Virtual Reality for Psychological Assessment in Clinical Practice

Thomas D. Parsons and Amanda S. Phillips
University of North Texas

Clinical psychologists are often asked to make recommendations regarding a client’s ability to function in everyday activities (e.g., work, classroom, or shopping). Common approaches include a combination of paper-and-pencil (including some computer automated) assessments, behavioral observations of the client’s performance during testing, and self-report measures about their perceived deficits. From this combination of assessments, observations, and self-reports, the psychologist is expected to make predictions about the client’s ability to return to the classroom, return to work, and successfully complete other activities of daily living. While there are advantages to this approach, there are also some shortcomings—perhaps most notable is the lack of ecological validity. Recent advances in virtual reality technologies allow for enhanced computational capacities for administration efficiency, stimulus presentation, automated logging of responses, and data analytic processing. These virtual environments allow for controlled presentations of emotionally engaging background narratives to enhance affective experience and social interactions. Within this context virtual reality can allow psychologists to offer safe, repeatable, and diversifiable assessments of real world functioning. Although there are a number of purported advantages of virtual reality technologies, there is still a need for establishing the psychometric properties of virtual reality assessments and interventions. This review investigates the advantages and challenges inherent in the application of virtual reality technologies to psychological assessments and interventions.

Keywords: virtual reality, neuropsychology, psychological assessment, ecological validity

Clinical psychologists are increasingly being asked to make prescriptive statements about every-day functioning. Unfortunately, results from many psychological tests are not easily generalizable to real-world functioning. Common approaches include a combination of history taking, self-reports, paper-and-pencil cognitive assessments, and the psychologist’s observations of the client’s behavior. From this combination, the psychologist is expected to make predictions about the client’s ability to return to the classroom, return to work, and successfully complete other activities of daily living. The limitations inherent in this process have led to increasing calls for assessment methods that provide more generalizable data about client functioning (Burgess et al., 2006; Jurado & Rosselli, 2007). Virtual reality (VR) assessments have been developed to provide an enhanced understanding of client functioning in activities of daily living (Campbell et al., 2009; Matheis et al., 2007).

While VR assessments are often presented as tools for neurocognitive assessment in research settings, this article aims to provide an introduction to VR assessment for clinical practice. First, current assessment methods are compared with assessment in virtual environments. Next, examples of VR assessments with clinical case examples are provided. Finally, considerations for the adoption of VR technologies in clinical practice are explored. We propose that the addition of VR to current psychological assessment batteries can improve the generalizability of test results and increase the utility and rele-
vance of psychologists’ recommendations to clients.

**Self-Reports and Behavioral Observations**

Although traditional assessment methods provide valuable information, there are a number of limitations in using this approach (see Table 1). Determining a client’s functional capabilities requires precise control over the environment and the ability to adjust the potency or frequency of stimuli (White et al., 2014). This control is difficult to ensure in the traditional assessment environment. Given that psychological tests are typically administered one-on-one in a controlled environment; the psychologist may not receive a clear picture of the client’s cognitive functions in everyday activities. Although psychologists often use information from parent and teacher reports to get an idea of the patient’s everyday functioning, studies indicate that the agreement between parents and teachers is modest. For example, the concordance between parents and teachers on diagnosing attention-deficit-hyperactivity disorder (ADHD) varies from .30 to .50 depending on the behavioral dimensions being rated (Biederman, Faraone, Milberger, & Doyle, 1993; Biederman, Keenan, & Faraone, 1990; Mitsis, McKay, Schulz, Newcorn, & Halperin, 2000; Newcorn et al., 1994; de Nijs et al., 2004). Additionally, there is often not a strong overlap between these rating scales and standard tests of cognitive functioning, suggesting that these assessments may be reflective of different aspects of behavior (see Table 1 for a list of potential confounds).

**Normative Comparison of Cognitive Performance**

For many psychologists, the way to get beyond these limitations is to emphasize a normative comparison approach to psychological assessment. For example, a psychologist can give a continuous performance test (CPT) to assess a cognitive construct (e.g., attentional processing). Results from the CPT reveal the client’s performance in areas of inattentiveness, impulsivity, sustained attention, and vigilance. The psychologist can compare the client’s CPT results with reference groups to determine whether a client is performing as would be

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Comparison of the Advantages and Disadvantages of Traditional Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>Controlled environment limits the generalizability of results.</td>
</tr>
<tr>
<td>Self-reports</td>
<td>Allows identification of areas of client functioning that have been impacted or preserved.</td>
</tr>
<tr>
<td></td>
<td>Provides information about the client’s perceptions of their performance.</td>
</tr>
<tr>
<td></td>
<td>Third-party reports provide data about client functioning in various environments.</td>
</tr>
<tr>
<td>Behavioral observation</td>
<td>Provides insights into how the client may react when faced with difficult tasks.</td>
</tr>
<tr>
<td>Normative comparison</td>
<td>Allows for measurement of expected performance.</td>
</tr>
<tr>
<td></td>
<td>Quantifies discrepancies in expected performance.</td>
</tr>
<tr>
<td></td>
<td>Tests are often measure abstract constructs rather than functioning.</td>
</tr>
<tr>
<td></td>
<td>Many cognitive tests were developed for use with normal populations, not for functional assessment.</td>
</tr>
<tr>
<td></td>
<td>Tests may not be representative of real-world situations.</td>
</tr>
<tr>
<td></td>
<td>Tests lack generalizability to real-world situations.</td>
</tr>
<tr>
<td></td>
<td>Test sensitivity and potency cannot be altered.</td>
</tr>
<tr>
<td></td>
<td>Time-consuming and expensive administration.</td>
</tr>
</tbody>
</table>
expected given their achievements and educational attainment.

An unfortunate limitation of this approach is that the tests used are typically based upon construct-driven assessments that do little to replicate the client’s everyday activities. Furthermore, many of the tests commonly used today were never meant for clinical assessments. Burgess and colleagues (2006) point out that cognitive construct measures like the Tower of London and the Wisconsin Card Sorting Test (WCST) were not originally designed to be used as clinical measures. Instead, these measures were found to be useful tools for cognitive assessment among normal populations and then later found their way into the clinical realm to aid in assessing constructs that are important to carrying out real-world activities. This approach forces the psychologist to rely on measures that were designed for purposes other than predictions of real-world functioning. Burgess and colleagues (2006) argued that we need assessments capable of broadening our understanding about the ways in which the brain enables persons to interact with their environment and organize everyday activities. Although many cognitive tests do give us some insight into the client’s everyday performance, they do not provide direct knowledge about shortcomings in the functional capabilities of the client, which limits the accuracy and utility of the psychologist’s recommendations (Chaytor & Schmitter-Edgecombe, 2003; Manchester, Priestley, & Jackson, 2004).

Virtual Environments

One potential answer to the issues raised above is the addition of virtual environments to the psychologists’ cognitive assessment battery, which allow clients to become immersed within a computer-generated simulation (Campbell et al., 2009; Matheis et al., 2007). Potential virtual environment use in assessment and rehabilitation of human cognitive processes is becoming recognized as technology advances, and has a number of advantages (see Table 2). While a complete listing of all the virtual environment-based psychological assessments is outside the scope of this article, we provide examples of virtual environments with case examples.

Virtual Classroom

Various virtual classrooms are being validated for assessment of supervisory attentional processing. Virtual Classrooms are simulation environments that were designed to assess potential attention deficits by embedding cognitive tasks (e.g., CPT; Stroop) into a virtual classroom environment (Díaz-Orueta et al., 2013; Iriarte et al., 2016; Parsons et al., 2007; Rizzo et al., 2006; see Table 3). In these virtual classrooms the participant is seated at one of the desks and is surrounded by desks, children, a teacher, and a white board much like they would be in a real-world classroom. Various cognitive tasks can be presented on the whiteboard in the front of the room and the participant performs a task (e.g., Stroop or CPT) with auditory (e.g., airplane passing overhead, a voice from the intercom, and the bell ringing) and visual (e.g., children passing notes, a child raising his hand, the teacher answering the classroom door, and principal entering the room) distractors in the background. The number and frequency of these distractors can be adjusted based on age, grade level, or other

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison of the Advantages and Disadvantages of Assessment in a Virtual Environment</strong></td>
</tr>
<tr>
<td>Advantages</td>
</tr>
<tr>
<td>Precise and consistent presentation of stimuli.</td>
</tr>
<tr>
<td>Clinician control of dynamic stimuli.</td>
</tr>
<tr>
<td>Greater ecological validity.</td>
</tr>
<tr>
<td>Availability of function led assessments.</td>
</tr>
<tr>
<td>Programs administer, record, score, and analyze data.</td>
</tr>
<tr>
<td>Environmental stimuli and test sensitivity can be adjusted.</td>
</tr>
<tr>
<td>Environment can be adjusted to reflect clients’ everyday surroundings.</td>
</tr>
</tbody>
</table>
Table 3
Recent Construct-Driven Studies (Within Past 10 Years) Using a Virtual Classroom Environment

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional Tests</th>
<th>Results</th>
</tr>
</thead>
</table>
| Adams et al. (2009) | *N* = 35 boys ages 8–14 years; 19 participants with ADHD were compared to 16 age-matched controls. | Comparison of participant performance on the Continuous Performance Test with and without the Virtual Classroom. | • Behavior Assessment System for Children (BASC)  
• VIGIL Continuous Performance Test | • Greater specificity was found for Virtual CPT.  
• Although differences between the two groups were not significant, a strong trend was observed for correct target identification and commission errors. |
| Bioulac et al. (2012) | *N* = 36 boys ages 7–10 years; 20 participants with ADHD were compared to 16 age-matched controls. | ADHD and controls children were first tested with the traditional computerized CPT. After 10 min they were tested with the virtual CPT. | • Continuous Performance Test (CPT II)  
• Conners' parents rating scale  
• Child Behavior Check List  
• State Trait Inventory Anxiety Inventory | • ADHD participants showed a significant performance decrement; decrease of correct hits; and increased reaction time.  
• ADHD children performed worse than controls on both the Virtual CPT and the traditional computerized CPT. |
| Díaz-Orueta et al. (2014) | *N* = 57 participants with ADHD between the ages of 6–16 years (26.3% female). | Convergent validity study between both a Virtual CPT and traditional computerized CPT. Compared children undergoing medical treatment with a non-medicated group | • Conners’ Continuous Performance Test  
• WISC-III (Selected subtests) | • Both Virtual CPT and the traditional computerized CPT showed significant correlations.  
• The Virtual CPT (but not Conners’ CPT) was able to differentiate between ADHD children with and without pharmacological treatment (inattention, impulsivity, processing speed, motor activity, and attention focus). |
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional Tests</th>
<th>Results</th>
</tr>
</thead>
</table>
| Gilboa et al. (2011)   | $N = 54.29$ with Neurofibromatosis type 1 (NF1); 69% female; mean age 12.2. 25 controls 72% female; mean age = 12.2. | Comparison of Virtual CPT and the traditional tests. Cross sectional design. | Conners' Parent Rating Scales-Revised | • Significant differences between the NF1 and controls on omission errors and commission errors in the Virtual CPT.  
• Poorer performance by NF1 children.  
• Significant correlations between number of targets correctly identified, the number of commission errors, and reaction time. |
Wechsler Abbreviated Scale of Intelligence (Matrix Reasoning and Vocabulary)  
Conners Parent Rating Scales-Revised | • Significant between group differences for number of targets correctly identified in the Virtual CPT.  
• 45% of the children with ABI suffered marked deficits in sustained attention on the Virtual CPT.  
• Attentional performance was found to be related to age, age at injury/diagnosis and treatment. |
| Iriarte et al. (2016)  | $N = 1282$ children ages 6 to 16.                                      | A normative study. Single application with descriptive design. | None | • Results were clustered into different categories for posterior analysis.  
• Differences by age and gender were analyzed, resulting in 14 groups, 7 per sex group.  
• Differences between visual and auditory attention were also obtained.  
(table continues) |
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional Tests</th>
<th>Results</th>
</tr>
</thead>
</table>
• Behavior Rating Inventory of Executive Function  
• Child Behavior Checklist | • VR-Stroop task correlated with D-KEFS and BRIEF.  
• Performance on the VR-Stroop task was correlated with paper-pencil Stroop task.  
• VR-Stroop more accurately reflected everyday behavioral functioning. |
| Nolin et al. (2009)   | $N = 8$ children with acquired brain injury, ages 8 to 12 years. | Repeated measures comparisons.                                                  | • VIGIL Continuous Performance Test                                                | • No difference between the Virtual CPT and the traditional computerized CPT on total of omissions.  
• Significantly more commissions and longer reaction times in the Virtual CPT. |
| Nolin et al. (2012)   | $N = 50$ 25 sports-concussed and 25 matched control adolescents. | Comparison of the traditional CPT and Virtual CPT was counterbalanced across participants. | • VIGIL Continuous Performance Test                                                | • Virtual CPT showed greater sensitivity to the subtle effects of sports concussion.  
• The sports concussion group reported more symptoms of cybersickness than the control group. |
| Parsons et al. (2007) | $N = 20$ 10 boys diagnosed with ADHD and 10 matched controls. | Intergroup comparison of participants with ADHD and normal controls.            | • SWAN Behavior Checklist  
• Conners’ CPT II  
• Stroop  
• Trail Making tests  
• NEPSY (Visual attention, design fluency, verbal fluency)  
• WISC-III (Digit Span, Coding, Arithmetic, Vocabulary)  
• Judgement of Line Orientation | • ADHD group exhibited more omission errors, commission errors, and overall body movement in the Virtual CPT.  
• ADHD group was more impacted by distraction in the Virtual CPT.  
• Virtual CPT was correlated with traditional ADHD assessment tools, behavior checklist, and traditional computerized CPT. |
Table 3 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional Tests</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parsons and Carlew</td>
<td>Two studies reported: Study 1: 50 undergraduate students (mean age = 20.37; 78% female).</td>
<td>Study 1: Normative study comparing Virtual Stroop to traditional tasks.</td>
<td>• Wechsler Test of Adult Reading</td>
<td>• Virtual Stroop task was correlated with traditional tasks and elicited an interference effect similar to those found in classic Stroop tasks.</td>
</tr>
<tr>
<td>(2016)</td>
<td></td>
<td>Study 2: Cross sectional design.</td>
<td>• Delis-Kaplan Executive Functioning System: Color Word Interference Test</td>
<td>• During the distraction condition of the Virtual Stroop ASD group performance declined.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Stroop task from Automated Neuropsychological Assessment Metrics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Wechsler Abbreviated Scale of Intelligence-Second Edition</td>
<td></td>
</tr>
<tr>
<td>Pollak et al.</td>
<td>N = 37 boys ages 9–17 years, with (n = 20) and without ADHD (n = 17).</td>
<td>Crossover design comparing Virtual Classroom on regular computer screen.</td>
<td>• Test of Variables of Attention (TOVA)</td>
<td>• ADHD group performed less well on all CPT tasks.</td>
</tr>
<tr>
<td>(2009)</td>
<td></td>
<td></td>
<td>• Short Feedback Questionnaire</td>
<td>• Virtual CPT showed effect sizes similar to the TOVA.</td>
</tr>
<tr>
<td>Pollak et al.</td>
<td>N = 27 16 boys and 11 girls, with clinical diagnosis of ADHD.</td>
<td>Double-blind, placebo-controlled, crossover design.</td>
<td>• Metilphenidate (MPH) reduced omission errors to a greater extent on the VR-CPT compared to the no VR-CPT and the TOVA, and decreased other CPT measures on all types of CPT to a similar degree.</td>
<td>• Children rated the VR-CPT as more enjoyable compared to the other types of CPT.</td>
</tr>
<tr>
<td>(2010)</td>
<td></td>
<td></td>
<td>• Virtual Classroom on regular computer screen</td>
<td></td>
</tr>
</tbody>
</table>

Note. ADHD = Attention-deficit/hyperactivity disorder; ASD = Autism Spectrum Disorder; BASC = Behavior Assessment System for Children; CPT = Continuous Performance Test; DKEFS = Delis–Kaplan Executive Function System; BRIEF = Behavior Rating Inventory of Executive Function; NEPSY = abbreviation of the term “neuropsychology”; NF1 = Neurofibromatosis type 1; TOVA = Test of Variables of Attention; VR = Virtual Reality; WISC-III = Wechsler Intelligence Scale for Children.
testing needs. In addition, there are a number of auditory distractors that can be adjusted, such as the sounds of vehicles passing by and ambient classroom noise. Other aspects of the virtual environment can also be adjusted, including seating position of the client, the number of virtual students, and the sex of the teacher (Rizzo et al., 2006).

The case of G. T. A psychologist receives a referral for assessment of attentional problems. The client is G. T., a 10-year-old biracial male who is experiencing difficulty sustaining attention during class and listening to instructions. The information from the parent and teacher versions of the behavior rating scales are somewhat inconsistent. His teacher endorses a number of attention problems for G. T. such as frequent forgetting of instructions before completing schoolwork; G. T. fidgets often; and G. T. inappropriately leaves his seat during class. According to his teacher, G. T. also appears to be easily distracted during teacher presentations of class material. However, the parent rating reveals only minor difficulties with attention.

As part of a larger psychological battery, G. T. was assented to complete Virtual Classroom CPT and Stroop tasks. The virtual classroom was modified to closely reflect his typical classroom environment by including the same number of students, assigned desk location, and same sex of the teacher. While immersed in the Virtual Classroom, G. T. sat at a desk in the middle of the virtual classroom that corresponded with where he sits in his real classroom at school. During the Stroop and CPT tasks, head movement tracking sensors logged G. T.’s frequent head movements as he looked away from the blackboard to observe vehicles passing by outside and virtual students moving around in the room. Moreover, the movement sensors logged his body movement and revealed that he was very physically active during the tasks. He showed poor performance (compared with norm-reference groups) on the cognitive tasks.

The results of testing indicated G. T. was showing several behavioral symptoms of ADHD. Review of the data showed that symptoms of inattention, such as maintaining attention, are likely because of being easily derailed by classroom and outdoor distractors. This also affected his ability to follow through with instructions and listen to the virtual teacher. Of particular note are the symptoms of hyperactivity that were directly recorded. He frequently moved his legs and arms and squirmed around in his seat. He had difficulty sitting still and remaining seated.

Several recommendations were made based on these observations. G. T. performed better when placed in the front of the virtual classroom, so it was recommended that this adjustment be made in his actual classroom. It was recommended that classroom windows remain closed if possible to reduce noise from outside sources. Additionally, it was recommended that people entering and exiting the room be reduced as much as possible. Finally, the specific distractors were provided to the parent to share with G. T.’s teacher, so when these distractors occur the teacher will know it is necessary to redirect G. T.’s attention.

In summary, the Virtual Classroom has several advantages as opposed to traditional pen-and-paper testing. It represents the typical classroom in which students with ADHD and autism often struggle to maintain attention and engagement, which increases ecological validity compared with traditional testing methods. In addition, functional impairments can be directly observed, which improves the quality of recommendations. The Virtual Classroom provides a record of which distractions caused the client to look away from the board and how distracting they were (how often the client looked away). These data can inform recommendations to parents and teachers to remove specific distractors from the learning environment. The impact of placing the client in different locations in the virtual classroom can also be assessed. The results of such testing allow for recommendations based on direct observation about where to seat the client in the classroom so they can perform at their best. The virtual classroom records the head and body movements of the child in real time. Thus, the level of activity is accurately recorded without the potential bias that may influence parent and teacher reports (Pas & Bradshaw, 2014; Sayal & Taylor, 2005). Collecting and recording these body movement data in an ecologically valid environment rather than relying on parent and teacher reports may be a better way of assessing the “H” in ADHD.
Virtual Shopping Tasks

Another virtual environment approach that may be of interest to clinicians is the use of a simulated shopping environment. A number of these virtual shopping environments have been developed to assess execution of everyday behaviors in a virtual shopping center (see Table 4). These virtual shopping environments have been developed in a manner that allows the psychologist to systematically vary the information load (that affects goal maintenance). For example, the Virtual Environment Grocery Store (VEGS) uses multiple adaptive trials in the assessment procedure by creating a pool of “multiple task assignments,” empirically determining their baseline difficulty, and then adding conditions in the environment that affect baseline task difficulty via the manipulation of the density of items on shelves, the similarity of packaging, and the intensity and types of realistic irrelevant distractions (e.g., loudness/type of music in the background and loudspeaker announcements). The VEGS platform offers a range of difficulties that may be used to make the tasks sufficiently complex to avoid floor and ceiling effects. The use of a simulated environment allows older adults who may be physically or behaviorally impaired to be safely assessed. This would not be possible with the traditional MET, which requires these tasks to be completed in a real-world shopping environment (Parsons, McPherson, & Interrante, 2013).

The case of L. S. A psychologist receives a referral for assessment of memory problems. The client is L. S., a 58-year-old married, Hispanic female who sought assessment after memory difficulties appear to have resulted in her frequently misplacing items around the house and forgetting to attend appointments. She also continually asks family members to help her with shopping. Furthermore, she reported that at times she would get up and go to her kitchen only to feel confused because she could not remember what she intended to do in the kitchen. She had a diagnosis of osteoarthritis, which decreased movement in her right knee, and denied any previous history of head-injury, substance abuse, or psychiatric illness. Data collected during the clinical interview suggested early onset Alzheimer’s disease, but results from the neuropsychological assessment were inconsistent.

To assess her cognitive functioning, L. S. was asked to complete the VEGS as part of her larger battery of tests. While immersed in the VEGS, L. S. often had difficulty with sustained concentration, which was demonstrated by the indirect route she took to the pharmacy and when shopping for items on the shopping list. She also had difficulty remembering how much money was budgeted for purchasing items and continually asked her psychologist what she was supposed to do next. She also performed poorly on various aspects of the prospective memory tasks. For example, she had to be reminded to go to the coupon machine after 5 min of shopping. She also had difficulty recalling the shopping items and frequently clicked on the shopping list for a reminder. These difficulties were notably amplified when distractors were present (e.g., announcements over the loud speaker; people shopping in the same aisle as her).

These results suggest that L. S. is demonstrating behavioral symptoms of suggestive of early onset Alzheimer’s disease, including difficulty with concentration, problem solving, and memory for recently learned information. With permission from L. S., her caregiver was shown a virtual replay of her activities in the VEGS. It revealed areas of difficulty and supported a recommendation that L. S. be provided assistance with tasks that require sustained attention such as shopping or household chores. The VEGS report also revealed that she needed assistance with tasks that require problem-solving skills such as budgeting. It was recommended that family members provide L. S. with frequent reminders of newly learned information such as dates and changes in routines on an easily accessible calendar and in-person.

Virtual Reality Apartment

Virtual Apartments are simulation environments that allow clinicians to see how the client behaves in their home environment. The purpose of the Virtual Apartment is to assess potential cognitive deficits using a simulated apartment. While a number of virtual apartment environments have emerged, there is a consistent emphasis upon (a) ecologically valid representations of the client’s everyday activities in a living situation; (b) presentation of either traditional cognitive tasks (e.g., CPT; Stroop) or
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional tests</th>
<th>Results</th>
</tr>
</thead>
</table>
| Atkins et al. (2015)                | $N = 44$ healthy young and $N = 39$ healthy old.                        | The two groups were compared on virtual and traditional measures.                | • MATRICS Consensus Cognitive Battery  
  • UCSD Performance-Based Skills Assessment-Brief  
  • Hopkins Verbal Learning Test  
  • Brief Visual Memory Test | • Each VR outcome measure displayed significant age-related differences.  
  • Traditional measures of cognitive functioning were significantly associated with VR performance across age groups.  
  • Virtual shopping performance differentiated between TBI patients and the control group.  
  • Measures of prospective memory, neurocognitive functioning, and psychosocial functioning were significantly associated with Virtual shopping performance among TBI patients. |
| Canty et al. (2014)                 | $N = 30$ severe TBI vs. $N = 24$ healthy controls.                      | Severe TBI ($n = 30$) vs. healthy control ($n = 24$).                            | • Lexical Decision Prospective Memory Task  
  • Hopkins Verbal Learning Test  
  • Trail Making Task  
  • Controlled oral word association  
  • Hayling Sentence Completion  
  • Letter Number Sequencing  
  • Sydney Psychosocial Reintegration Scale | • Virtual shopping performance differentiated between TBI patients and the control group.  
  • Measures of prospective memory, neurocognitive functioning, and psychosocial functioning were significantly associated with Virtual shopping performance among TBI patients. |
| Carelli, Morganti, Weiss, Kizony, & Riva (2008) | $N = 24$ healthy adults.                                                 | Descriptive study.                                                              | n/a                                                                              | • Virtual supermarket may be a useful tool in executive assessment, particularly due to its temporal and accuracy measures. |
| Erez, Weiss, Kizony, & Rand (2013)  | $N = 20$ children with TBI vs. $N = 20$ healthy controls.               | Comparison of TBI ($n = 20$) to healthy controls ($n = 20$) in a virtual mall.  | n/a                                                                              | • Outcome measures of the VMall successfully differentiated between children with TBI and healthy controls. |
| Josman et al. (2006)                | $N = 20$ stroke patients.                                               | Comparison of Virtual Supermarket to traditional tests.                         | Behavioral Assessment of Dysexecutive Syndrome (BADS)                             | • Virtual shopping outcome measures (# items purchased, # correct actions, duration pauses/stops) were significantly associated with the key search subtest of the BADS.  
  • Results revealed significant differences in executive functioning between post-stroke patients and the control group on virtual shopping outcome measures.  
  • Virtual shopping outcome measures were also significantly associated with the BADS. |
<p>| Josman et al. (2014)                | $N = 24$ stroke and 24 matched controls.                               | Comparison of stroke ($n = 24$) and control ($n = 24$) participants.            | Behavioral Assessment of Dysexecutive Syndrome                                    | |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional tests</th>
<th>Results</th>
</tr>
</thead>
</table>
| Josman, Scheniderman, Klinger, and Shevil (2009) | $N = 30$ schizophrenia and $N = 30$ healthy controls. | Comparison of schizophrenia ($n = 30$) to healthy control ($n = 30$) performance. | Behavioral Assessment of Dysexecutive Syndrome | • Virtual shopping outcome measures were sensitive to differences in executive functioning between schizophrenia patients and controls and differentiated between patients with differing levels of executive function.  
• Virtual shopping outcome measures were negatively related to symptoms of schizophrenia and the BADS. |
| Kang et al. (2008) | $N = 20$ stroke and $N = 20$ matched controls. | Comparison of Stroke ($n = 20$) and control ($n = 20$) performance. | n/a | • Virtual shopping outcome measures consistently differentiated between stroke patients (unilateral brain lesion) and the control group.  
• Virtual shopping outcome measures revealed that patients walked a significantly longer distance.  
• This suggests that the Virtual Supermarket may be a useful tool for assessing cognitive planning in patients with Parkinson’s. |
| Klinger, Chemin, Lebreton, and Marié (2006) | $N = 5$ Parkinson’s and $N = 5$ age-matched healthy controls. | Parkinson’s ($n = 5$) to age-matched healthy control ($n = 5$) performance. | n/a | n/a |
| Okahashi et al. (2013) | $N = 10$ brain damage and $N = 10$ age-matched controls and $N = 10$ healthy-older adults compared to $N = 10$ healthy-young. | Comparison of brain damage ($n = 10$) performance to age-matched control ($n = 10$); Healthy-old ($n = 10$) and healthy-young ($n = 10$). | Behavioral Assessment of Dysexecutive Syndrome | • Performance on the virtual shopping task was significantly associated with conventional cognitive assessments.  
• Older participants and patients with brain damage scored significantly worse on the virtual shopping tasks.  
• Virtual shopping tasks successfully differentiated between all three groups.  
• Virtual shopping outcome measures were significantly associated with Zoo Map subtest. |
| Rand, Basha-Abu Rukan, Weiss, and Katz (2009) | $N = 9$ post-stroke ($n = 9$) and $N = 20$ healthy-young and $N = 20$ healthy aging | Comparison of post-stroke ($n = 9$) vs. healthy-young ($n = 20$) vs. healthy-old ($n = 20$). | Behavioral Assessment of Dysexecutive Syndrome (Zoo Map subtest) | • Virtual shopping tasks successfully differentiated between post-stroke patients and the control groups with post-stroke patients taking significantly longer to complete the shopping task.  
• Virtual shopping time successfully differentiated between post-stroke patients and the control groups with post-stroke patients taking significantly longer to complete the shopping task. |
| Rand, Katz, and Weiss (2007) | $N = 14$ stroke and $N = 23$ healthy control children and $N = 44$ young adults; and $N = 26$ older adults. | Comparison of stroke ($n = 14$) vs. healthy control (children $n = 23$; young adults $n = 44$; older adults $n = 26$). | n/a | n/a |
### Table 4 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Research design</th>
<th>Traditional tests</th>
<th>Results</th>
</tr>
</thead>
</table>
| Raspelli et al. (2012)             | $N = 9$ stroke and $N = 10$ healthy-young; and $N = 10$ healthy-old.    | Comparison of stroke ($n = 9$) vs. healthy-young ($n = 10$) vs. healthy-old. | • Test of Everyday Attention  
• Iowa Gambling Task  
• Stroop Test                                                     | • Significant differences were found between all three groups on two outcome measures of the virtual shopping task (i.e., time, errors).  
• The TEA (but not the IGT or Stroop) was significantly associated with virtual shopping performance in post-stroke participants. |
| Werner, Rabinowitz, Klinger, Korczyn, and Josman (2009) | $N = 10$ mild cognitive impairment and $N = 30$ matched controls.      | Comparison of MCI ($n = 30$) vs. control ($n = 40$). | Behavioral Assessment of Dysexecutive Syndrome                                  | 4 of the 8 virtual shopping outcome measures were associated with performance on the BADS.  
• Virtual shopping performance successfully differentiated between MCI patients and the control group.  
• Virtual shopping performance was moderately correlated with traditional neuropsychological tests.  
• Virtual shopping performance was able to differentiate between MCI patients and the control group; however, it was unable to differentiate MCI subtypes. |
| Zygouris et al. (2015)             | $N = 34$ mild cognitive impairment and $N = 20$ matched controls.       | Comparison of MCI ($n = 34$) vs. control ($n = 21$). | • Mini Mental  
• Rey-Osterrieth Complex Figure Test  
• Rey Auditory Verbal Learning Test  
• Rivermead Behavioural Memory Test  
• Test of Everyday attention  
• Trail Making test  
• Functional Rating Scale for Symptoms of Dementia  
• Functional Cognitive Assessment Scale |                                                                                                                                  |

*Note.* BADS = Behavioral Assessment of Dysexecutive Syndrome; MCI = Mild Cognitive Impairment; TEA = Test of Everyday Attention; IGT = Iowa Gambling Task; TBI = Traumatic Brain Injury; UCSD = University of California, San Diego; VR = Virtual Reality; VMall = Virtual Mall.
everyday activities (without embedded cognitive stimuli) in the environment; (c) low and high distraction conditions; and (d) logging of behavioral metrics (e.g., ambient body movement, head/eye gaze).

The ClinicaVR team has extended the virtual classroom paradigm to a virtual apartment that superimposes construct-driven stimuli (e.g., Stroop and CPT) onto a large TV. The Virtual Apartment Bimodal Stroop (VABS) is a 9.6-min task. Participants are seated in the living room, in front of a flat-screen TV. A kitchen is located to the left of the TV and a window is located to the right of the TV. The task builds on the unimodal (visually mediated) Stroop and measures cognitive interference using reaction time (RT), commission errors and omission errors, and RT variability. The VABS extends the traditional Stroop paradigm via the inclusion of bimodal (auditory and visually mediated) stimuli.

During the task, distractors appear in different field of view locations in the environment. Some distractors are audio—visual: School Bus passing on the street and SUV viewed through window on the right; iPhone ringing and vibrating on the table (in front of participant); Toy Robot moving and making noise on the floor (center). Auditory distractors included: Crumple Paper (left); Drop Pencil (left); Doorbell (left); Clock (left) Vacuum Cleaner (right); Jack Hammer (right); Sneeze (left) Jet Noise (center). Visual distractors included: Paper airplane (flying from left to right in front of the participant), Woman walking in the kitchen (center). That condition was designed to assess RTs (simple and complex), selective attention (matching the auditory and visual stimuli), and external interference control (environmental distracters).

In a preliminary study, Henry and colleagues (2012) with 71 healthy adult participants found that the VR-Apartment Stroop is capable of eliciting the Stroop effect with bimodal stimuli. Initial validation data also suggested that measures of the VR-Stroop significantly correlate with measures of the Elevator counting with distractors (ranging from .38 to .62), the Continuous Performance Task (CPT-II; ranging from .32 to .42), and the Stop-it task (ranging from .37 to .39). Results from regression indicated that commission errors and variability of RTs at 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.

**Virtual reality apartment medication management assessment.** A different Virtual Apartment has been developed by Kurtz, Baker, Pearlson, and Astur (2007) to assess medication management skills among those with schizophrenia. The Virtual Reality Apartment Medication Management Assessment (VRAMMA) lasts a maximum of 23 min. The environment consists of a four-bedroom apartment containing a living room, bedroom, kitchen, and bathroom. During the practice phase, clients are asked to do a number of tasks that help build familiarity with the virtual environment: using a joystick to navigate to the living room and turn on the TV, checking the time on the interactive clock on top of the TV, turning on a light in the bedroom, checking the medication reminder post-it note in the kitchen (with a different prescription than the actual trial) after turning on the light and turning off the stove, and opening the medicine cabinet in the bathroom and taking out pills. During the testing phase clients start in the living room and a message is displayed that gives them the medications to take and the time they must be taken. They must then use all the items from the practice phase to take the correct medication at the correct time (in 15 min). Several auditory distractors that simulate real-life distractions were included in the virtual apartment: a phone ringing, a doorbell ringing, a dog barking, and a police siren. These auditory distractors are introduced every 3 min. Significant events, such as turning on lights or checking the clock, and the location and movement of the client are recorded. The variables recorded during testing include: quantitative errors, qualitative errors, time discrepancy, total distance traveled, clock checks, and reminder note checks. The VRAMMA is able to distinguish between healthy controls and those with schizophrenia, who perform worse. Kurtz et al. (2007) conducted a validation study with 25 schizophrenia patients and 18 matched healthy controls. They found patients with schizophrenia made more quantitative errors concerning the number of pills taken ($p = .001$), were less likely to take pills at the correct time ($p = .01$), and checked the clock less often ($p = .001$).

The VRAMMA has several advantages compared with traditional testing. The VRAMMA
shows good convergent validity with the Medication Management Ability Assessment (Patterson et al., 2002), a role-play task involving medication management that typically would be administered in an office setting. Clinicians are able to directly observe client behavior in a more ecologically valid environment rather than relying on self-report measures, which may be biased by a lack of insight from clients with schizophrenia (see Mintz, Dobson, & Romney, 2003 for a meta-analysis) or cognitive and emotional disruptions (Anticevic, Schleifer, & Youngsun, 2015). The VRAMMA approximates an environment in which many clients would likely need to manage medication and includes common tools needed to carry out this activity, such as a reminder note and a clock. It also allows clinicians to assess when the process of performing this task is disrupted and make specific recommendations from these observations.

The case of J. A. A psychologist receives a referral for assessment of concentration, and a resurgence of visual and auditory hallucinations. The client is J. A., a 23-year-old, single White male who was diagnosed with schizophrenia in 2013 after hospitalization for a single episode of psychosis. He was released from the psychiatric hospital after 1 month and given a prescription for antipsychotic medication. J. A. continued meeting regularly with a psychiatrist. Several months after being released from the psychiatric hospital, J. A. began experiencing difficulty with concentration, and a resurgence of visual and auditory hallucinations. To ascertain whether medication adherence was at issue, J. A. was asked to complete the VRAMMA. He took fewer than the required number of pills and took the pills 3 min after the 15-min time lapse had expired. While in the virtual apartment, J. A. traveled around the apartment longer than necessary to complete the task and was distracted by the sound of the barking dog.

Based on these results, it appears that J. A. is having difficulty taking the correct number of pills and using cues in the environment to take pills at the correct time. Positive symptoms are related to distance traveled in VRAMMA (Kurtz et al., 2007). J. A. traveled a significant distance in the virtual apartment, which shows his medication adherence is low. It was recommended that J. A. use an automated pill dispenser that holds the correct number of pills for each day of the week, reminds J. A. to take medication if not taken at a scheduled time, and sends a notification to his psychiatrist if he stops taking medication. J. A. was advised to reduce auditory distractions at home by turning on a fan to reduce noise from outside sources.

Considerations in the Adoption of Virtual Reality Technologies

As can be seen there are some specific cases in which virtual reality-based psychological assessments may offer the psychologist with ecologically valid assessments of day-to-day activities. That said, there are a number of considerations that go into the decision to add new technologies to one’s battery of tests (see Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). While some of these issues are concerns related to the current generation of virtual reality based assessments, others reflect outdated concerns from an earlier generation of platforms., for example, it used to be the case that the equipment needed to conduct such assessments was bulky and expensive. Recent advances in virtual reality technology have made the use of simulations in assessment more feasible and affordable. Smaller, easier-to-use equipment and reduced cost make virtual environment assessment a practical tool psychologists can use to gather precise functional data and to provide customized recommendations to clients (Bohil et al., 2011). Furthermore, validated virtual environments with automated stimulus presentation, data capture, and scoring are emerging that include sample characteristics for norm-referenced assessment. However, there are other concerns that continue to this day. The dearth of established guidelines for the development, administration and interpretation of these assessments could lead to important psychometric pitfalls. At minimum, all virtual reality-based psychological assessments must have (and many now do) standardized instructions for administration and methods for scoring and interpreting test results provided in a test manual. While some virtual environments are being designed for limit testing, more work needs to be done in this area.
Need for VE-Based Neuropsychological Assessments to Be Sufficiently Standardized

While the use of virtual environments for assessment is an emerging area of application, adoption will require substantial research and development to establish acceptable psychometric properties and clinical utility. Although a review of VR therapies has revealed statistically large effects on a number of affective domains (Parsons & Rizzo, 2008), findings must be interpreted with caution given that some VR studies do not include control groups, and many are not randomized clinical trials, limiting the confidence that the enhancements were caused by the VR intervention. An important resolution to clinical heterogeneity of outcome measures in virtual environment research is the standardization of outcomes and the measures used to assess these outcomes. The selection of outcome measures for standardization need to be relevant to the client’s treatment and health status as well as psychometrically sound. Another pressing need among psychologists is the identification of VE-based assessments that reflect relevant underlying cognitive and behavioral capacities for assessments of varying degrees of psychological deficits. VE-based assessments must demonstrate relevance beyond that which is available through simpler means of assessment. As such, there is specific need for VE-based assessments to be sufficiently standardized within the range and nature of responses available to participants within the virtual environment to allow for reliable measurement.

VE-based assessment studies have often sought to establish construct validity by demonstrating significant associations between virtual environments and paper-and-pencil assessments (e.g., virtual classroom assessments). 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 model using a function led approach accurately predicts relevant real-world behavior may be more important than large-magnitude associations with traditional construct driven paper-and-pencil tests (e.g., virtual shopping tasks).

Issues for Use of Virtual Environments in Specific Patient Populations

In addition to psychometric and technical issues, clinicians, researchers, and policymakers will need to scrutinize emerging VE-based assessments to ensure adherence to legal, ethical, and human safety guidelines. The matching of specific technologies to the needs and capacities of the client will also require careful consideration by psychologists. How will virtual environments be experienced by certain clinical populations? In pediatrics and geriatrics, human guidance is critical for safeguarding the client’s full comprehension of assessment use and instruction. Geriatric clients in particular may find adjusting to virtual platforms, on the whole, difficult (Miller et al., 2014).

Although virtual environments have been successfully applied to the study of age differences in spatial navigation among both healthy and demented elderly, virtual environment-based tasks may be complicated by visual, auditory, or motor impairment. In comparison with younger controls, aging patients may perform more poorly on virtual environment-based tasks simply because of the normative aging process or because of 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, overexertion, and musculoskeletal irritation. There is need for more rigorous research methods including more consistent and strenuous reporting of exercise dosages and adherence. Moffatt (2009) suggests a number of helpful methodological practices in assessing older adults in research studies of navigation skills, including: (a) allowing aging patients to practice and ensure maximum familiarization with the computer platform, (b) including measures of computer experience, visual ability, and motor function, and (c) including assessments requiring the same sen-
sorimotor capacities, but not physical navigation.

A potential barrier to adoption of virtual reality technology among clinicians is concerns about the ability of older adults to use this technology. Wandke, Sengpiel, and Sönksen (2012) have outlined several pervasive myths about older adults and human-computer interaction. They state that the myth that older adults are not interested in computers is not true. Dyck and Smither (1994) conducted a survey of adults over age 55 and found these older adults were less computer anxious and had more positive attitudes about computers than adults under 30. Wandke et al. (2012) also address the myth that “you can’t teach an old dog new tricks.” While there are decreases in brain plasticity in older age, this does not mean that learning ceases. It may be the case that some older adults have had frustrating experiences that lead to giving up on learning how to use new technologies. However, this effect should not be overgeneralized, as older adults are often interested in using newer technologies (Sayago, Sloan, & Blat, 2011).

When assessing older adults for memory performance, it is important to avoid invoking stereotype threat. Chasteen, Bhattacharyya, Horhota, Tam, and Hasher (2005) found that invoking stereotype threat about memory abilities in older adults harms performance on memory tasks, particularly when these adults are aware their memory is being assessed. Subtle and unambiguous age-related stereotypes have also been found to influence older adults’ performance on a number of cognitive and physical tasks (see Lamont, Swift, & Abrams, 2015 for a meta-analysis), such as and map learning (Meneghetti, Muffato, Suitner, De Beni, & Borella, 2015) driving a car (Lambert et al., 2016), and hand grip strength (Swift, Lamont, & Abrams, 2012). It is possible that assessment in a virtual environment such as the VEGS could reduce stereotype threat by obscuring the true purpose of the task.

Simulation technology may also be problematic for individuals with autism spectrum disorder. Given pronounced sensory issues commonly found in this population, the head mounted display or even the graphical interface may be experienced as intolerable. Moreover, there is concern that too intense a stimulus presentation may aggravate sensory processing difficulties. This is an important concern though there is no evidence from two different studies with students diagnosed with autism that they experience negative effects over and above those experienced by students without autism (Parsons & Carlew, 2016; Wallace et al., 2010). However, while these two studies tend to suggest that negative effects were self-reported as low, they involved screen-based virtual environments. As we adopt newer and more immersive technologies (i.e., HMDs) it is important to consider the potential negative effects (i.e., dizziness, sickness, and displacement) to ensure that wearable technologies (e.g., HMDs) can provide an acceptable space for children to use; especially children with disabilities. With this said, there is some evidence that suggests children do not experience HMDs any more negatively than screen-based media (Peli, 1998). Although more work is needed in this area, these findings support the potential of VR technology for continued greater approximations (Bohil et al., 2011) of cognitive processes in the real world.

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 virtual environment. There is also the potential of unintended negative effects of exposure to virtual environments—stimulus intensity, if taken too far, may exacerbate rather than ameliorate a deficit. High-fidelity virtual environments may be confusing for these individuals and increase negative behaviors after exposure to the environment. Flat-screen presentation of virtual environments has proven to be an acceptable alternative to full immersion with the environment, and may be more appropriate for certain clinical groups (Attree et al., 1996).

Summary and Conclusions

There are many different types of tests available to psychologists for determining a client’s level of functioning. The challenge for psychologists is choosing tests that provide accurate information for making prescriptive statements to clients, parents, and teachers based on the best evidence available. Self-reports can be helpful in collecting data on specific areas of functioning, but also suffer from lack of agreement among informants (Biederman et al., 1993, 1990; Mitsis et al., 2000; Newcorn et al., 1994; de Nijs et al., 2004) potential bias from
clients (Schwarz, 1999), and provides post hoc appraisal of behavior. Normative comparison of performances on cognitive assessments allows the psychologist to determine if a client’s performance is similar or divergent from peers, but provides limited information about daily functioning because they are construct-driven rather than function-led.

The addition of virtual reality to a psychological battery provides an opportunity for psychologists to obtain more ecologically valid data about client functioning in simulations of real-world environments. Virtual environments allow the psychologist to have greater control of dynamic perceptual stimuli and the sensitivity of the test, while also capturing data about client performance in activities of daily living (Bohil et al., 2011). The computerized nature of these tests allows for the accurate capture of neurobehavioral data, as well as precise recording and scoring of neuropsychological test results (Campbell et al., 2009). Several virtual environments have already been developed to for psychologists to use in neuropsychological assessment, such as the virtual classroom, virtual grocery store, and the virtual apartment. Although more validation studies need to be conducted with virtual reality assessments, the benefits of using this technology for understanding daily functioning are clear. In addition, smaller and more affordable equipment makes virtual reality a viable option for use in psychological assessment.

Preferably, virtual assessments will be 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 generalizable 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 psychological battery and the minimization of human interaction.

Virtual environments may add to an existing psychological battery when the psychologist is attempting to make accurate predictions about a person’s behavior within the real world. In a virtual environment, the psychologist can measure functional output of constructs within the complexity of a real-world environment. For example, in a virtual classroom, selective attention can be measured by conducting tests such as the CPT in a real world environment. In a virtual environment grocery store, prospective memory may be assessed using a real-world task like remembering to pick up a prescription at the pharmacy. Cognitive interference can be assessed in a virtual apartment that includes common distractors found in an everyday environment.

Technological innovations, such as virtual reality, allow psychologists to expand our methods for designing and implementing assessments capable of collecting information that provides an accurate picture of client limitations. These advances improve the prescriptive statements psychologists dispense by providing the opportunity to observe client functioning in real-world environments—a practice that might otherwise be infeasible because of clients’ behavioral and physiological impairments. By adopting virtual reality as a method for assessing clients, psychologists increase the potential positive impact of neuropsychological assessment for improving the daily functioning of clients through accurate understanding of neuropsychological deficits and directly relevant recommendations.

References
Atkins, A. S., Stroescu, I., Spagnola, N. B., Davis, V. G., Patterson, T. D., Narasimhan, M., . . . Keefe,


Members of Underrepresented Groups:
Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.

- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.

- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.

- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit http://www.apa.org/pubs/authors/review-manuscript-ce-video.aspx.