Neuropsychological Rehabilitation, 2015
http://dx.doi.org/10.1080/09602011.2015.1109524

The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology

Thomas D. Parsons, Anne R. Carlew, Jonlih Magtoto, and Kiefer Stonecipher

Computational Neuropsychology and Simulation, Department of Psychology, University of North Texas, Denton, USA

(*Received 2 January 2015; accepted 14 October 2015*)

The assessment of executive functions is an integral task of neuropsychological assessment. Traditional measures of executive function are often based on hypothetical constructs that may have little relevance to real-world behaviours. In fact, some traditional tests utilised today were not originally developed for clinical use. Recently, researchers have been arguing for a new generation of "function-led" neuropsychological assessments that are developed from directly observable everyday behaviours. Although virtual environments (VEs) have been presented as potential aides in enhancing ecological validity, many were modelled on construct-driven approaches found in traditional assessments. In the current paper, we review construct-driven and function-led VE-based neuropsychological assessments of executive functions. Overall, function-led VEs best represent the sorts of tasks needed for enhanced ecological validity and prediction of real-world func
tioning.

Keywords: Virtual environments; Neuropsychological assessment; Ecological validity; Executive function

Correspondence should be addressed to Thomas D. Parsons, Ph.D, Associate Professor of Psychology, Director, Computational Neuropsychology and Simulation (CNS) Lab, Department of Psychology, University of North Texas, 1155 Union Circle #311280, Denton, TX 76203, USA. http://psychology.unt.edu/~tparsons/, E-mail: Thomas.Parsons@unt.edu

No potential conflict of interest was reported by the authors.

INTRODUCTION

The assessment of executive functioning represents a principal objective of neuropsychological evaluations. Neuropsychological measures of executive functions aim to assess a number of constructs: selective attention, inhibitory control, planning, problem solving, and some aspects of short-term memory (Burgess & Simons, 2005; Chan, Shum, Toulopoulou, & Chen, 2008; Stuss, Shallice, Alexander, & Picton, 1995; see Diamond, 2013 for review). Norman and Shallice (1986) have proposed a model of executive functioning that emphasises a supervisory attentional system (SAS). According to the SAS model, executive functions involve two systems: (1) a Contention Scheduling System that is responsible for automatic and overlearned behaviours; and (2) a Supervisory Attentional System that is responsible for controlled processing of non-routine and novel tasks. Stuss et al. (1995) identified tasks that may be used to assess executive functioning based on the SAS. For example, the Stroop task involves conflict, and the Wisconsin Card Sorting Test, Trail Making Test and verbal fluency all involve mental switching. A difficulty for the neuropsychological assessment of the dysexecutive syndrome is that patients' performance on an executive function test may have little or no predictive value for how they may perform in a real-world situation (Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Chaytor, Schmitter-Edgecombe, & Burr, 2006). To address this issue, neuropsychologists are increasingly emphasising the need for tasks that represent real-world functioning and tap into a number of executive domains (Chaytor & Schmitter-Edgecombe, 2003; Jurado & Rosselli, 2007).

Virtual environments (VEs) are increasingly considered as potential aids in enhancing the ecological validity of neuropsychological assessments (Campbell et al., 2009; Parsons, 2011; Renison, Ponsford, Testa, Richardson, & Brownfield, 2012; Schultheis, Himelstein, & Rizzo, 2002). Part of this increased interest is due to recent (past 10–15 years) enhancements in 3D rendering and shading capabilities that have allowed for greatly improved textures in computer graphics. Earlier virtual reality (VR) equipment suffered a number of limitations, such as being large and unwieldy, difficult to operate, and very expensive to develop and maintain. Over the past decade, researchers have steadily progressed in making VE hardware and software more reliable, cost effective, and acceptable in terms of size and appearance (Bohil, Alicea, & Biocca, 2011). The VEs of today are advanced computer interfaces that allow patients to become immersed within a computer-generated simulation of everyday activities.

Given that VEs represent a special case of computerised neuropsychological assessment devices (Bauer et al., 2012; Schatz & Browndyke, 2002) they have enhanced computational capacities for administration efficiency, stimulus presentation, automated logging of responses, and data analytic processing. Since

VEs allow for precise presentation and control of dynamic perceptual stimuli, they can provide ecologically valid assessments that combine the control and rigour of laboratory measures with a simulation that reflects real-life situations (Bohil et al., 2011). Additionally, the enhanced computation power allows for increased accuracy in the recording of neurobehavioural responses in a perceptual environment that systematically presents complex stimuli. Such simulation technology appears to be distinctively suited for the development of ecologically valid environments, in which 3D objects are presented in a consistent and precise manner (Schultheis et al., 2002). VE-based neuropsychological assessments can provide a balance between naturalistic observation and the need for exacting control over key variables (Campbell et al., 2009). Importantly, the development of VEs is not meant to replace neuropsychologist involvement, but rather to augment the neuropsychological assessment to enhance its predictive ability. To dismiss the neuropsychologist from the neuropsychological assessment would be to sacrifice valuable qualitative observations. In summary, VE-based neuropsychological assessments allow for real-time measurement of multiple neuropsychological abilities in order to assess complex sets of skills and behaviours that may more closely resemble real-world functional abilities (Matheis et al., 2007).

WHAT CONSTITUTES AN ECOLOGICALLY VALID ASSESSMENT OF EXECUTIVE FUNCTIONING?

Verisimilitude and veridicality

A difficult issue facing neuropsychologists interested in adding VEs to their assessments of real-world functioning is the question of what constitutes an ecologically valid assessment of executive functioning. Early attempts at a definition of ecological validity (Franzen & Wilhelm, 1996) for neuropsychological assessment emphasised two requirements: (1) veridicality, in which the participant's performance on a construct-driven measure should predict some feature(s) of the participant's day-to-day functioning (e.g., vocational status); and (2) verisimilitude, in which the requirements of a neuropsychological measure and the testing conditions should resemble requirements found in a participant's activities of daily living (Spooner & Pachana, 2006). Early discussions of verisimilitude in neuropsychology emphasised that the technologies current to the time could not replicate the environment in which the behaviour of interest would ultimately take place (Goldstein, 1996). Almost 20 years later, most neuropsychological assessments represent outdated technologies (e.g., paper-and-pencil assessments; static stimuli) that are yet to be validated with respect to real-world functioning (Rabin, Burton, & Barr, 2007). Furthermore, current verisimilitude assessments are somewhat conflicted in that while they focus on cognitive constructs (e.g., attention, executive function, memory), they are used for identifying functional abilities (Chaytor & Schmitter-Edgecombe, 2003).

Construct-driven and function-led assessments

Burgess et al. (2006) argue that the majority of neuropsychological assessments currently in use today were developed to assess cognitive "constructs" without regard for their ability to predict "functional" behaviour. For example, one of the most widely used measures of executive function is the Wisconsin Card Sorting Test (WCST). The WCST was not originally developed as a measure of executive functioning. Instead, the WCST was preceded by a number of sorting measures that were developed from observations of the effects of brain damage (e.g., Weigl, 1927). Nevertheless, in a single study by Milner (1963), patients with dorsolateral prefrontal lesions were found to have greater difficulty on the WCST than patients with orbitofrontal or nonfrontal lesions. While data from the WCST do appear to provide information relevant to the constructs of "set shifting" and "working memory", the data do not necessarily offer information that would allow a neuropsychologist to predict what situations in everyday life require the abilities that the WCST measures. Furthermore, it has been shown that patients with frontal lobe pathology do not always differ from control subjects on the WCST (Stuss et al., 1983).

CONSTRUCT-DRIVEN VIRTUAL ENVIRONMENTS

Although VE-based neuropsychological assessments have been proposed as a potential answer to the requirements for generalisability of everyday functioning, many of the VEs that have been developed simply recreate construct-driven assessments in a simulated environment. For example, a number of early VE-based neuropsychological assessments were modelled on the WCST (Elkind, Rubin, Rosenthal, Skoff, & Prather, 2001; Pugnetti, Mendozzi, Attree et al., 1998; Pugnetti, Mendozzi, Motta et al., 1995). One of the first VEs modelled on the WCST required patients to reach the exit of a virtual building through the use of environmental cues (e.g., categories of shape, colour, and number of portholes) that aided in the correct selection of doors leading from room to room (Pugnetti, Mendozzi, Attree et al., 1998; Pugnetti, Mendozzi, Motta et al., 1995). Similar to the WCST, after a fixed number of successful trials the correct choice criteria (e.g., categories) were changed so that the patient had to shift cognitive set and devise a new choice strategy in order to pass into the next room. Pugnetti, Mendozzi, Attree et al. (1998) compared neurologically impaired patients and non-impaired controls on both the VR task and the WCST. While the controls performed more successfully on both tests, weak correlations were found between the VR task and the WCST. As a result, there is a question about whether the WCST and the VE-based assessment were measuring different functions. It is important to note that the Pugnetti version had a heavy reliance on navigating through a building and this may have confounded the results.

A more current VE-based neuropsychological assessment modelled on the WCST did not have the potentially confounding effects of navigation. In the Virtual Reality Look for a Match Test (VRLFAM), Elkind et al. (2001) developed a beach scene, in which participants were asked to deliver frisbees, sodas, popsicles, and beach balls to umbrellas. Each umbrella had one of the four objects on it (differing in type, colour, and number). As participants delivered the objects, they received verbal feedback (e.g., "That's it" or "That's not what I want"). Following the WCST, the participant had 128 turns to match (twice) to the appropriate target (in the following order: 10 times to colour, 10 to object, and 10 times to number−in that order) and successfully complete the task. Results from a comparison of healthy control performance on the VRLFAM and the WCST indicated that all performance scales (with the exception of WCST perseverative errors) were directly related (Elkind et al., 2001). An unfortunate limitation of modelling VE-based neuropsychological assessments on the WCST is that the virtual analogues, like the original WCST, may not be able to differentiate between patients with frontal lobe pathology and control subjects (Stuss et al., 1983). Furthermore, while data from the VE-based assessments, like the WCST, do appear to provide information relevant to the constructs of "set shifting" and "working memory", the VE assessments seem to do little to extend ecological validity.

Following the VE-based WCST, a number of other construct-driven VEs have been developed. For instance, various virtual classroom environments (e.g., ClinicaVR Digital Media Works; AULA) have emerged that include construct-driven Stroop (Lalonde, Henry, Drouin-Germain, Nolin, & Beauchamp, 2013) and Go/No-Go stimuli (Parsons, Bowerly, Buckwalter, & Rizzo, 2007). Like the WCST studies mentioned above, these VEs have been compared to traditional paper-and-pencil assessments. Typically, they have good convergent and discriminant validity (Diaz-Orueta et al., 2014; Parsons et al., 2007; Parsons & Courtney, 2014). Furthermore, they at times offer enhanced classification of attentional deficits when distractors (e.g., bell ringing, teacher answering the classroom door, principal entering the room) are introduced into the VE (Adams, Finn, Moes, Flannery, & Rizzo, 2009; Bioulac et al., 2012; Iriate et al., 2012; Pollak, Shomaly et al., 2010; Pollak, Weiss et al., 2009). The classroom paradigm has been extended to a virtual apartment that superimposes construct-driven stimuli (e.g., Stroop and Continuous Performance Test (CPT)) onto a large television set in the living room (Henry, Joyal, & Nolin, 2012). Results from regression analysis indicated that commission errors and variability of reaction times in the VR-Apartment Stroop were significantly predicted by scores of the Elevator task and the CPT-II. These preliminary results suggest that the VR-Apartment Stroop is an interesting measure of cognitive and motor inhibition for adults.

While the actual cognitive neuropsychological construct assessed may vary, the construct-driven VE paradigm consistently involves the superimposing of construct-driven stimuli onto some aspect of the environment. Unfortunately, such tests are little more than adaptations of outmoded conceptual and experimental frameworks. These construct-driven tests fail to represent the actual functional capacities inherent in executive functions. Measures like the WCST were not originally designed to be used as clinical measures (Burgess et al., 2006). Instead, they were found to be useful tools for cognitive neuropsychological assessment and normal populations and then later found their way into the clinical realm to aide in assessing constructs that are important to carrying out real-world activities. One explanation for the prevailing emphasis upon construct-driven assessments is technological advances (e.g., new methods of brain imaging). This rapid rate of methodological innovation in the field of executive function has not matched the rate of technological advance. As a result, the majority of the early research papers using the new technologies (e.g., functional brain imaging) understandably leaned towards using the previously existing tools because there was a need to validate new technologies.

NEED FOR FUNCTION-LED ASSESSMENTS

Burgess et al. (2006) suggest that future development of neuropsychological assessments should result in tests that are "representative" of real-world "functions" and proffer results that are "generalisable" for prediction of the functional performance across a range of situations. According to Burgess et al. (2006) a "function-led approach" to creating neuropsychological assessments will include neuropsychological models that proceed from directly observable everyday behaviours backward to examine the ways in which a sequence of actions leads to a given behaviour in normal functioning; and the ways in which that behaviour might become disrupted. As such, he calls for a new generation of neuropsychological tests that are "function led" rather than purely "construct driven". These neuropsychological assessments should meet the usual standards of reliability, but discussions of validity should include both sensitivity to brain dysfunction and generalisability to real-world function.

A number of investigators have argued that performance on traditional tests of executive function (e.g., WCST, Stroop Test) has little correspondence to activities of daily living. As such, neuropsychologists are left uncertain of the efficacy of these tests for predicting the way in which patients will manage in their everyday lives (Bottari, Dassa, Rainville, & Dutil, 2009; Manchester, Priestley, & Howard, 2004; Sbordone, 2008). According to Chan et al. (2008) most of these traditional measures assess at the impairment level and do not capture the complexity of response required in the many multistep tasks found in everyday activities. It is important to note that a number of function-led tests of executive function have been developed to assess real-world planning (e.g., Zoo Map and Six Elements subtests of the Behavioural Assessment of Dysexecutive Syndrome; Wilson, Alderman, Burgess, Emslie, & Evans, 1996) and self-regulation (e.g., the Revised Strategy Application Test, Levine, Dawson, Boutet, Schwartz, & Stuss, 2000; Sustained Attention to Response Test, Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; see Chan et al., 2008 for review).

Multiple errands paradigm for function-led assessments

Shallice and Burgess (1991) developed the Multiple Errands Test (MET) as a function-led assessment of multitasking. The MET requires the patient to perform a number of relatively simple but open-ended tasks in a shopping context. Participants are required to achieve a number of simple tasks without breaking a series of arbitrary rules. The MET has been shown to have increased sensitivity (over traditional neuropsychological measures) to elicit and detect failures in executive function (e.g., distractibility and task implementation deficits). It has also been shown to be better at predicting behavioural difficulties in everyday life (Alderman, Burgess, Knight, & Henman, 2003). Furthermore, the MET has been found to have strong inter-rater reliability (Dawson et al., 2009; Knight, Alderman, & Burgess, 2002), and performance indices from the MET were able to predict significantly severity of everyday life executive problems in persons with traumatic brain injury (TBI) (Cuberos-Urbano et al., 2013).

FUNCTION-LED VIRTUAL ENVIRONMENTS

Virtual errands tasks

Potential limitations for the MET are apparent in the obvious drawbacks of experiments conducted in real-life settings (e.g., Bailey, Henry, Rendell, Phillips, & Kliegel, 2010). Logie, Trawley, and Law (2011) point out a number of limitations in the MET: (1) time consuming; (2) transportation is required for participants; (3) consent from local businesses; (4) lack of experimental control; and (5) difficulty in adapting tasks for other clinical or research settings. McGeorge et al. (2001) modelled a Virtual Errands Test (VET) on the original MET. However, the VET tasks were designed to be more vocationally oriented in format, containing work-related errands as opposed to the shopping errands used in the MET. In a study involving five adult patients with brain injury and five unimpaired matched controls, participants completed both the real-life MET and the VET. Results revealed that performance was similar for real-world and VE tasks. In a larger study comparing 35 patients with prefrontal neurosurgical lesions to 35 controls matched for age and estimated IQ (Morris, Kotitsa, Bramham, Brooks, & Rose, 2002), the VE scenario was found to successfully differentiate between participants with brain injuries and controls. A limitation of these early VEs is that the graphics were unrealistic, and performance assessment involved video recording test sessions with subsequent manual scoring.

In the past decade a number of VEs with enhanced graphics (and usability) have been developed to model the function-led approach found in the MET. In addition to the VEs for assessment of nonclinical populations (Logie et al., 2011), a number of virtual errand protocols have been developed to evaluate executive functions of clinical populations (see Table 1 for examples of Virtual Errand Protocols over the past 10 years). For example, virtual shopping scenarios (see Parsons, McPherson, & Interrante, 2013 for review) offer an advanced computer interface that allows the clinician to immerse the patient within a computer-generated simulation that reflects activities of daily living. They involve a number of errands that must be completed in a real environment following certain rules that require problem solving. Since these programmes allow for precise presentation and control of dynamic perceptual stimuli, they have the potential to provide ecologically valid assessments that allow the control of laboratory measures within simulations that reflect real-life situations.

A number of other function-led VEs are being modelled to reflect the multitasking demands found in the MET. The Multitasking in the City Test (MCT) is modelled on the MET and involves an errand-running task that takes place in a virtual city (Jovanovski, Zakzanis, Campbell, et al., 2012, 2012). The MCT can be distinguished from existing VR and real-life METs. For instance, MCT tasks are performed with less explicit rule constraints. This contrasts with the MET, in which participants must abide by certain rules (not travelling beyond a certain spatial boundary and not entering a shop other than to buy something). This difference was intentional in the MCT because the researchers aimed to investigate behaviours that are clearly not goal-directed. The MCT is made up of a virtual city that includes a post office, drug store, stationery store, coffee shop, grocery store, optometrist's office, doctor's office, restaurant/pub, bank, dry cleaners, pet store, and the participant's home. Although all buildings in the MCT VE can be entered freely, interaction within them is possible only for those buildings that must be entered as part of the task requirements. The MCT was used to compare a sample of post-stroke and TBI patients to an earlier sample of

TABLE 1
Assessment of executive functioning using virtual errands protocols

Study	*Virtual environment*	*Traditional tests*	*Design*	*Outcome*
Carelli, Morganti, Weiss, Kizony, and Riva (2008)	Virtual Supermarket	N/A	1 within subjects (n = 20)	Results suggest the virtual supermarket may be a useful tool in executive assessment, particularly due to its temporal and accuracy measures.
Cipresso et al. (2013)	Virtual Multiple Errands Test (VMET)	MMSE; DS; Short Story Recall Test; TMT A, B, & B-A; FAB; Corsi span; Corsi Block Task; phonemic fluency test; semantic fluency test; disyllabic word test; ToL; Token test; Street Completion Test; State and Trait Anxiety Index; BDI	16 within subjects. OCD group (n = 15) vs. Control n = 15)	While performing routine tasks in the VMET, patients with OCD had more difficulties working with breaks in volition than normal controls.
Josman et al. (2014)	Virtual Action Planning Supermarket (VAP-S)	BADS, OTDL-R	3 within subjects. Stroke (n = 24) vs. Control (n = 24)	Results revealed there were significant differences between the research group and healthy controls in two outcome measures of the VAP-S. Participants who had a stroke showed a modest relationship between the BADS and number of purchases in the VAP-S. The VAP-S demonstrated decent predictive validity of IADL performance.

(*Continued*)

TABLE 1 Continued.

Study	*Virtual environment*	*Traditional tests*	*Design*	*Outcome*
Josman, Schenirderman, Klinger, and Shevil (2009)	Virtual Action Planning Supermarket (VAP-S)	BADS	2 within subjects. Schizophrenia (n = 30) vs. Control (n = 30) 3	The VAP-S was sensitive to differences in executive functioning between schizophrenia patients and controls, and appeared to be a feasible task to use for assessments of this population.
Jovanovski, Zakzanis, Ruttan, et al. (2012)	Multitasking in the City Test (MCT)	WTAR, ToMM, COWAT, Semantic Fluency, WCST, MSET, TMT, WAIS—III, JOL, RCFT, CVLT -II; WMS- III, BDI, BAI, FSBS	16 within-subjects (n = 13 stroke or TBI)	Results of data analyses suggest that patients can be differentiated from normal samples. Significant correlations between the MCT and standardised neuropsychological assessments were also established, suggesting the MCT is a valid assessment of executive functioning.
Jovanovski, Zakzanis, Campbell, et al. (2012)	Multitasking in the City Test (MCT)	TMT, ToL, MSET, WTAR, JOL, Star Cancellation, RBMT-extended	8 within subjects (n = 30)	The MCT was found to have low associations with other tests of executive function, aside from the MSET. The MSE was correlated with the MCT plan score. The MCT had more ecological validity than traditional tests of executive function, and appeared to be a valid measure of assessing integration of executive functions into meaningful behaviour.
Law, Trawley, Brown, Stephens, and Logie (2013)	Edinburgh Virtual Errands Test (EVET)	N/A	2 (between-subjects) Good vs. Poor factor plan x 2 (within-subjects) single task, dual task (n = 40)	Participants were able to navigate the environment and perform tasks. Factor demand and factor plan were shown to have little effect on ability to complete tasks.

Logie et al. (2011)	EVET	Word Recall Task, Working Memory Verbal Span, Working Memory Spatial Span, Travelling Salesman Task, Breakfast Task	6 within subjects (n = 165)	The EVET scores could be predicted by measures of retrospective memory, visuospatial planning and spatial working memory. A limitation of the EVET is its lack of generalisability to other scenarios. Authors suggested modifications could make it more applicable to other research questions.
McGeorge et al. (2001)	Virtual Building Environment (VET)	N/A	1 within subjects TBI (n = 5) vs Control (n = 5)	The control sample performed significantly better on the task, completing more errands and having higher quality written plans.
Morris et al. (2002)	Virtual Bungalow Task	N/A	Patients with prefrontal neurosurgical lesions (n = 35) vs. Controls (n = 35)	The VE scenario was found to successfully differentiate between participants with brain injuries and controls. Patients showed increased rule breaking behaviours and deficits in strategy formation compared to controls.
Patil, Cogoni, Zangrando, Chittaro, and Silani (2014)	Virtual Reality Moral Dilemma Scenarios	Traditional text scenarios	2 within subjects (n = 40)	Many of the participants made non-utilitarian judgements in textual versions of the dilemmas, but demonstrated utilitarian behaviours in the VR scenarios. This suggests that the VEs altered the contextual information, evoking more emotional responses to the dilemmas.

(*Continued*)

TABLE 1 Continued.

Study	Virtual environment	Traditional tests	Design	Outcome
Pedroli et al. (2013)	Virtual Multiple Errands Test (VMET)	15 (within-subjects) MMSE, AVLT, DS, Corsi's Span, Supra-span, Short Story, ToL, Verbal Fluency Test, Benton's JOL, WAIS-R, Laiacona's Naming Test, TMT, State Trait Anxiety Index, BDI	15 within subjects Parkinson's disease (n = 3) vs. Control (n = 21)	The VMET showed good reliability. However, the System Usability Scale suggests that the VMET may need minor improvements to increase usability for patients with cognitive disorders. Nevertheless, the VMET seems to be a reliable and ecologically valid measure.
Rand, Rukan, Weiss, and Katz (2009)	Virtual Multiple Errands Test (VMET)	Zoo Map Test, IADL, MET	4 within post-stroke (n = 9) vs. young adults (n = 20, mean = 26.3) vs. older adults (n = 20, mean age = 64)	The VMET was able to distinguish between healthy and post-stroke groups as well as between older and younger individuals. Overall, the VMET seems to be a sensitive measure of executive function in both brain-injured and healthy individuals.
Raspelli et al. (2012)	Virtual Multiple Errands Test (VMET)	MMSE, Star Cancellation Test, Token Test, Street's Completion Test, Test of Attentional Performance, Stroop Test, Iowa Gambling Task, DEX, ADL, IADL, State Trait Anxiety Index, BDI	13 within subjects post-stroke (n = 9) vs. Young adults (n = 10, mean = 26) vs. Older adults (n = 20, mean age = 64)	Stroke patients showed the highest number of errors and slowest reaction time on the VR-MET, followed by older adults and younger adults. These findings suggest the VMET is a valid tool for the assessment of executive functioning.

Sauzéon et al. (2014)	Virtual Human Object Memory for Everyday Scenes (Virtual HOMES)	8 (within-subjects) MMSE, MRT, Spatial span, Stroop task, CVLT , SSQ, NTQ, QAM	9 within subjects Alzheimer's disease (n = 16) vs. Younger adults (n = 23) vs. Older adults (n = 23)	Results revealed that, compared to younger participants, older individuals showed signs of decline in free recall, but Alzheimer's patients demonstrated deficits in several aspects of executive functioning. These results also indicate that the virtual assessment is useful for identifying age-related differences and effects of Alzheimer's disease on cognitive functioning.
Toet and van Schaik (2013)	Virtual Town	N/A	3 between-subjects (Pleasant vs. Unpleasant vs. Neutral olfactory groups) (n = 69)	Participants navigated a town while smelling either pleasant, unpleasant, or neutral odours. Odours did not seem to impact visual attention.
Werner, Rabinowitz, Klinger, Korczyn, and Josman (2009)	Virtual Action Planning Supermarket (VAP-S)	BADS	2 within-subjects MCI (n = 30) vs. Control (n = 40)	The VAP-S was able to discriminate between MCI and the control group. This suggests that the VAP-S is a valid measure of executive function.

MMSE = Mini-Mental State Examination, DS = Digit Span, TMT A & B = Trail Making Test Parts A & B, FAB = Frontal Assessment Battery, ToL = Tower of London, STAI = State Trait Anxiety Inventory, BDI = Beck Depression Inventory, BADS = Behavioural Assessment of the Dysexecutive Syndrome, OTDL-R = Observed Tasks of Daily Living–Revised, AVLT = Auditory-Verbal Learning Test, JOL = Judgement of Line Orientation, WAIS-R = Wechsler Adult Intelligence Scale–Revised, IADL = Instrumental Activities of Daily Living, MET = Multiple Errands Test, DEX = Dysexecutive Questionnaire, ADL = Activities of Daily Living, MRT = Mental Rotation Task, CVLT = California Verbal Learning Task, SSQ = Simulator Sickness Questionnaire, NTQ = New Technology Questionnaire, QAM = Self-Evaluation Questionnaire.

normal controls. Jovanovski, Zakzanis, Ruttan, et al. (2012) found that although the patient sample developed adequate plans for executing the tasks, their performance of the tasks revealed a greater number of errors. The MCT was significantly correlated with a rating scale completed by significant others.

A number of other virtual assessments modelled on the MET have been created and validated in samples with stroke or injury-related brain deficits. These protocols are often placed in living or work settings (see Table 2 for examples of Function-Led Virtual Environments for assessing executive functioning from the past 10 years): Virtual Office Tasks (Jansari, Froggatt, Edginton, & Dawkins, 2013; Lamberts, Evans, & Spikman, 2009; Montgomery, Ashmore, & Jansari, 2011); Virtual Apartment/Home Tasks (Saidel-Goley, Albiero, & Flannery, 2012; Sweeney, Kersel, Morris, Manly, & Evans, 2010); Virtual Park (Buxbaum, Dawson, & Linsley, 2012); Virtual Library Task (Renison et al., 2012); Virtual Anticipating Consequences Task (Cook et al., 2013); Virtual Street Crossing (Avis, Gamble, & Schwebel, 2014; Clancy, Rucklidge, & Owen, 2006; Davis, Avis, & Schwebel, 2013; Nagamatsu et al., 2011); and Virtual Kitchen (Cao, Douguet, Fuchs, & Klinger, 2010).

Driving simulator paradigm for function-led assessments

Another area of function-led assessment can be found in VE-based neuropsychological assessments that use driving simulators. The successful operation of a motor vehicle requires coordination of multiple functional behaviours. Given its complexity, driving is often an ability that becomes difficult for clinical populations. It is important to note that the literature on neuropsychological assessment using driving simulation is vast and beyond the scope of this review. Perhaps the most comprehensive review of executive function assessments in relation to fitness to drive is Asimakopulos et al.'s (2012). For additional articles about driving simulation see Calhoun and Pearlson (2012) or Schultheis, Rebimbas, Mourant, and Millis (2007). Unfortunately, these reviews do not look at the construct-driven versus function-led approach that is discussed herein. As a result, we would be remiss if we did not attempt to give a general summary of the efforts of researchers in this area.

VR driving simulators have been used to investigate driving performance in individuals with attention deficit hyperactivity disorder (Barkley, Anderson, & Kruesi, 2007; Barkley, Murphy, O'Connell, & Connor, 2005; Cox et al., 2008; Knouse, Bagwell, Barkley, & Murphy, 2005), alcohol or drug impairment (Allen et al., 2009; Barkley, Murphy, O'Connell, Anderson, & Connor, 2006), and brain injury (Liu, Miyazaki, & Watson, 1999; Milleville-Pennel, Pothier, Hoc, & Mathe, 2010; Schultheis & Mourant, 2001; Schultheis et al., 2007; Wald & Liu, 2001; Wald, Liu, & Reil, 2000).

TABLE 2
Assessment of attention and executive functioning using function-led VEs

Study	Virtual environment	Traditional tests	Design	Outcome
Avis et al. (2014)	Virtual Reality Pedestrian Environment (VRPE)	N/A	1 within subjects Children with EDS (n = 33, M age = 12.93) vs. Controls (n = 33)	Children with EDS were found to be much more likely to get hit in the VE. Attention to traffic was not found to be significantly different between groups, but decision-making time did vary across groups, suggesting that children with EDS may take longer to determine whether it is safe to cross.
Buxbaum et al. (2012)	Virtual Park	Bell Cancellation Test; RBIT; Fluff Test; Real World Navigation Test	5 within subjects Right hemisphere stroke (n = 70) vs. Controls (n = 10)	Results indicated strong reliability, validity, and specificity in the ability of the VRLAT to evaluate hemispatial neglect in comparison to traditional neuropsychologicalal assessments.
Cao et al. (2010)	Therapeutic Virtual Kitchen (TVK)	N/A	E1: 1 within subjects (n = 13) E2: 1 within subjects (n = 7 ; 4 TBI, 2 stroke, 1 meningoencephalitis)	Results revealed that a daily life task can be virtually simulated to assess executive functioning, and patients with TBI and normal participants could complete the task.
Cook et al. (2013)	Virtual Anticipating Consequences Task (VR-AC)	N/A	1 within subjects Adolescents with TBI (n = 15) vs. Controls (n = 13)	No significant differences were observed in the number of short-term consequences in the two groups. However, the TBI group had significantly more long-term consequences than the control group.

(*Continued*)

TABLE 2 Continued.

Study	Virtual environment	Traditional tests	Design	Outcome
Clancy et al. (2006)	Road Crossing Virtual Apparatus	K-SADS-PL; Conners' Rating Scales–Revised; CBCL; New Zealand Socioeconomic Index of Occupational Status; word reading, spelling, pseudoword decoding (WIAT-II); Block Design, Vocabulary (WAIS-III)	10 within subjects Children with ADHD ($n = 24$) vs. Control ($n = 24$)	ADHD and control adolescents' behaviours were compared in road-crossing safety (including speed, attention, and other safety behaviours) in a VE. ADHD adolescents had low margins of safety compared to controls, and were hit twice as often as controls. Participants with ADHD walked slower and used less of the available space to cross the road.
Davis et al. (2013)	Virtual Reality Pedestrian Environment (VRPE)	N/A	1 within subjects adolescents deprived of sleep ($n = 55$)	Sleep restriction seems to be correlated with an increase in risky behaviours, as sleep-deprived participants did not pay as much attention to traffic, made poorer decisions when crossing, and took longer to make those decisions.
Jansari et al. (2013)	Virtual Office Assistant	N/A	2 × 2 (between-subjects) Smokers ($n = 36$) vs. Non-Smokers ($n = 36$)	Results from prospective memory tasks revealed nicotine improved performance in smokers, but not in non-smokers. Non-smokers performed better than smokers on selective and adaptive thinking.
Lamberts et al. (2009)	Virtual Executive Secretarial Task (VEST)	DEX, EOS, TMT, 15 Words Test, BADS	6 within subjects TBI ($n = 35$) vs. Control ($n = 57$)	The VEST successfully discriminated between the two groups. Further, the VEST was able to give valuable information on real-life functioning in individuals with brain injury.

Montgomery et al. (2011)	Virtual Office Environment	N/A	2 between-subjects. Alcohol vs. Placebo ($n = 40$)	Significant impairments in both executive functioning and prospective memory were found in the alcohol group. These findings suggest that people who have had even a modest amount of alcohol may not realise the extent of their planning abilities and performance impairment.
Nagamatsu et al. (2011)	Street crossing	N/A	1 within-subjects ($n = 33$, older adults)	Results revealed older adults “at risk” for falls had significantly slower walking speed and made poorer decisions while crossing the street when faced with a secondary task. This indicates these at-risk patients may have reduced executive function in the realm of planning and decision making when faced with increased cognitive load.
Renison et al. (2012)	Virtual Library Task (VLT)	Real Library Task, WTAR, LM-II, DS, WCST, Brixton Spatial Anticipation Test; Zoo Map Test; MSET	9 within subjects TBI ($n = 30$) vs. Control ($n = 30$)	The VLT Task was highly positively correlated with the Real Library Task, and was able to discriminate between executive function in people with TBI and a control group. Results suggest the VLT can be used reliably to test patients with TBI.

(*Continued*)

TABLE 2 Continued.

Study	Virtual environment	Traditional tests	Design	Outcome
Saidel-Goley et al. (2012)	Virtual Apartment	DS, BDI-II, TMT, Corsi Block-Tapping Test	5 within subjects ($n = 45$)	Results showed positive correlations between dissociation and performance on tasks that assessed attention and working memory, indicating that nonclinical dissociation seems to enhance cognitive functioning. Comparison of performance on the VR task and standardised neuropsychological assessments also supported increased ecological validity of VR assessments.
Sweeney et al. (2010)	Bungalow Task	N/A	2 within subjects TBI ($n = 17$) vs. Control ($n = 17$)	The VRT revealed differences between individuals with TBI and controls in executive functioning. Specifically, the virtual task revealed differences in planning and prospective memory between TBI and control groups, with the TBI group showing greater deficits.

K-SADS-PL = Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version; CBCL = Child Behaviour Checklist; WIAT-II = Wechsler Individual Achievement Test—Second Edition, WAIS-III = Wechsler Adult Intelligence Scale—Third Edition; WTAR = Wechsler Test of Adult Reading, ToMM = Test of Memory Malingering, COWAT = Controlled Oral Word Association Test, WCST = Wisconsin Card Sorting Test, MSET = Modified Six Elements Test, TMT = Trail Making Test, RCFT = Rey-Osterreith Complex Figure Test, JOL = Judgement of Line Orientation Task, CVLT-II = California Verbal Learning Test II, WMS-II = Wechsler Memory Scale III; BAI = Beck Anxiety Inventory, FSBS = Frontal Systems Behaviour Scale, ToL = Tower of London, RBMT = Rivermead Behavioural Memory Test, DEX = Dysexecutive Questionnaire, EOS = Executive Observation Scale, BADS = Behavioural Assessment of the Dysexecutive Syndrome; LM = Logical Memory II (WMS), DS = Digit Span.

While driving simulators may not assess driving capabilities in a manner exactly the same as an on-road driving test, both tests have their limitations as indicators of actual driving performance, owing to different methods and demand characteristics. Further, past research has shown that VR simulators have evidence of validity for predicting actual driving performance and risks (Bedard, Parkkari, Weaver, Riendeau, & Dahlquist, 2010; Lee, Cameron, & Lee, 2003; Lee, Lee, Cameron, & Li-Tsang, 2003).

While there is currently limited empirical evidence to determine the efficacy of driving simulation for function-led assessment of executive functioning, Bedard et al. (2010) found preliminary support for assessing driving performance in relation to neuropsychological functioning using driving simulations. Furthermore, it may be best to view driving simulator results as complementary assessment data to traditional neuropsychological assessment data (Milleville-Pennel et al., 2010).

FUNCTIONAL OUTCOMES

In the development of function-led assessments for VEs, researchers should take care to place emphasis on tuning each task to capture the desired outcomes to be measured. For example, patients with frontal lobe damage often present with deficits in a variety of everyday behaviours: shopping, cooking, balancing a chequebook, medication adherence, and driving. While virtual errands tests (discussed earlier) have proven to be useful and quite effective at capturing a number of these domains, there is a need to develop VE-based neuropsychological assessments that proceed backward from directly observable everyday behaviours (to examine the ways in which a sequence of actions in the VE occurs) to a prediction of such behaviours occurring in everyday activities. One way to implement this would be to develop a driving simulator that taps into the sorts of deficits found in poor driving performance (e.g., car accidents, traffic violations, obstacles collided with, etc.) that might be obtained from proxy or Department of Motor Vehicles (DMV) records. This information should be combined with the type information that neuropsychologists receive while identifying and quantifying behavioural problems associated with frontal lobe dysfunction. Next, driving simulators may be developed that measure critical functional outcomes such as car accident risks, traffic violations, and collisions (Calhoun & Pearlson, 2012; Parsey & Schmitter-Edgecombe, 2013). Once developed, the driving simulator can be used by the neuropsychologist for comparisons of driving performance in the virtual world to traditional assessments of frontal lobe functioning and with previous records, such as DMV reports. An additional advantage of the driving simulator would be that there would be a record of the sequences of actions leading to each aspect of driving performance that may enlighten and enhance prediction of such poor driving performance in the real world.

VE-based neuropsychological assessments afford several important advantages for clinical application in the assessment of deficits. Some positive attributes of these VE-based neuropsychological assessments include enhanced ecological validity, simulation customisability, affordability, safety and efficiency, applicability to a wide range of impairments, user-friendly interfaces, data capture, and real-time analyses of performance (Parsons, 2011). That said, some VEs do little more than place traditional construct-driven stimuli into various simulations of real-life environments. Hence, the difficult issue facing neuropsychologists interested in adding VEs to their assessments of real-world functioning is the question of what constitutes an ecologically valid assessment of executive functioning. As Burgess et al. (2006) have pointed out, the majority of neuropsychological assessments currently in use today were developed to assess cognitive "constructs" without regard for their ability to predict "functional" behaviour. VEs that simply recycle these construct-driven paradigms run the risk of perpetuating a declining emphasis in the world of neuropsychological assessment. That said, there are a number of VEs reviewed that meet the standards laid out by Burgess et al. (2006) and emphasise a function-led approach to assessing executive functions using simulated environments.

ISSUES RELATED TO ADOPTION OF VIRTUAL ENVIRONMENTS

Need for VE-based neuropsychological assessments to be sufficiently standardised

While the use of function-led VEs for neuropsychological assessment is an emerging area of application, adoption will require substantial research and development to establish acceptable psychometric properties and clinical utility. An important resolution to clinical heterogeneity of outcome measures in VE research is the standardisation of outcomes and the measures used to assess these outcomes. The selection of outcome measures for standardisation need to be relevant to the patient's treatment and health status, as well as psychometrically sound. Furthermore, VE-based measures need to be fully validated. Findings from VEs need to be interpreted with caution, given that some virtual reality studies do not include control groups, and many are not randomised clinical trials (Parsons & Rizzo, 2008).

Another pressing need among neuropsychologists is the identification of VE-based neuropsychological assessments that reflect relevant underlying cognitive and behavioural capacities for assessments of varying degrees of neuropsychological deficits. VE-based neuropsychological assessments must demonstrate relevance beyond that which is available through simpler means of assessment. As such, there is a specific need for VE-based neuropsychological assessments to be sufficiently standardised within the range and nature of responses available to participants within the VE to allow for reliable measurement. Through the amassing of multiple studies of various clinical populations, VE-based neuropsychological assessments may reveal relevant responses that can be catalogued and defined as measurable factors in a VE. This will require large-scale research trials for validation of measures and development of norms.

Researchers of VE-based neuropsychological assessments have often sought to establish construct validity by demonstrating significant associations between VEs and paper-and-pencil neuropsychological assessments (e.g., Armstrong et al., 2013; Matheis et al., 2007; Parsey & Schmitter-Edgecombe, 2013; Parsons, Courtney, & Dawson, 2013). In the area of function-led assessment, multiple cognitive domains may be involved in the simulation of real-world tasks, and associations with traditional construct-driven tests may be necessarily lower than is typically desired to establish construct validity. In this context, the degree to which a VE-based neuropsychological model using a function-led approach accurately predicts relevant real-world behaviour may be more important than large-magnitude associations with traditional construct-driven paper-and-pencil tests (Renison et al., 2012). Future research should consider this issue in the design of function-led VE-based neuropsychological assessment studies. In addition to these technical issues, clinicians, researchers, and policymakers will need to scrutinise emerging VE-based neuropsychological assessments to ensure adherence to legal, ethical, and human safety guidelines. Finally, the matching of specific technologies to the needs and capacities of the patient will also require careful consideration be neuropsychologists (Schultheis & Rothbaum, 2002). For neuropsychological assessment of clinical populations, increased research is needed. This research will require large participant pools consisting of patients and normal controls from various samples.

Issues for the use of virtual environments in specific patient populations

An additional issue to be considered is how VEs will be experienced by certain patient populations. In paediatrics and geriatrics, human guidance is crucial to ensuring the full comprehension of assessment use and instruction. Geriatric patients, in particular, may find adjusting to virtual platforms particularly difficult (Miller et al., 2014). Ideally, function-led virtual assessments are added to flexible assessment batteries tailored to each individual within the context of the presenting question. Thus, traditional construct-driven measures should not be abandoned. In some circumstances, construct-driven assessments may be more appropriate in terms of assessing a specific construct that is generalisable across environments. For instance, working memory may be more easily assessed by a simple span task. The allure of the virtual assessment lies primarily in enriching stimulus presentation, logging additional variables, and database building rather than the automation of the entire neuropsychological battery and the minimisation of human interaction (Parsey & Schmitter-Edgecombe, 2013; Parsons & Rizzo, 2008). VEs may add to an existing neuropsychological battery when the neuropsychologist is attempting to make accurate predictions about a person's behaviour in the real world. In a VE, the neuropsychologist can measure functional output of constructs within the complexity of a real-world environment. For example, in a VE grocery store, prospective memory may be assessed using a real-world task, such as remembering to pick up a prescription at the pharmacy (Parsons et al., 2013).

Simulation technology may also be problematic for individuals with autism spectrum disorder. Individuals with pronounced sensory issues may find the head-mounted display or even the graphical interface intolerable. Furthermore, individuals with severe psychiatric conditions that cause limited self-awareness, high suggestibility, and/or an altered sense of reality (e.g., hallucinations, delusions) may respond undesirably to immersion in a VE. High-fidelity VEs may confuse these individuals and increase negative behaviours following exposure. Flat-screen presentation of VEs have proven to be an acceptable alternative to full immersion within the environment, and may be more appropriate for certain clinical groups (Attree et al., 1996).

Although VEs have been successfully applied to the study of age differences in spatial navigation among both healthy and demented elderly, VE-based tasks may be complicated by visual, auditory, or motor impairment (Moffat, 2009). In comparison to younger controls, ageing patients may perform more poorly on VE-based tasks simply due to the normative ageing process or lack of experience with computers. Maximum effort should be exerted to ensure equitability in sensorimotor capacities between younger and older adult subjects. A systematic review by Miller et al. (2014) introduced concern regarding the feasibility of home-use VE and gaming systems for physical rehabilitation of older adults. Such systems could be therapeutic to existing physical impairment or could be preventative. A main limitation is the low quality of studies investigating the effectiveness of these systems in older adult populations. Furthermore, some studies cited heightened fall risk, over-exertion, and musculoskeletal irritation. There is a need for more rigorous research methods including more consistent and strenuous reporting of exercise dosages and adherence. Moffat (2009) suggests a number of helpful methodological practices in assessing older adults in research studies of navigation skills: (1) allowing ageing patients to practise and ensure maximum familiarisation with the computer platform; (2)

including measures of computer experience, visual ability, and motor function; and (3) including assessments requiring the same sensorimotor capacities, but not physical navigation.

Verisimilitude of virtual-function led assessments and rapidly changing technology

One challenge neuropsychologists may face in the future is maintaining the verisimilitude of virtual function-led assessments in the face of changing technology. For example, technology has rapidly changed the everyday task of purchasing various items. In a short period of time, the typical modality of payment has changed from cash to cheque to card, with a new method of paying by smart phone becoming more and more popular. Another example is found in in-car entertainment, where radio gave way to compact disks, which were replaced by MP3 players, which are being replaced by voice controlled music selections. Importantly, each of these tasks reflect a slightly different neuropsychological process. One approach to addressing this issue is to develop VEs that allow the neuropsychologist to adjust the graphics, stimuli, and task parameters via an interactive user interface. While function-led VEs are often developed to assess deviations from a healthy control's performance, this approach delivers sub-optimal experiences for above-average or below-average trainees. When this limitation is coupled with the need for enhancements in everyday technologies (e.g., moving from finger-controlled radio tuning to voice-controlled command of music selections), there is a risk that the VE-based assessment will become less relevant. To counteract this, VEs need to be developed that allow for tailoring scenarios to changes in everyday technologies and tractability for customisation of scenarios to the specific needs and abilities of individuals. To achieve this capability, VEs should be developed with scenario-generation systems that allow for modifications of the VE. Again, these virtual scenarios have a "shelf life", and they become less effective as new tactics, techniques, and procedures are adopted to respond to an ever-changing world. Hence the VE must be capable of being re-configured to account for changes to everyday technologies that the scenarios are meant to emulate. This is enabled by easily configurable graphics that can be changed to preserve their verisimilitude (Fox, Arena, & Bailenson, 2009).

Given the above suggestion of flexible VEs with user interfaces that allow for changes to graphics and stimulus presentation, neuropsychological test developers may be able to move beyond heterogeneity of VE-based assessments. One way to ensure optimal test development with minimal variability is to align test development with certain guidelines. Parsons (2011) suggests four key criteria to ensure that VE-based assessments meet high standards of internal, external, and ecological validity: (1) Correspondence: The tasks performed within VEs should correspond to the pertinent aspects of real-world activities and environments. (2) Representativeness: The tasks developed should be representative of persons who are performing the tasks. In addition to consideration of tasks and outcome measures, it is also important to consider subject populations. (3) Expedience: The tasks should have practical consequences on real-world functioning and the test results should reflect and predict real-world phenomena. (4) Relevance: The tasks need to be relevant to both the neurocognitive domains and the relations between these domains that result from interaction within the real world.

A further issue to keep in mind is that the plethora of VEs that may emerge to assess the same construct/function may result in a redundancy of measures—each with a different software and hardware platform. There is a need for the development of systemic methods to minimise this variability. The development of systemised development standards would be very helpful in offering a standard way (including a checklist and flow diagram) to improve the quality of such efforts. Furthermore, it ensures that readers have the basic information necessary to evaluate the quality of the VE. Recently, a joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology was published that sets forth their position on appropriate standards and conventions for computerised neuropsychological assessment devices (Bauer et al., 2012). For VE-based neuropsychological assessments to move forward, similar standards should be developed to address: (1) hardware/software platforms (e.g., Unity, Unreal game engines) for VEs; (2) data security; (3) psychometric development issues; (4) population-specific issues (cultural, experiential, and disability factors); (5) use of VEs and reporting services; (6) quality assurance for response validity and effort in the VE; (7) marketing and performance claims made by developers of VEs; and (8) standardised guidelines for administration and interpretation of performance in VEs.

CONCLUSIONS

Given the above, it is possible that VE-based neuropsychological assessments developed as function-led assessments can meaningfully inform a neuropsychologist's predictive statements about a patient's real-world functioning. We acknowledge that not considering these aspects may lead to our missing of decrements in many aspects of cognition that are critical to competence in everyday life. We reviewed some progress that has been made in various areas of VE-based neuropsychological assessment. Herein, we proffered a general scientific approach to the neuropsychology of executive function that stresses the importance of analysing the demands made by real-world situations and then trying to mimic them in the lab. To the extent that our approach is correct, we hope to encourage VE-based neuropsychologists to shift from batteries that are construct driven only to batteries that include VE-based function-led assessments as a starting point for both the development of new and better clinical tests of executive function, and also for basic neuropsychological investigations.

REFERENCES

Adams, R., Finn, P., Moes, E., Flannery, K., & Rizzo, A. (2009). Distractibility in attention/deficit/hyperactivity disorder (ADHD): The virtual reality classroom. *Child Neuropsychology*, *15*, 120–135.

Alderman, N., Burgess, P. W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*, *9*, 31–44.

Allen, A. J., Shashwath, A. M., Skudlarski, P., Calhoun, V., Astur, R., Ruopp, K. C., & Pearlson, G. D. (2009). Effects of alcohol on performance of distraction task during simulated driving. *Alcoholism: Clinical and Experimental Research*, *33*, 617–625.

Armstrong, C., Reger, G., Edwards, J., Rizzo, A., Courtney, C., & Parsons, T. (2013). Validity of the virtual reality stroop task (VRST) in active duty military. *Journal of Clinical and Experimental Neuropsychology*, *35*, 113–123.

Asimakopulos, J., Boychuck, Z., Sondergaard, D., Poulin, V., Ménard, I., & Korner-Bitensky, N. (2012). Assessing executive function in relation to fitness to drive: A review of tools and their ability to predict safe driving. *Australian Occupational Therapy Journal*, *59*, 402–427.

Attree, E. A., Brooks, B. M., Rose, F. D., Andrews, T. K., Leadbetter, A. G., & Clifford, B. R. (1996, July). Memory processes and virtual environments: I can't remember what was there, but I can remember how I got there. Implications for people with disabilities. In *ECDVRAT: 1st European Conference on Disability, Virtual Reality and Associated Technologies. Reading*, UK (Vol. 118).

Avis, K., Gamble, K., & Schwebel, D. (2014). Does excessive daytime sleepiness affect children's pedestrian safety? *SLEEP*, *37*, 283–287.

Bailey, P. E., Henry, J. D., Rendell, P. G., Phillips, L. H., & Kliegel, M. (2010). Dismantling the "age–prospective memory paradox": The classic laboratory paradigm simulated in a naturalistic setting. *The Quarterly Journal of Experimental Psychology*, *63*(4), 646–652.

Barkley, R. A., Anderson, D. L., & Kruesi, M. (2007). A pilot study of the effects of atomoxetine on driving performance in adults with ADHD. *Journal of Attention Disorders*, *10*, 306–316.

Barkley, R., Murphy, K., O'connell, T., Anderson, D., & Connor, D. (2006). Effects of two doses of alcohol on simulator driving performance in adults with attention-deficit/hyperactivity disorder. *Neuropsychology*, *20*, 77–87.

Barkley, R. A., Murphy, K. R., O'Connell, T., & Connor, D. F. (2005). Effects of two doses of methylphenidate on simulator driving performance in adults with attention deficit hyperactivity disorder. *Journal of Safety Research*, *36*, 121–131.

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*, 177–196.

Bedard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of driving performance using a simulator protocol: Validity and reproducibility. *American Journal of Occupational Therapy*, *64*, 336–340.

Bioulac, S., Lallemand, S., Rizzo, A., Philip, P., Fabrigoule, C., & Bouvard, M. P. (2012). Impact of time on task on ADHD patients' performances in a virtual classroom. *European Journal of Paediatric Neurology*, *16*, 514–521.

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, *12*, 752–762.

Bottari, C., Dassa, C., Rainville, C., & Dutil, E. (2009). The criterion-related validity of the IADL profile with measures of executive functions, indices of trauma severity and sociodemographic characteristics. *Brain Injury*, *23*, 322–335.

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *The Journal of the International Neuropsychological Society*, *4*, 547–558.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., MA COATES, L. A. U. R. E., Dawson, D. R., . . . Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, *12*, 194–209.

Burgess, P. W., & Simons, J. S. (2005). Theories of frontal lobe executive function: Clinical applications. In P. W. Halligan & D. T. Wade (Eds.), *Effectiveness of Rehabilitation for Cognitive Deficits* (pp. 211–231). Oxford: Oxford University Press.

Buxbaum, L., Dawson, A., & Linsley, D. (2012). Reliability and validity of the virtual reality lateralized attention test in assessing hemispatial neglect in right-hemisphere stroke. *Neuropsychology*, *26*, 430–441.

Calhoun, V. D., & Pearlson, G. D. (2012). A selective review of simulated driving studies: Combining naturalistic and hybrid paradigms, analysis approaches, and future directions. *NeuroImage*, *59*, 25–35.

Campbell, Z., Zakzanis, K. K., Jovanovski, D., Joordens, S., Mraz, R., & Graham, S. J. (2009). Utilizing virtual reality to improve the ecological validity of clinical neuropsychology: An FMRI case study elucidating the neural basis of planning by comparing the tower of London with a three-dimensional navigation task. *Applied Neuropsychology*, *16*, 295–306.

Cao, X., Douguet, A. S., Fuchs, P., & Klinger, E. (2010). Designing an ecological virtual task in the context of executive functions: A preliminary study. In P. Sharkey & J. Sanchez (Eds.), *Proceedings of the 8th International Conference Disability, Virtual Reality and Associated Technologies* (pp. 71–78). Chile: Vĩna del Mar/Valparaiso.

Carelli, L., Morganti, F., Weiss, P. L., Kizony, R., & Riva, G. (2008). A virtual reality paradigm for the assessment and rehabilitation of executive function deficits post stroke: Feasibility study. *Virtual Rehabilitation*, *2008*, 99–104.

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*, 201–216.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*, 181–197.

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*, *21*, 217–227.

Cipresso, P., La Paglia, F., La Cascia, C., Riva, G., Albani, G., & La Barbera, D. (2013). Break in volition: A virtual reality study in patients with obsessive-compulsive disorder. *Experimental Brain Research*, *229*, 443–449.

Clancy, T. A., Rucklidge, J. J., & Owen, D. (2006). Road-Crossing safety in virtual reality: A comparison of adolescents with and without ADHD. *Journal of Clinical Child & Adolescent Psychology*, *35*, 203–215.

Cook, L., Hanten, G., Orsten, K., Chapman, S., Li, X., Wilde, E., Schnelle, K. P., Levin, H. (2013). Effects of moderate to severe traumatic brain injury on anticipating consequences of actions in adults: A preliminary study. *Journal of the International Neuropsychological Society*, *19*, 508–517.

Cox, D., Moore, M., Burket, R., Merkel, R., Mikami, A., & Kovatchev, B. (2008). Rebound effects with long-acting amphetamine or methylphenidate stimulant medication preparations among adolescent male drivers with attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology*, *18*, 1–10.

Cuberos-Urbano, G., Caracuel, A., Vilar-López, R., Valls-Serrano, C., Bateman, A., & Verdejo-García, A. (2013). Ecological validity of the multiple errands test using predictive models of dysexecutive problems in everyday life. *Journal of Clinical and Experimental Neuropsychology*, *35*, 329–336.

Davis, A., Avis, K., & Schwebel, D. (2013). The effects of acute sleep restriction on adolescents' pedestrian safety in a virtual environment. *Journal of Adolescent Health*, *53*, 785–790.

Dawson, D. R., Anderson, N. D., Burgess, P., Cooper, E., Krpan, K. M., & Stuss, D. T. (2009). Further development of the multiple errands test: Standardized scoring, reliability, and ecological validity for the baycrest version. *Archives of Physical Medicine and Rehabilitation*, *90*, S41–S51.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168.

Diaz-Orueta, U., Garcia-Lopez, C., Crespo-Eguilaz, N., Sanchez-Carpintero, R., Climent, G., & Narbona, J. (2014). AULA virtual reality test as an attention measure: Convergent validity with conners' continuous performance test. *Child Neuropsychology*, *20*, 328–342.

Elkind, J. S., Rubin, E., Rosenthal, S., Skoff, B., & Prather, P. (2001). A simulated reality scenario compared with the computerized Wisconsin card sorting test: An analysis of preliminary results. *CyberPsychology & Behavior*, *4*, 489–496.

Fox, J., Arena, D., & Bailenson, J. N. (2009). Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology*, *21*, 95–113.

Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 91–112). Boca Raton, FL: St Lucie Press.

Goldstein, G. (1996). Functional considerations in neuropsychology. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 75–89). Delray Beach, FL: GR Press/St. Lucie Press.

Henry, M., Joyal, C. C., & Nolin, P. (2012). Development and initial assessment of a new paradigm for assessing cognitive and motor inhibition: The bimodal virtual-reality stroop. *Journal of Neuroscience Methods*, *210*, 125–131.

Iriate, Y., Diaz-Orueta, U., Cueto, E., Irazustabarrena, P., Banterla, F., & Climent, G. (2012). AULA – Advanced virtual reality tool for the assessment of attention: Normative study in Spain. *Journal of Attention Disorders*, *20*, 1–27.

Jansari, A. S., Froggatt, D., Edginton, T., & Dawkins, L. (2013). Investigating the impact of nicotine on executive functions using a novel virtual reality assessment. *Addiction*, *108*, 977–984.

Josman, N., Kizony, R., Hof, E., Goldenberg, K., Weiss, P., & Klinger, E. (2014). Using the virtual action planning-supermarket for evaluating executive functions in people with stroke. *Journal of Stroke and Cerebrovascular Diseases*, *23*, 879–887.

Josman, N., Schenirderman, A. E., Klinger, E., & Shevil, E. (2009). Using virtual reality to evaluate executive functions among persons with schizophrenia: A validity study. *Schizophrenia Research*, *115*, 270–277.

Jovanovski, D., Zakzanis, K., Campbell, Z., Erb, S., & Nussbaum, D. (2012). Development of a novel, ecologically oriented virtual reality measure of executive function: The multitasking in the city test. *Applied Neuropsychology: Adult*, *19*, 171–182.

Jovanovski, D., Zakzanis, K., Ruttan, L., Campbell, Z., Erb, S., & Nussbaum, D. (2012). Ecologically valid assessment of executive dysfunction using a novel virtual reality task in patients with acquired brain injury. *Applied Neuropsychology*, *19*, 207–220.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*, 213–233.

Knight, C., Alderman, N., & Burgess, P. (2002). Development of a simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation*, *12*, 231–255.

Knouse, L. E., Bagwell, C. L., Barkley, R. A., & Murphy, K. R. (2005). Accuracy of self-evaluation in adults with ADHD: Evidence from a driving study. *Journal of Attention Disorders*, *8*, 221–234.

Lalonde, G., Henry, M., Drouin-Germain, A., Nolin, P., & Beauchamp, M. H. (2013). Assessment of executive function in adolescence: A comparison of traditional and virtual reality tools. *Journal of Neuroscience Methods*, *219*, 76–82.

Lamberts, K. F., Evans, J. J., & Spikman, J. M. (2010). A real-life, ecologically valid test of executive functioning: The executive secretarial task. *Journal of Clinical and Experimental Neuropsychology*, *32*(1), 56–65.

Law, A. S., Trawley, S. L., Brown, L. A., Stephens, A. N., & Logie, R. H. (2013). The impact of working memory load on task execution and online plan adjustment during multitasking in a virtual environment. *The Quarterly Journal of Experimental Psychology*, *66*, 1241–1258.

Lee, H. C., Cameron, D., & Lee, A. G. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis and Prevention*, *35*, 797–803.

Lee, H. C., Lee, A. H., Cameron, D., & Li-Tsang, C. (2003). Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. *Journal of Safety Research*, *34*, 453–459.

Levine, B., Dawson, D., Boutet, I., Schwartz, M. L., & Stuss, D. T. (2000). Assessment of strategic self-regulation in traumatic brain injury: Its relationship to injury severity and psychosocial outcome. *Neuropsychology*, *14*, 491–500.

Liu, L., Miyazaki, M., & Watson, B. (1999). Norms and validity of the drivr: A virtual reality driving assessment for persons with head injuries. *Cyberpsychology & Behavior*, *2*, 53–67.

Logie, R. H., Trawley, S., & Law, A. (2011). Multitasking: Multiple, domain-specific cognitive functions in a virtual environment. *Memory & Cognition*, *39*, 1561–1574.

Manchester, D., Priestley, N., & Howard, J. (2004). The assessment of executive functions: Coming out of the office. *Brain Injury*, *18*, 1067–1081.

Matheis, R. J., Schultheis, M. T., Tiersky, L. A., DeLuca, J., Millis, S. R., & Rizzo, A. (2007). Is learning and memory different in a virtual environment? *The Clinical Neuropsychologist*, *21*, 146–161.

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Sala, S. D., & Milne, A. B. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence: Teleoperators and Virtual Environments*, *10*, 375–383.

Miller, K. J., Adair, B. S., Pearce, A. J., Said, C. M., Ozanne, E., & Morris, M. M. (2014). Effectiveness and feasibility of virtual reality and gaming system use at home by older adults for enabling physical activity to improve health-related domains: A systematic review. *Age and Ageing*, *43*, 188–195.

Milleville-Pennel, I., Pothier, J., Hoc, J. M., & Mathe, J. F. (2010). Consequences of cognitive impairments following traumatic brain injury: Pilot study on visual exploration while driving. *Brain Injury*, *24*, 678–691.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology (Chicago)*, *9*, 90–100.

Moffat, S. D. (2009). Aging and spatial navigation: What do we know and where do we go? *Neuropsychology Review*, *19*, 478–489.

Montgomery, C., Ashmore, K., & Jansari, A. (2011). The effects of a modest dose of alcohol on executive functioning and prospective memory. *Human Psychopharmacology: Clinical and Experimental*, *26*, 208–215.

Morris, R. G., Kotitsa, M., Bramham, J., Brooks, B., & Rose, F. D. (2002). Virtual reality investigation of strategy formation, rule breaking and prospective memory in patients with focal prefrontal neurosurgical lesions. In P. Sharkey, C. S. Lányi, & P. Standen (Eds.), *Proceedings of the 4th International Conference on Disability, Virtual Reality & Associated Technologies, Veszprém, Hungary* (pp. 101–108). Reading, UK: The University of Reading.

Nagamatsu, L., Voss, M., Neider, M., Gaspar, J., Handy, T., Kramer, A., & Liu-Ambrose, T. (2011). Increased cognitive load leads to impaired mobility decisions in seniors at risk for falls. *Psychology and Aging*, *26*, 253–259.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18). New York: Plenum.

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

Parsons, T. D. (2011). Neuropsychological assessment using virtual environments: Enhanced assessment technology for improved ecological validity. In *Advanced computational intelligence paradigms in healthcare 6. Virtual reality in psychotherapy, rehabilitation, and assessment* (pp. 271–289). Berlin Heidelberg: Springer.

Parsons, T. D., Bowerly, T., Buckwalter, J. G., & Rizzo, A. A. (2007). A controlled clinical comparison of attention performance in children with ADHD in a virtual reality classroom compared to standard neuropsychological methods. *Child Neuropsychology*, *13*, 363–381.

Parsons, T. D., & Courtney, C. (2014). An initial validation of the virtual reality paced auditory serial addition test in a college sample. *Journal of Neuroscience Methods*, *222*, 15–23.

Parsons, T. D., Courtney, C. G., & Dawson, M. E. (2013). Virtual reality stroop task for assessment of supervisory attentional processing. *Journal of Clinical and Experimental Neuropsychology*, *35*(8), 812–826.

Parsons, T. D., McPherson, S., & Interrante, V. (2013, March). Enhancing neurocognitive assessment using immersive virtual reality. In *1st Workshop on Virtual and Augmented Assistive Technology (VAAT)* (pp. 27–34). Orlando, FL: IEEE.

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*, 250–261.

Parsons, T. D., & Rizzo, A. A. (2008). Initial validation of a virtual environment for assessment of memory functioning: Virtual reality cognitive performance assessment test. *Cyberpsychology and Behavior*, *11*, 17–25.

Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social Neuroscience*, *9*(1), 94–107.

Pedroli, E., Cipresso, P., Serino, S., Pallavicini, F., Albani, G., & Riva, G. (2013). Virtual multiple errands test: Reliability, usability and possible applications. *Annual Review of Cybertherapy and Telemedicine*, *191*, 48–52.

Pollak, Y., Shomaly, H. B., Weiss, P. L., Rizzo, A. A., & Gross-Tsur, V. (2010). Methylphenidate effect in children with ADHD can be measured by an ecologically valid continuous performance test embedded in virtual reality. *CNS Spectrums*, *15*, 125–130.

Pollak, Y., Weiss, P. L., Rizzo, A. A., Weizer, M., Shriki, L., Shalev, R. S., & Gross-Tsur, V. (2009). The utility of a continuous performance test embedded in virtual reality in measuring ADHD-related deficits. *Journal of Developmental & Behavioral Pediatrics*, *30*, 2–6.

Pugnetti, L., Mendozzi, L., Attree, E. A., Barbieri, E., Brooks, B. M., Cazzullo, C. L., . . . Psychol, C. (1998). Probing memory and executive functions with virtual reality: Past and present studies. *CyberPsychology and Behavior*, *1*, 151–161.

Pugnetti, L., Mendozzi, L., Motta, A., Cattaneo, A., Barbieri, E., & Brancotti, S. (1995). Evaluation and retraining of adults' cognitive impairments: Which role for virtual reality technology? *Computers in Biology and Medicine*, *25*, 213–227.

Rabin, L. A., Burton, L. A., & Barr, W. B. (2007). Utilization rates of ecologically oriented instruments among clinical neuropsychologists. *The Clinical Neuropsychologist*, *21*(5), 727–743.

Rand, D., Rukan, S., Weiss, P., & Katz, N. (2009). Validation of the virtual MET as an assessment tool for executive functions. *Neuropsychological Rehabilitation*, *19*, 583–602.

Raspelli, S., Carelli, L., Morganti, F., Pedroli, E., Cipresso, P., Poletti, B., . . . Riva, G. (2012). Validating the Neuro VR-based virtual version of the multiple errands test: Preliminary results. *Presence: Teleoperators and Virtual Environments*, *21*, 31–42.

Renison, B., Ponsford, J., Testa, R., Richardson, B., & Brownfield, K. (2012). The ecological and construct validity of a newly developed measure of executive function: The virtual library task. *Journal of the International Neuropsychological Society*, *18*, 440–450.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). "Oops!": Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*, 747–758.

Saidel-Goley, I., Albiero, E., & Flannery, K. (2012). An evaluation of nonclinical dissociation utilizing a virtual environment shows enhanced working memory and attention. *Cyberpsychology, Behavior, and Social Networking*, *15*, 112–116.

Sauzéon, H., N'Kaoua, B., Arvind Pala, P., Taillade, M., Auriacombe, S., & Guitton, P. (2014). Everyday-like memory for objects in ageing and Alzheimer's disease assessed in a visually complex environment: The role of executive functioning and episodic memory. *Journal of Neuropsychology*. doi:10.1111/jnp.12055

Sbordone, R. J. (2008). Ecological validity of neuropsychological testing: Critical issues. In A. M. Horton & D. Wedding (Eds.), *The neuropsychology handbook* (3rd ed., pp. 367–394). New York: Springer.

Schatz, P., & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. *Journal of Head Trauma Rehabilitation*, *17*, 395–410.

Schultheis, M. T., Himelstein, J., & Rizzo, A. A. (2002). Virtual reality and neuropsychology: Upgrading the current tools. *The Journal of Head Trauma Rehabilitation*, *17*, 378–394.

Schultheis, M. T., & Mourant, R. (2001). Virtual reality and driving: The road to better assessment for cognitively impaired populations. *Presence: Teleoperators and Virtual Environments*, *10*, 431–439.

Schultheis, M. T., Rebimbas, J., Mourant, R., & Millis, S. R. (2007). Examining the usability of a virtual reality driving simulator. *Assistive Technologies*, *19*, 1–10.

Schultheis, M. T., & Rothbaum, B. O. (2002). Ethical issues for the use of virtual reality in the psychological sciences. In S. Bush & M. Drexler (Eds.), *Ethical issues in clinical neuropsychology* (pp. 243–280). Lisse, NL: Swets & Zeitlinger.

Shallice, T. I. M., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*, 727–741.

Spooner, D. M., & Pachana, N. A. (2006). Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology*, *21*, 327–337.

Stuss, D. T., Benson, D. F., Kaplan, E. F., Weir, W. S., Naeser, M. A., Lieberman, I., . . . Ferrill, D. (1983). The involvement of orbitofrontal cerebrum in cognitive tasks. *Neuropsychologia*, *21*, 235–248.

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of the New York Academy of Sciences*, *769*, 191–212.

Sweeney, S., Kersel, D., Morris, R. G., Manly, T., & Evans, J. J. (2010). The sensitivity of a virtual reality task to planning and prospective memory impairments: Group differences and the efficacy of periodic alerts on performance. *Neuropsychological Rehabilitation*, *20*, 239–263.

Toet, A., & van Schaik, M. G. (2013). Visual attention for a desktop virtual environment with ambient scent. *Frontiers in Psychology*, *4*, 1–11.

Wald, J., & Liu, L. (2001). Psychometric properties of the driVR: A virtual reality driving assessment. *Studies in Health Technology Information*, *81*, 564–566.

Wald, J. L., Liu, L., & Reil, S. (2000). Concurrent validity of a virtual reality driving assessment in persons with brain injury. *Cyberpsychology & Behavior*, *3*, 643–654.

Weigl, E. (1927). Zur psychologie sogenannter abstraktionsprozesse. *Zeitschrift fur Psychologie*, *103*, 2–45. [Translated by M. Rioch and reprinted as: On the psychology of so-called processes of abstraction. *Journal of Abnormal and Social Psychology*, 36, 3–33, 1948].

Werner, P., Rabinowitz, S., Klinger, E., Korczyn, A., & Josman, N. (2009). Use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, *27*, 301–309.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome (BADS)*. Bury St Edmunds: Thames Valley.