moore, D.J., Mungas, D., … Heaton, R.K. (2017). Performance of Hispanics and Non-Hispanic Whites on the NIH Toolbox Cognition Battery: The roles of ethnicity and language backgrounds. *The Clinical Neuropsychologist*, *31*, 783–797. doi: 10.1080/13854046.2016.1276216

Fratti, S., Bowden, S.C., & Cook, M.J. (2016). Reliability and validity of the CogState computerized battery in patients with seizure disorders and healthy young adults: Comparison with standard neuropsychological tests. *Clin Neuropsychol*, *31*, 569–586. doi: 10.1080/13854046.2016.1256435

Fuiji, D. (2017). *Conducting a culturally informed neuropsychological evaluation*. Washington, DC: American Psychological Association.

Gershon, R.C., Wagster, M.V., Hendrie, H.C., Fox, N.A., Cook, K.F., & Nowinski, C.J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology*, *80*, S2–S6. doi: 10.1212/WNL.0b013e3182872e5f

Ghate, M., Mehendale, S., Meyer, R., Umlauf, A., Deutsch, R., Kamat, R., … Alexander, T. (2015). The effects of antiretroviral treatment initiation on cognition in HIV-infected individuals with advanced disease in Pune, India. *Journal of Neurovirology*, *21*(4), 391–398.

Golden, C.J., Purisch, A.D., & Hammeke, T.A. (1979). *The Luria-Nebraska Neuropsychological Battery: A manual for clinical and experimental uses*. Lincoln, NE: University of Nebraska Press.

Grant, I., & Heaton, R.K. (2015). Ralph M. Reitan: A founding father of neuropsychology. *Archives of Clinical Neuropsychology*, *30*(8), 760–761. doi: 10.1093/arclin/acv077

Green, M.F., Nuechterlein, K.H., Gold, J.M., Barch, D.M., Cohen, J., Essock, S., … Marder, S.R. (2004). Approaching a consensus cognitive battery for clinical trials in schizophrenia: The NIMH-MATRICS conference to select cognitive domains and test criteria. *Biological Psychiatry*, *56*(5), 301–307. doi: 10.1016/j.biopsych.2004.06.023

Green, M.F., Olivier, B., Crawley, J.N., Penn, D.L., & Silverstein, S. (2005). Social cognition in schizophrenia: Recommendations from the measurement and treatment research to improve cognition in schizophrenia new approaches conference. *Schizophrenia Bulletin*, *31*(4), 882–887. doi: 10.1093/schbul/sbi049

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Gupta, S., Iudicello, J.E., Shi, C., Letendre, S., Knight, A., Li, J., … Atkinson, J.H. (2014). Absence of neurocognitive impairment in a large Chinese sample of HCV-infected injection drug users receiving methadone treatment. *Drug and Alcohol Dependence*, *137*, 29–35.

Harris, M.E., Ivnik, R.J., & Smith, G.E. (2002). Mayo's older americans normative studies: Expanded AVLT recognition trial norms for ages 57 to 98. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 214–220. doi: 10.1076/Jcen.24.2.214.995

Heaton, R.K., Grant, I., & Matthews, C.G. (1991). *Comprehensive norms for an expanded Halstead-Reitan battery: Demographic corrections, research findings, and clinical applications*. Odessa, FL: Psychological Assessment Resources.

Heaton, R.K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M.J., Marcotte, T.D., … Grant, I. (2001). Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology*, *16*(1), 75–91. doi: 10.1016/S0887-6177(99)00062-1

Heaton, R.K., Miller, S.W., Taylor, J.T., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults*. Lutz, FL: Psychological Assessment Resources, Inc.

Heaton, R.K., Cysique, L.A., Jin, H., Shi, C., Yu, X., Letendre, S., … Marcotte, T.D. (2008). Neurobehavioral effects of human immunodeficiency virus infection among former plasma donors in rural China. *Journal of Neurovirology*, *14*(6), 536–549.

Hermann, B.P., Seidenberg, M., Schoenfeld, J., Peterson, J., Leveroni, C., & Wyler, A.R. (1996). Empirical techniques for determining the reliability, magnitude, and pattern of neuropsychological change after epilepsy surgery. *Epilepsia*, *37*(10), 942–950.

Hestad, K.A., Menon, J.A., Serpell, R., Kalungwana, L., Mwaba, S.O., Kabuba, N., … Heaton, R.K. (2016). Do neuropsychological test norms from African Americans in the United States generalize to a Zambian population?. *Psychological Assessment*, *28*(1), 18.

Holtzer, R., Goldin, Y., Zimmerman, M., Katz, M., Buschke, H., & Lipton, R.B. (2008). Robust norms for selected neuropsychological tests in older adults. *Archives of Clinical Neuropsychology*, *23*(5), 531–541. doi: 10.1016/j.acn.2008.05.004

Homer, B.D., Solomon, T.M., Moeller, R.W., Mascia, A., DeRaleau, L., & Halkitis, P.N. (2008). Methamphetamine abuse and impairment of social functioning: A review of the underlying neurophysiological causes and Behavioral implications. *Psychological Bulletin*, *134*(2), 301–310. doi: 10.1037/0033-2909.134.2.301

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1992a). Mayo's Older Americans Normative Studies: WAIS-R norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6*(S1), 1–30.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1992b). Mayo's Older Americans Normative Studies: Updated AVLT norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6*(S1), 83–104.

Jacobson, N.S., & Truax, P. (1991). Clinical-significance - A statistical approach to defining meaningful change in psychotherapy-research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. doi: 10.1037//0022-006x.59.1.12

Kabuba, N., Menon, J.A., Franklin, D.R., Heaton, R.K., & Hestad, K.A. (2017). Use of Western neuropsychological test battery in detecting hiv-associated neurocognitive disorders (HAND) in Zambia. *AIDS and Behavior*, *21*(6), 1717–1727.

Kamat, R., Ghate, M., Gollan, T.H., Meyer, R., Vaida, F., Heaton, R.K., … Mehendale, S. (2012). Effects of Marathi-Hindi

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

bilingualism on neuropsychological performance. *Journal of the International Neuropsychological Society*, *18*(2), 305–313.

Kamat, R., McCutchan, A., Kumarasamy, N., Marcotte, T.D., Umlauf, A., Selvamuthu, P., … Bharti, A.R. (2017). Neurocognitive functioning among HIV-positive adults in southern India. *Journal of Neurovirology*. [Epub ahead of print].

Kaplan, E. (1988). The process approach to neuropsychological assessment. *Aphasiology*, *2*(3-4), 309–311. doi: 10.1080/02687038808248930

Kaplan, E., Fein, D., Morris, R., & Delis, D. (1991). *The WAIS-R as a neuropsychological instrument*. San Antonio, TX: Psychological Corporation.

Kliegel, M., Altgassen, M., Hering, A., & Rose, N.S. (2011). A process-model based approach to prospective memory impairment in Parkinson's disease. *Neuropsychologia*, *49*(8), 2166–2177. doi: 10.1016/j.neuropsychologia.2011.01.024

Ko, J., Rosen, A.B., Simpson, K.J., & Brown, C.N. (2014). Cross-cultural adaption and reliability of the Korean version of the identification of functional ankle instability. *Medicine and Science in Sports and Exercise*, *46*(5), 203–203.

Kramer, J.H., Mungas, D., Possin, K.L., Rankin, K.P., Boxer, A.L., Rosen, H.J., … Widmeyer, M. (2014). NIH EXAMINER: Conceptualization and development of an executive function battery. *Journal of the International Neuropsychological Society*, *20*(1), 11–19.

Kreutzer, J.S., DeLuca, J., & Caplan, B. (Eds) 2011). *Encyclopedia of Clinical Neuropsychology. Halstead-Reitan Neuropsychology Test Battery* (pp. 1201–1205). New York: Springer.

Larrabee, G.J. (2008). Flexible vs. fixed batteries in forensic neuropsychological assessment: Reply to Bigler and Hom. *Archives of Clinical Neuropsychology*, *23*(7-8), 763–776. doi: 10.1016/j.acn.2008.09.004

Lenehan, M.E., Summers, M.J., Saunders, N.L., Summers, J.J., & Vickers, J.C. (2016). Does the Cambridge Automated Neuropsychological Test Battery (CANTAB) distinguish between cognitive domains in healthy older adults? *Assessment*, *23*(2), 163–172. doi: 10.1177/1073191115581474

Lezak, M.D. (1976). *Neuropsychological assessment*. Oxford, England: Oxford University Press.

Loftus, E. (1971). Memory for intentions: The effect of presence of a cue and interpolated activity. *Psychonomic Science*, *23*, 315–316.

Lucas, J.A., Ivnik, R.J., Willis, F.B., Ferman, T.J., Smith, G.E., Parfitt, F.C., … Graff-Radford, N.R. (2005). Mayo's older African Americans normative studies: Normative data for commonly used clinical neuropsychological measures. *Clinical Neuropsychologist*, *19*(2), 162–183. doi: 10.1080/13854040590945265

Luria, A.R. (1966). *Higher cortical functions in man*. New York: Springer.

Machulda, M.M., Ivnik, R.J., Smith, G.E., Ferman, T.J., Boeve, B.F., Knopman, D., … Tangalos, E.G. (2008). Mayo's older Americans normative studies: Visual form discrimination and copy trial of the Rey-Osterrieth complex figure. *Journal of Clinical and Experimental Neuropsychology*, *29*(5), 377–384. doi: 10.1080/13803390701850817

Malda, M., van de Vijver, F.J.R., Srinivasan, K., Transler, C., & Sukumar, P. (2010). Traveling with cognitive tests: Testing the validity of a KABC-II adaptation in India. *Assessment*, *17*(1), 107–115. doi: 10.1177/1073191109341445

Manly, J.J. (2008). Critical issues in cultural neuropsychology: Profit from diversity. *Neuropsychology Review*, *18*(3), 179.

Manly, J.J., Jacobs, D.M., Touradji, P., Small, S.A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, *8*(3), 341–348.

Marcotte, T.D., & Grant, I. (Eds.) (2009). *Neuropsychology of everyday functioning*. New York: Guilford Press.

McCaffrey, R.J. (Ed.) (2007). Automated neuropsychological assessment metrics [Special issue]. *Archives of Clinical Neuropsychology*, *22S*, S1.

McDonald, S. (2017). Emotions are rising: The growing field of affect neuropsychology. *Journal of the International Neuropsychological Society*, *23*, 719–731.

McDonald, S., Flanagan, S., & Rollins, J. (2011). *The awareness of social inference test* (Revised). Sydney, Australia: Pearson Assessment.

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *18*, 219–238.

McSweeny, A.J., Naugle, R.I., Chelune, G.J., & Luders, H. (1993). "T scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist*, *7*(3), 300–312.

Merikle, P.M., Smilek, D., & Eastwood, J.D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition*, *79*(1), 115–134.

Moore, D.J., Palmer, B.W., Patterson, T.L., & Jeste, D.V. (2007). A review of performance-based measures of functional living skills. *Journal of Psychiatric Research*, *41*(1-2), 97–118. doi: 10.1016/j.psychires.2005.10.008

Nell, V. (1999). Luria in Uzbekistan: The vicissitudes of cross-cultural neuropsychology. *Neuropsychology Review*, *9*(1), 45–52. doi: 10.1023/A:1025643004782

Nell, V., Myers, J., Colvin, M., & Rees, D. (1994). Neuropsychological assessment of organic solvent effects in South-Africa - Test selection, adaptation, scoring, and validation issues. *Environmental Research*, *63*, 301–318.

Nuechterlein, K.H., Green, M.F., Kern, R.S., Baade, L.E., Barch, D.M., Cohen, J.D., … Marder, S.R. (2008). The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203–213. doi: Doi 10.1176/Appi.Ajp.2007.07010042

Parsey, C.M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *The Clinical Neuropsychologist*, *27*(8), 1328–1361. doi: 10.1080/13854046.2013.834971

Parsons, T.D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience*, *9*, 660. doi: Artn 66010.3389/Fnhum.2015.00660

Piatt, A.L., Fields, J.A., Paolo, A.M., Koller, W.C., & Troster, A.I. (1999). Lexical, semantic, and action verbal fluency in Parkinson's disease with and without dementia. *Journal of Clinical and Experimental Neuropsychology*, *21*(4), 435–443. doi: 10.1076/Jcen.21.4.435.885

Piatt, A.L., Fields, J.A., Paolo, A.M., & Troster, A.I. (1999). Action (verb naming) fluency as an executive function measure: Convergent and divergent evidence of validity. *Neuropsychologia*, *37*(13), 1499–1503. doi: 10.1016/S0028-3932(99)00066-4

Pijnenborg, G.H., Withaar, F.K., Evans, J.J., van den Bosch, R.J., Timmerman, M.E., & Brouwer, W.H. (2009). The predictive value of measures of social cognition for community functioning in schizophrenia: Implications for neuropsychological

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

assessment. *Journal of the International Neuropsychological Society*, *15*(2), 239–247. doi: 10.1017/S1355617709090341

Pinkham, A.E., Hopfinger, J.B., Pelphrey, K.A., Piven, J., & Penn, D.L. (2008). Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophrenia Research*, *99*(1-3), 164–175. doi: 10.1016/j.schres.2007.10.024

Purisch, A.D. (2001). Misconceptions about the Luria-Nebraska Neuropsychological Battery. *NeuroRehabilitation*, *16*(4), 275–280.

Raskin, S.A. (2004). *Memory for intentions screening test*. Paper presented at the Journal of the International Neuropsychological Society.

Reeves, D.L., Winter, K.P., Bleiberg, J., & Kane, R.L. (2007). ANAM Genogram: Historical perspectives, description and current endeavors. *Archives of Clinical Neuropsychology*, *22*(Suppl. 1), S15–S37.

Reitan, R.M. (1985). *Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tuscon, AZ: Neuropsychology Press.

Reitan, R.M. (1994). Ward Halstead's contributions to neuropsychology and the Halstead-Reitan Neuropsychological Test Battery. *Journal of Clinical Psychology*, *50*(1), 47–70. doi: 10.1002/1097-4679(199401)50:1 < 47::Aid-Jclp2270500106 > 3.0.Co;2-X

Reitan, R.M., & Davidson, L.D. (1974). *Clinical neuropsychology: Current status and applications*. Washington, DC: Winston.

Ruffieux, N., Njamnshi, A.K., Mayer, E., Sztajzel, R., Eta, S.C., Doh, R.F., … Hauert, C.A. (2010). Neuropsychology in Cameroon: First normative data for cognitive tests among school-aged children. *Child Neuropsychology*, *16*(1), 1–19. doi: 10.1080/09297040902802932

Russell, E.W., Neuringer, C., & Goldstein, G. (1970). *Assessment of brain damage: A neuropsychological key approach*. New York: Interscience.

Sahakian, B.J., & Owen, A.M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, *85*(7), 399–402.

Sawrie, S.M., Chelune, G.J., Naugle, R.I., & Luders, H.O. (1996). Empirical methods for assessing meaningful neuropsychological change following epilepsy surgery. *Journal of the International Neuropsychological Society*, *2*(6), 556–564.

Shany-Ur, T., & Rankin, K.P. (2011). Personality and social cognition in neurodegenerative disease. *Current Opinion in Neurology*, *24*(6), 550–555. doi: 10.1097/WCO.0b013e32834cd42a

Shi, C., Kang, L., Yao, S., Ma, Y., Li, T., Liang, Y., … Zhang, C. (2015). The MATRICS consensus cognitive battery (MCCB): Co-norming and standardization in China. *Schizophrenia Research*, *169*(1), 109–115.

Sliwinski, M.J., Mogle, J.A., Hyun, J., Munoz, E., Smyth, J.M., & Lipton, R.B. (2016). Reliability and validity of ambulatory cognitive assessments. *Assessment*. [Epub ahead of print]. doi: 10.1177/1073191116643164

Smith, G.E., Wong, J.S., Ivnik, R.J., & Malec, J.F. (1997). Mayo's older American normative studies: Separate norms for WMS-R logical memory stories. *Assessment*, *4*(1), 79–86.

Tate, R. (2010). *A compendium of tests, scales and questionnaires: The practitioner's guide to measuring outcomes after acquired brain impairment*. New York: Psychology Press.

Temkin, N.R., Heaton, R.K., Grant, I., & Dikmen, S.S. (1999). Detecting significant change in neuropsychological test performance: A comparison of four models. *Journal of the International Neuropsychological Society*, *5*(4), 357–369.

Tranel, D. (2009). The Iowa-Benton School of Neuropsychological Assessment. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric and neuromedical disorders*. New York: Oxford University Press.

Twamley, E.W., Woods, S.P., Dawson, M.S., Narvaez, J.M., & Jeste, D.V. (2007). Remembering to remember: Prospective memory impairment in schizophrenia. *Schizophrenia Bulletin*, *33*(2), 578–578.

Valciukas, J.A., Levin, S.M., Nicholson, W.J., & Selikoff, I.J. (1986). Neurobehavioral assessment of Mohawk Indians for subclinical indications of methyl mercury neurotoxicity. *Archives of Environmental Health*, *41*(4), 269–272.

Weinstein, C.S., Fucetola, R., & Mollica, R. (2001). Neuropsychological issues in the assessment of refugees and victims of mass violence. *Neuropsychology Review*, *11*(3), 131–141. doi: 10.1023/A:1016650623996

Wilson, B., Alderman, N., Burgess, P., Emslie, H., & Evans, J.J. (1996). *Behavioural assessment of the Dysexecutive Syndrome (BADS). Manual*. London: Harcourt Assessment.

Woods, S.P., Carey, C.L., Troster, A.I., Grant, I., & HIV Neurobehavioral Research Center Group. (2005). Action (verb) generation in HIV-1 infection. *Neuropsychologia*, *43*(8), 1144–1151. doi: 10.1016/j.neuropsychologia.2004.11.018

Woods, S.P., Dawson, M.S., Weber, E., Gibson, S., Grant, I., Atkinson, J.H., & HIV Neurobehavioral Research Center Group. (2009). Timing is everything: Antiretroviral nonadherence is associated with impairment in time-based prospective memory. *Journal of the International Neuropsychological Society*, *15*(1), 42–52. doi: 10.1017/S1355617708090012

Woods, S.P., Iudicello, J.D., Dawson, M.S., Moran, L.M., Carey, C.L., Letendre, S.L., … Grant, I. (2007). HIV-associated prospective memory deficits predict functional dependence. *Clinical Neuropsychologist*, *21*(4), 701–701.

Woods, S.P., Iudicello, J.E., Moran, L.M., Carey, C.L., Dawson, M.S., Grant, I., & HIV Neurobehavioral Research Center Group. (2008). HIV-Associated prospective memory impairment increases risk of dependence in everyday functioning. *Neuropsychology*, *22*(1), 110–117. doi: 10.1037/0894-4105.22.1.110

Woods, S.P., Iudicello, J.E., Morgan, E.E., Cameron, M.V., Doyle, K.L., Smith, T.V., & HIV Neurobehavioral Research Center Group. (2016). Health-related everyday functioning in the internet age: HIV-associated neurocognitive disorders disrupt online pharmacy and health chart navigation skills. *Archives of Clinical Neuropsychology*, *31*(2), 176–185. doi: 10.1093/arclin/acv090

Woods, S.P., Morgan, E.E., Dawson, M., Scott, J.C., & Grant, I., HIV Neurobehavioral Research Center Group. (2006). Action (verb) fluency predicts dependence in instrumental activities of daily living in persons infected with HIV-1. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 1030–1042. doi: 10.1080/13803390500350985

Woods, S.P., Scott, J.C., Sires, D.A., Grant, I., Heaton, R.K., & Troster, A.I., HIV Neurobehavioral Research Center Group. (2005). Action (verb) fluency: Test-retest reliability, normative standards, and construct validity. *Journal of the International Neuropsychological Society*, *11*(4), 408–415. doi: 10.1017/S1355617705050460

Woods, S.P., Iudicello, J.E., Morgan, E.E., Verduzco, M., Smith, T.V., & Cushman, C., HIV Neurobehavioral Research Center Group.

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press

(In press). Household everyday functioning in the Internet age: Online shopping and banking skills are affected in HIV-associated neurocognitive disorders. *Journal of the International Neuropsychological Society*.

Zhu, J.J., & Tulsky, D.S. (2000). Co-norming the WAIS-III and WMS-III: Is there a test-order effect on IQ and memory scores? *Clinical Neuropsychologist*, *14*(4), 461–467. doi: 10.1076/Clin.14.4.461.7197

Zogg, J.B., Woods, S.P., Sauceda, J.A., Wiebe, J.S., & Simoni, J.M. (2012). The role of prospective memory in medication adherence: A review of an emerging literature. *Journal of Behavioral Medicine*, *35*(1), 47–62. doi: 10.1007/s10865-011-9341-9

https://doi.org/10.1017/S1355617717001060 Published online by Cambridge University Press