Increasing the Ecological Validity of Traditional Neuropsychological Assessments

By

Reanna Trexler

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Arts in Psychology

Middle Tennessee State University

August, 2024

Thesis Committee:

Dr. Paul Foster, Chair

Dr. Kimberly Ujcich Ward

Dr. James Loveless

ABSTRACT

The purpose of this study was to create a version of a traditional neuropsychological assessment, the Hopkins Verbal Learning Test – Revised (HVLT-R), that is ecologically valid by adding an audio distraction to the standardized test setting. Additionally, a subjective memory questionnaire (SMQ) was given to determine if the participants' views about their own memory may be related to the difference in scores on the HVLT-R. Forty-eight healthy college-aged individuals were given the HVLT-R with and without an audio distraction, in addition to other standardized assessments and questionnaires. The participants were predicted to perform worse on average on the version of the HVLT-R with the audio distraction. Overall, the difference in scores from the HVLT-R without audio to the audio distraction condition was not significant. Also, scores on the SMQ were not shown to be related to the difference scores of the HVLT-R with and without an audio distraction.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

Every so often, patients will arrive at neuropsychological clinics complaining of memory deficits but will perform within normal limits on formal memory assessments. These clients may make statements such as, “I cannot remember conversations,” “I cannot remember names,” or “I cannot remember what I was doing and will forget to complete a task that I started.” However, on formal neuropsychological evaluations they are not found to have impaired memory functioning. Several studies provide evidence that individuals complaining of cognitive dysfunction in their daily lives do not always present as impaired when assessed using traditional neuropsychological measures. Personal reports of daily cognitive difficulties are not always related to objective measures of cognitive performance (Akbar & Finlayson, 2021), and subjective memory complaints may not be related to objective memory impairment in elderly individuals (Caramelli & Beato, 2008). At least three reasons can be used to explain the discrepancy between subjective memory complaints and the results of formal, objective neuropsychological testing. First, an individual may have historically performed at a high average or superior level on formal memory testing, but at the time of the evaluation, they may score within the average range. Therefore, the patient may be experiencing true memory decline but is still functioning within normal limits. A second possibility may be that these individuals are not accurately assessing and judging their own memory functioning. This scenario might arise with elderly individuals who misattribute normal age-related decline to a possible degenerative disease. The final possibility involves the potential issue of ecological validity. Specifically, patients are tested in optimal environments of one-on-one interactions that are free from distractions and stressors. However, many of these individuals might experience problems when outside of the clinic setting and in situations that have distractions and stressors present.

The purpose of the current study is to test the idea of ecological validity and how it plays a role in these issues of standardized memory testing.

As stated, one specific reason as to why these deficits do not appear on standardized tests could be the result of poor ecological validity. Ecological validity refers to the ability to translate results of memory and cognitive assessments conducted in a standardized environment to the client's memory and cognitive abilities in their everyday life (Chaytor et al., 2006). Research settings and clinical environments are two areas in which ecological validity is important. A controlled research environment may include administering standardized tests to participants in a laboratory setting, where distractions are minimal, and the examiner must refrain from straying from the procedure as to maintain control of potential extraneous variables. A neuropsychology clinic is a controlled environment as they are also free from distractions, and clinicians must maintain standardization when administering the tests to ensure the results are reliable and valid. However, the results from this type of procedure do not translate to everyday life because the environment does not resemble the real world.

Currently, neuropsychological assessments have been shown to be valid for diagnosing and tracking decline in cognitive abilities (Kibby et al., 1998; Makatura et al., 1999; Youn et al., 2020), but there is limited literature that shows that neuropsychological assessments are ecologically valid, and there is minimal research on improving the ecological validity of traditional paper-pencil neuropsychological tests. Much research has been conducted to change the way that neuropsychological assessments are administered so that they are ecologically valid. However, most of the research consists of modifying the assessment so that it is conducted in a naturalistic or simulated environment. Additionally, the current approach for assessing the ecological validity of current assessments has mainly consisted of taking already established and

normed tests and administering them in natural environments. An attempt to create norms for a test in these environments has not occurred. Importantly, these assessments should accurately reflect the patient's everyday abilities.

Chaytor et al. (2006) states that neuropsychologists tend to assume that poor performance on traditional clinical tests means that the client will also have poor cognitive functioning in everyday life. However, patients may arrive at the clinic with memory complaints, yet are performing within a normal range on the objective assessments. Observing a client in a clinical setting may not accurately reflect their cognitive abilities in everyday situations because the testing conditions do not reflect their typical everyday life. The clinic lacks the distractions and other stressors that are present in everyday, natural environments.

An important aspect of memory assessments is to provide information about client's everyday memory functioning based on the given assessment. The present paper aims to explore the issue of ecological validity of neuropsychological tests for memory functioning. The goal is to propose that a superior approach for increasing ecological validity is to develop and norm these tests in situations that resemble real life. The idea is that this specific type of assessment can be used during limits testing with a client who complains of memory and cognitive problems but performs within normal limits on objective assessment. Therefore, the test would be administered in relation to the referral question. To achieve this objective, the problems of standardized assessment settings will first be discussed, followed by a theory of why these problems with standardized assessment settings may occur. Next, a discussion of problems with the ecological validity of current neuropsychological tests will be mentioned, as well as potential reasons for poor ecological validity. Finally, possible solutions to increasing ecological validity of current formal neuropsychological assessments will be discussed.

Subjective Memory Complaints, but Average Objective Memory Scores

As stated, people that present to the clinic with memory complaints, or other cognitive complaints, sometimes do not perform at an impaired level on the standardized clinical assessment that is administered by the neuropsychologist (Caramelli & Beato, 2008; Pearman et al., 2014). Patients may complain that, in their daily lives, they do not remember certain tasks that they were attempting to complete, or they may not remember conversations that they had with other people. The phenomenon of patients complaining of their own memory but scoring within a normal range when in a clinical setting has been shown in patient populations, such as in people with multiple sclerosis (MS) and mild traumatic brain injury (mTBI). The phenomenon has also been demonstrated in healthy adults and healthy elderly adults, in which they subjectively perceive their memory to be worse than what is shown on objective assessments.

Demonstrating Discrepancies Between Subjective and Objective Measures

Some studies have been conducted that discuss the discrepancies that can be observed between subjective and cognitive assessments. Akbar and Finlayson (2021) did a study to determine if there is a relation between self-reports of cognitive function and neuropsychological test performance in patients with multiple sclerosis. The study included 12 participants between the ages of 28 to 73 years. The researchers used a convergent parallel mixed-methods design using both a semi-structured interview and a neuropsychological assessment. The semi-structured interview was conducted to determine the cognitive daily function of someone with MS, and a list of questions was used by the interviewer. The neuropsychological assessment included the California Verbal Learning Test Second Edition (CVLT) and Brief Visuospatial Memory Test- Revised (BVMT) for measuring visual and verbal aspects of memory. The Symbol Digit Modalities (SDM) Test and Delis-Kaplan Executive Function System (D-KEFS) Trail

Making Test were used to measure information processing speed. Selective and divided attention was measured with the D-KEFS Color-Word Interference Test and Paced Auditory Serial Addition Test. General executive functioning was measured using the D-KEFS Sorting Test and Tower Test. Verbal Fluency from the D-KEFS was used to measure word finding abilities. The participants complained about four primary cognitive areas which included memory, word finding, processing speed, and executive functioning. The results of the objective measures showed that 43% of participants had difficulties with memory, 33% had difficulties with processing speed, and 0% had difficulties with word finding. These findings demonstrate that objective measures and subjective cognitive abilities are not highly related. Cognitive abilities are perceived by the participants to be reduced compared to what is shown by the objective measures.

Caramelli and Beato (2008) produced similar results that demonstrate the discrepancy between subjective and objective cognitive measures. The study was conducted by administering the Mini-Mental Status Exam (MMSE) to a group of elderly individuals with and without subjective memory complaints (SMCs). All participants scored in the intact range for cognitive functioning, even if they complained of cognitive issues in their daily lives. Pearman et al. (2014) also completed a study with comparable findings when they observed relationships between memory performance and subjective memory complaints in 516 older adults. The adults had an average age of 85 years old and were assessed over a 6-year time span. The participants completed a memory complaint questionnaire and an episodic memory assessment throughout this time. The results revealed that a relation between the objective and subjective memory assessments did not exist. Memory complaints were not related to objective memory performance or changes in performance overtime. Therefore, the results of both studies provide

evidence that there may be extraneous environmental factors influencing individuals' cognitive abilities in the real world that are not present in traditional cognitive assessments.

Investigating the Discrepancies Between Subjective and Objective Measures

As shown previously, several studies have demonstrated that there is a discrepancy between subjective and objective measures for memory and cognitive functioning. As a result, other studies have been conducted to investigate the discrepancies between these two types of assessments. Some studies have been conducted that introduce a distractor to a measure of cognitive functioning that had previously been shown to lack results, as the patients were not performing as impaired, but were claiming to have impairments. One study done by Johansson (2021) used a computerized version of an attention task to determine the attentional abilities of those with mild traumatic brain injury and added a distractor within the assessment. During the distraction component of the test, the mTBI group had a much slower reaction time than the control group. Leavitt and Katz (2006) showed similar results in patients with Fibromyalgia. The researchers used traditional measures of memory and attention, and also included cognitive measures with stimulus competition. On the traditional measures, the patients scored in the average range. However, on the measures with stimulus competition, they scored significantly below average. Importantly, the patient group and control group scored similarly only on the traditional measures. These findings suggest that traditional measures are missing a component that could potentially demonstrate how patients are performing in their natural environments.

Distractions are shown to be related to a decrease in performance on neuropsychological assessments that have previously been shown to have no differences between healthy individuals and those with subjective memory complaints. These assessments ultimately lead to recommendations for clients to improve their daily lives. However, if the tests are not accurately

measuring their everyday cognitive abilities, then recommendations made based on these measures may not be adequate. Adding a distraction to an existing measure may be an important component of increasing the ecological validity of the measure.

Further Evidence of Discrepancies on Subjective and Objective Measures

Some researchers have attempted to further assess the discrepancies between objective and subjective measures in an attempt to determine the nature of these discrepancies. Schaeffer et al. (2022) assessed measures of executive functions that were thought to be ecologically valid and compared them to self-reported measures of executive functions and a control task. Fifty-nine participants ranging from ages 18 to 44 years were involved in the study. The outcomes showed that the supposedly ecologically valid measures of executive functioning were not better at detecting subjective deficits than the control task. Additionally, the tasks were not related to the self-reported impairments by the participants. Overall, the results demonstrate a lack of awareness of ecological validity, as the measures that were supposedly ecologically valid were no different than the control task in relation to the subjective measures of executive functioning.

Neuropsychological measures may not provide evidence about a client's cognitive abilities in everyday life, so it may be useful to determine if a relationship exists between objective and subjective measures. Furthermore, it may be possible to modify the objective measure to enhance the relationship with the subjective measure. A study done by Carrasco et al. (2017) analyzed objective memory tasks that represent daily memory performance and SMCs to determine if a relationship exists. The participants were 65 to 89 years old and 3,921 people participated in the study. The findings demonstrated a low correlation between SMCs and memory performance on the standardized measures. However, the findings did show that SMCs were positively related to health concerns. Subjective health concerns could be related to an

increase in distractibility and a decrease in attentional abilities, which could be affecting their memory and cognitive abilities (Carrasco et al., 2017). Stated otherwise, cognitive demands of a task can cause decreased information processing (Johansson, 2021; Leavitt & Katz, 2006). These results reiterate the idea that distractions could be contributing to cognitive deficits in the real world, and traditional cognitive measures do not accurately assess everyday cognitive abilities because they may not be ecologically valid. As such, ecologically valid measures are important so that neuropsychologists can use these measures to accurately diagnose their clients and provide adequate treatment recommendations.

Functional Cerebral Capacity

The idea that information from one's environment can reduce cognitive functioning may relate to the theory of functional cerebral capacity (FCC). *Functional cerebral capacity* is the idea that the brain has limited functional capacity, and the use of its resources can become demanding in the presence of intense stimuli (Harrison, 2015). That is, one may experience a decrease in the ability to inhibit unnecessary information (Leavitt & Katz, 2006). As Harrison (2015) explains, FCC can be conceptualized by comparing the theory to weightlifting. Individuals with different strength levels can differ in the amount of weight that they can lift. Similar to lifting weights, individuals with high levels of functional capacity are able to handle higher levels of stress and are able to handle strong emotional stimuli. These individuals can operate in a wide range of settings because they have a high limit for their functional capacity. Individuals with low functional capacity may operate appropriately in low stress situations but may have difficulty in moderate to high levels of stress.

Stressful environments could be described as a specific type of distractor because these types of environments can inhibit cognitive functioning. Distractions are shown to be related to

deficits in cognitive abilities. However, two conflicting and demanding cognitive functions may also cause conflicts in cognition (Carmona et al., 2009). Kinsbourne (1980) proposed an idea that the relatedness of a task and how close tasks are to each other in cerebral physical space correlates with the enablement or impairment of performance on multiple tasks at once. Specific brain areas are responsible for controlling specific processes, but multiple other brain areas might also be involved in these similar tasks, meaning that these brain areas might be competing (Comer et al., 2015). Performance increases when tasks are related and are close in proximity, and performance decreases when tasks are unrelated and are further away in physical cerebral space. When two similar tasks are asked of an individual at the same time, they will be able to perform those tasks at a similar, or better, level as if they were to perform them at two different times. However, if an individual is asked to perform two different tasks at the same time, they will perform at a lower level due to the requirement of added programming (Kinsbourne, 1980). The concept that Kinsbourne (1980) proposes is that of functional cerebral space. The more complex a task, the increasing level of pressure it puts on one's available resources, which will ultimately decrease cerebral capacity.

Haier et al. (1988) provides evidence for the theory of functional cerebral space, as the researchers examined regional brain glucose of normal young adult males while they completed abstract reasoning tests. By doing this, the researchers would be able to determine a relationship between cognitive ability and metabolic rates, which would allow them to assess brain structural patterns. The study included 30 healthy men between ages 18 and 30 years, and each participant was assigned to a different attention task. Positron emission tomography scans were given after the tasks were completed. Overall, the researchers found that poor performance on the tasks was related to an increased use of cortical activity throughout the brain. The result could indicate an

inefficiency of the neural circuit, where more neurons than necessary are used to complete the task. Therefore, those with more efficient circuits will have higher performance and less cortical activity broadly spread throughout the brain. In other words, they would perform better on the task because they are not using as much energy and networks throughout the brain.

Cappell et al. (2010) conducted a study that provides a good example of functional cerebral capacity. The idea was that older adults have lower levels of FCC and must use more brain functions than younger adults to complete a task and achieve the same outcome. Thus, using more resources for a task will cause them to have fewer resources available, and they therefore will have less resources to use for more demanding tasks. The study included 21 young adults (average age of 20.8) and 23 older adults (average age of 68.4). The participants first completed a set of neuropsychological assessments, such as Digit Span, Spatial Span, Wisconsin Card Sorting Test, Reading Span, the MMSE, and Digit Symbol. Working memory tasks were practiced so that they knew what to expect when in the functional magnetic resonance imaging (fMRI) scanner. Next, the participants completed verbal working memory tasks while in the scanner. Tasks fluctuated in difficulty level to determine differences in ability levels. Results showed that all age groups used some of the same brain regions during the working memory tasks. As the task load increased, young adults used additional brain areas compared to older adults. However, the older adults used the same brain areas in the low task load session as the younger adults used in the high task load session. The over-activation of brain areas in older adults therefore seems to be an attempt to compensate for lower functional abilities, as they scored similarly to the younger adults on these tasks. As the tasks increase in difficulty, the individual may attempt to block distractions while engaging in the task, which may have caused a decrease in performance and less brain areas to be used in comparison to the younger adults.

Cappell et al. (2010) found that, during the most difficult memory tasks, the older adults experienced under-activation and worse performance. They concluded that this may be due to the task going beyond their available resources. Hence, the individual reached their capacity and could no longer engage in proper processes to complete the task. The older adults could not recruit other resources to support these processes. All in all, the studies done by Cappell et al. (2010) and Haier et al. (1988) provide evidence that supports the theory of functional cerebral capacity. Based on this evidence, functional cerebral capacity is an important concept to understand in relation to ecological validity of standardized tests. Test developers must understand real-world implications on cognitive abilities in order to accurately measure daily cognitive functioning through formal assessments. These implications include distractions and the limited capacity of individuals to function in complex environments. However, research indicates that current neuropsychological assessments are lacking in ecological validity, which demonstrates a need for measures that includes an assessment of how distractions and stressors impact how individuals' function in their environment.

Problems with Ecological Validity of Current Neuropsychological Tests

Given the importance of ecological validity, it is surprising that little evidence exists to support the ecological validity of traditional neuropsychological assessments. Neuropsychological assessments are assumed to assess a patient's memory as it applies to the real world (Davidson et al., 2016). Contrary to how neuropsychologists use these measures, many of these tests have weak relationships between the clinic and real-world cognitive performance. Although subjective memory questionnaires can be used to assess memory functioning in everyday life, they rely on the client's subjective information which may or may not be accurate (Makatura et al., 1999). Therefore, an objective measure of memory and

cognitive functioning is a better option for assessing these abilities. Several studies have been conducted to assess the ecological validity of current measures. However, these studies actually seem to assess predictive validity, rather than ecological validity.

The CVLT is a neuropsychological assessment that measures memory encoding, learning, retrieval, and recognition of verbal information (Crean, 2003; Williams et al., 2021). Although widely used, the CVLT has little evidence for the ecological validity of the test. For instance, Williams et al. (2021) conducted a study to determine the relation between verbal declarative memory and employment because learning and memory are likely to be involved in many work environments. Specifically, instructions relayed from bosses to employees are meant to be learned and remembered for later. So, they examined word-list learning from the CVLT to determine process strategies as they related to employment status. The researchers conducted the study using a community sample of 179 individuals between the ages of 27 and 58 years who were involved in a comparison group within an NIH-funded study about the effects of HIV on memory. The participants were confirmed with negative HIV results. The participants were administered the CVLT- Second Edition in which they were read 16 words and were given a series of five learning trials. They were instructed to recall the words after a short delay and a long delay, recognize the words by giving Yes/No answers, and a trial in which they must decide if the words listed were in the original list (forced choice trial). Semantic clustering, serial clustering, and subjective clustering were calculated by the CVLT software, and these strategies were also used to compare to employment on the basis that these strategies may be used in occupational settings. The overall results showed that employed participants had higher cognitive capacities than those that were unemployed, as the employed group had higher semantic clustering scores than the unemployed group. Therefore, the findings show that the processes

used during the CVLT can be predictive of employment status in the real world. Creating semantic categories may be useful memory strategies in the workplace and this could be why they are used by employed individuals.

Williams et al. (2021) states that understanding the relation between memory and employment can be useful for intervention planning and recommendations for those who are planning to return to work. Given these findings, the CVLT was thought to have ecological validity. Crean (2003) had similar results in their study of patients with mild traumatic brain injury (mTBI) and their ability to return to work. The CVLT was determined to be able to predict work outcomes, as individuals with cognitive impairment were at risk for impairment on the job. Furthermore, the assessments used in this study seemed to predict those who would return to work more accurately than predicting those who would not return to work. Kibby et al. (1998) also claimed to assess the ecological validity of the CVLT based on a sample of participants with closed head injuries. They administered a questionnaire that assessed job performance to compare to CVLT scores. Certain subtests of the CVLT were significantly related to job performance, such as long-term recognition memory and total recalled words. However, ecological validity is not established for the CVLT based on this evidence. Rather, the studies by Crean (2003), Kibby et al. (1998), and Williams et al. (2021) were predicting an outcome of returning to work, job performance, or status of employment, and hence were investigating predictive validity of the CVLT, which is different than ecological validity.

Whereas the previous studies were assessing predictive validity and not ecological validity, other studies have used methods that more closely assess the idea of ecological validity. A study done by Davidson et al. (2016) determined that the CVLT is not ecologically valid. The goal of the study was to assess memory of people with mild cognitive impairment (MCI) and

determine the relation between subjective memory complaints and standardized assessments of memory. The researchers wanted to determine if the CVLT and Logical Memory (LM) standardized assessments were generalizable to the real-life memory abilities of the participants. The study was conducted on 53 older adults aged 57 to 84 years who were currently MCI patients at the regional memory disorders/dementia clinic. All participants were administered a neuropsychology battery that consisted of the Montreal Cognitive Assessment (MoCA), which is a brief measure of cognitive ability, the Forward and Backward Digit Span from the Wechsler Adult Intelligent Scale (WAIS) as a measure of working memory, and the Wisconsin Card Sorting Test (WCST), Boston Naming Test, and Stroop as measures of executive functions. All participants were also given two episodic memory tests, the Logical Memory I and the CVLT-Second Edition. The 5-word delayed recall subtest from the MoCA was also included in the episodic memory assessment. Once their visit was complete in the lab, they received a phone call at a random time from one of the researchers in which they were asked questions from a questionnaire about their experience in the lab. After the questionnaire, they were asked questions about how they perceive their own memory. The questions consisted of concepts about the vividness and clarity of their memory, and the interest they experienced during the assessments. As expected, the MCI patients scored lower than the control group on the neuropsychological assessments and the measures of episodic memory. For the self-reported questionnaire, the MCI patients reported having worse memories of the lab and testing session than the control group. Specifically, they reported that they had less vivid memories about specific details of their visit to the lab. The difference in memories of the lab could be due to the MCI patients being too distracted by the tasks to pay attention to the lab. Pearson correlations were used to correlate the standardized memory measures with the memories of the lab visit.

Although a positive correlation seems to exist between self-reported memory and clinical memory assessment, the correlation was not found to be significant. Therefore, extraneous factors may have influenced these findings. These findings show that the neuropsychological assessments and subjective recollections of the lab are not significantly related, so competent assessments for cognitive abilities in the real world may need to be created.

As previously shown, researchers have investigated ecological validity by examining the relation between subjective measures of memory and performance on formal neuropsychological tests of memory. Other studies have been conducted to investigate the ecological validity of neuropsychological tests by administering a test with purported ecological validity in succession with the measure that they are investigating. The relation between the results on these measures allows for an interpretation of the ecological validity of the test. The Wechsler Memory Scale (WMS) is another neuropsychological assessment that lacks evidence for ecological validity. A study conducted by Rzezak et al. (2017) compared traditional neuropsychological memory tests, including the WMS, to measures of everyday memory which were subtests from the RBMT, to determine the correlation between these two types of measures. The participants included 31 patients with temporal lobe epilepsy caused by hippocampal sclerosis (TLE-HS) and 34 healthy control participants. The results indicated that the TLE-HS group scored significantly lower on the traditional and everyday memory tests compared to the control group. However, the RBMT was only related to half of the traditional memory measures, so the full battery of the traditional memory assessment may not be ecologically valid, as they are not completely related to the everyday memory assessment, the RBMT.

Another study by Makatura et al. (1999) also compared traditional memory assessments with the RBMT to determine the usefulness of these traditional memory assessments in

evaluating functional memory impairment on 119 adults with brain injuries. These traditional memory assessments included the WMS and the Luria Nebraska Neuropsychological Battery (LNNB). These measures were also compared to scores on the Clinician's Memory Rating Scale (CMRS), which is a subjective measure of memory based on the clinician's assessment. The CRMS was used to determine that there were four specific groups of participants- those with severe memory impairment, moderate memory impairment, mild memory impairment, and no memory impairment groups were observed. The main findings demonstrated that the RBMT was more accurate in detecting everyday memory functional abilities than the WMS or LNNB, as indicated by the impairment levels that were created using the CMRS. The WMS was slightly more accurate in identifying the severity of memory impairment than the LNNB, but both were much less accurate than the RBMT.

Drozdick and Cullum (2010) completed a study in which they sought to investigate the ecological validity of the WMS and WAIS by using the Texas Functional Living Scale (TFLS) as comparison. The participants were split into clinical groups and normative groups. Overall, the results showed that the WMS and the TFLS had small to moderate correlations between each other. However, correlation for only the clinical group between the WMS and TFLS ranged from moderate to high, which indicates that there is a moderate to high correlation between memory and functional abilities when assessing someone who is showing impairment. To determine everyday functional abilities of patients, it may be better to use the TFLS and traditional measures combined, rather than traditional measures alone.

Similar to the research conducted with the CVLT, the correlation between a traditional clinical measure of memory and a functional living scale is not true ecological validity but is instead predictive validity. Noting that the clinical measures of memory are related to a

functional living scale allows for a prediction of how well the clients may function, but not how the distractions and stressors of everyday life may impact memory functioning. Ecological validity of an assessment is an important aspect of measuring cognitive functioning because it is essential in determining the patient's memory and cognitive abilities in daily life. Additional studies should be conducted to determine how these results can be translated to cognitive abilities in daily life. As such, future research must be consistent with defining and measuring ecological validity, and more studies must be conducted to determine the ecological validity of neuropsychological assessments. Additionally, the aforementioned studies used patient populations to assess the ecological validity of neuropsychological tests. Future research would benefit from using healthy individuals, as neuropsychological tests are also normed on this population. Due to the fact that the ecological validity of neuropsychological tests is debatable, other studies have been conducted to change the way in which the assessment process is conducted as a method to increase the ecological validity of the neuropsychological assessment process entirely.

Attempts to Increase Ecological Validity of Neuropsychological Assessment Processes

Everyday activities involve many cognitive processes that interact with each other, and this can be a difficult process for individuals with cognitive deficits. Although these are tasks that may not take much thought, several cognitive abilities are still involved in these tasks. People with cognitive deficits may experience difficulties with these commonplace tasks (Robertson et al., 2017). Observing a client in a clinical setting may not accurately reflect their level of cognitive functioning in everyday situations because the testing conditions do not reflect their typical everyday life (Chaytor et al., 2006). Clients can receive more accurate treatment and recommendations from their provider when ecologically valid measures are used, as it more

accurately relates to their true level of cognitive abilities. Due to the importance of ecological validity, many researchers have attempted to increase the ecological validity of neuropsychological assessment processes by assessing patients in naturalistic environments and simulated environments using virtual reality.

Weakley and Schmitter-Edgecombe (2019) also assessed how individuals complete real-world tasks. The authors state that many people have experiences in which they are interrupted when completing a task, and subsequently do not complete the initial task. The deficit that these individuals are experiencing may be because the process of returning to a task after an interruption requires the use of multiple cognitive processes (Robertson et al., 2017). Being able to return to an initial task is important to ensure that one is completing everyday tasks. Thus, Weakley and Schmitter-Edgecombe (2019) sought to determine how clinical measures of cognitive abilities relate to cognitive performance on real-world tasks, and if assessing cognitive abilities in real-world environments might be more ecologically valid than clinic-based tests. Participants were age 50 years or older, in which 32 participants had MCI and 64 participants were healthy older adults. The participants first completed a testing session that included neuropsychological tests in a lab or office setting. The next testing session consisted of the participants completing Instrumental Activities of Daily Living (IADL) assessments in a specific apartment used for this study. Video cameras were used to watch the participants and score their behaviors. The interruption task took place in the apartment, in which participants had to respond to tasks in the apartment as they would in their own home. These tasks included answering the phone when it rang or opening the door when someone knocked. During the interruption task, MCI participants performed at a slower rate compared to the healthy adults. Furthermore, MCI participants made more errors than the healthy adults when completing tasks in the interruption

condition, in which they also substituted more tasks than the healthy adults. Based on these results, patients with MCI seem to be more affected by interruptions in their environment than healthy adults. An important result is that those with MCI performed similarly to controls on tasks that did not include interruptions, leading to the conclusion that interruptions are related to increased deficits in cognition in individuals who already experience impairment, and therefore negatively impacts performance. Although the results indicate that interruptions affect individuals with MCI more significantly than the cognitively healthy older adults, the IADL task is not a measure of cognitive functioning. Rather, the task is a measure of physical functioning that could potentially demonstrate impairments that are caused by cognitive impairments. The comparison of the IADL task and traditional cognitive measures is therefore invalid, as the IADL task is not directly measuring cognitive impairments in the same way as the traditional measures.

Imitating real-world situations to understand cognitive performance is an important aspect of ecological validity, as it is imperative to know how people behave in their everyday environments. However, one's home is not the only environment in which they must use their cognitive abilities to complete tasks. Grocery stores and other establishments are also part of people's everyday experiences. Grocery store tasks were of particular interest in a study done by Robertson et al. (2017). The study was conducted on individuals with neurologic conditions and healthy controls in a lab that was transformed to resemble a grocery store setting. During the assessment, the participants were asked to complete two main steps. The first step was to create a recipe and prepare a list of items that would be needed from the "grocery store." The second step was to execute the grocery shopping within the lab. Similar to the results from Weakley and Schmitter-Edgecombe (2019), the participants with neurologic conditions had lower performance than the control group on the shopping task, as they required more cues and more time to

complete the task. Scores from the shopping task were able to differentiate between the neurologic and control groups. The participants also completed traditional neuropsychological tests, and the results from these measures were not able to differentiate between the two groups. So, the naturalistic environment was better for determining everyday functioning than the traditional measures. Even when assessing older adults compared to younger adults when completing a computer simulated grocery shopping task, older adults had a more difficult time completing the shopping task, especially in the face of a distraction (Knight et al., 2008). Specifically, performance on memory tasks seems to decrease much more in older adults than younger participants when a distraction is added. Also, when distractions are included, older adults fail to use compensatory strategies when they normally would if there were no distractions (Knight et al., 2008).

Overall, naturalistic environments can be used to determine how a client functions and uses their cognitive abilities in their everyday environments, which is an ecologically valid way to assess memory and cognitive functioning. Moreover, adding a distraction to this naturalistic environment may further the ecological validity of the assessment, as people consistently experience distractions in everyday life. Nonetheless, faults still exist within these studies. First, the studies by Knight et al. (2008), Robertson et al. (2017), and Weakley and Schmitter-Edgecombe (2019) did not strictly use cognitive measures for the real-world tasks, but instead also included behavioral functioning. So, the researchers are not independently increasing ecological validity, but are also indirectly attempting to increase criterion validity, which is not a true measure of ecological validity. Furthermore, many factors are considered as to why assessment in real-life settings is inconvenient, such as travel, limited standardization, and safety. For example, the simulated grocery store that was made in the lab in the study by Robertson et

al. (2017) is not convenient or reasonable to create for every client. Therefore, other options should be considered when attempting to increase the ecological validity of cognitive measures.

Virtual Reality (VR) is another option that has been studied to try and increase ecological validity of the neuropsychological assessment process. VR can be used to increase the ecological validity of studying human language and communication, which is typically done in a controlled lab setting (Peeters, 2019). VR has also been shown to be used for rehabilitation of patients with acquired brain injuries (Mendes et al., 2021). Ecologically valid forms of rehabilitation are important as they more accurately reflect the patient's everyday life compared to when rehabilitation is conducted in a controlled setting. The most relevant use of VR is the use in assessment of patients with cognitive impairments. Because neuropsychological assessments appear to lack ecological validity, VR has become of interest to measure clients in a way that represents their true levels of cognitive abilities in the real world (Giglioli et al., 2019).

One specific study by Giglioli et al. (2019) was conducted to assess the ecological validity of Serious Games. Serious Games are computer games that are used for investigating behavioral functioning, which can be done in 2D on a computer or 3D with a VR headset. The study was conducted in which they used 2D and 3D versions of Serious Games to assess executive functions and compared both versions with standard clinical tasks. The participants included were 94 healthy individuals, with half of the participants in the 3D group and half in the 2D group, and all participants completed standard tasks and self-report measures. Self-report questionnaires included measures that assess cognitive flexibility, attention, impulsiveness, inhibition, planning, and cognitive flexibility. The standard tasks were completed on computers and assessed the same abilities as the self-report measures. The Serious Game task was a simulated situation in a spaceship in which the participant was to discover new land for living.

Eight missions were to be completed during this simulation. These missions were intended to be tasks that assess attention, impulsivity, inhibition, planning, and cognitive flexibility. Generally, the 3D version of the Serious Game task seemed to be better at assessing cognitive deficits than the 2D version of the Serious Game, as it generated lower reaction times, higher correct answers, and less perseveration with attentional abilities, inhibition control, and cognitive shifting. The 3D version provided more accurate responses than the 2D version. The 3D task was more highly correlated than the 2D task to the self-report questionnaires. Importantly, the VR assessment of executive functions was correlated with the standardized paper-pencil assessments and provided a naturalistic environment for the participants to function. An ecologically valid assessment using VR was created due to the naturalistic environment and the assessment of the same abilities as the standardized assessments. Based on these results, the 3D version of the Serious game task seemed to allow participants to act according to a more natural environment. Although the results seem to show that VR is an ecologically valid method to conduct neuropsychological assessments, other modes of measurement may prove to be more efficient.

Computerized versions of neuropsychological tests and the paper-pencil versions of these tasks may not be ecologically valid as they do not include distractions that are typically involved in everyday life (Asbee et al., 2022). These environments may lack the ecological validity needed to translate results to real life. Alternatively, experimenting in a strictly controlled environment is beneficial as it allows for conclusions to be made about how one variable may influence another (Peeters, 2019). Therefore, increasing ecological validity of paper-pencil tests used in the clinic is important so that both of the important aspects of controlling variance and increasing translation to real-world cognitive abilities can be accomplished. Virtual reality assessments may seem to be one way to achieve this goal. However, VR is not practical to have

in the clinic. The equipment is expensive, the testing room may be too small, the clinician may have to teach clients the proper use of the equipment, and it is not "real life," it only simulates this experience, so the true ecological validity of VR remains uncertain. Hence, the goal is to find a way to increase the ecological validity of traditional paper-pencil measures by adding cognitive processes that occur during natural, everyday cognitive tasks.

Weakley et al. (2019) explains that older adults tend to rely on compensatory strategies, such as calendars or lists, to support their completion of complex cognitive tasks in their everyday lives. The researchers decided to analyze adults in their homes to determine how they used compensatory strategies and the impact of these techniques on their cognitive abilities. The idea is that analyzing an individual in their own home would increase the ecological validity of the assessments. Interestingly, people with average cognition levels were shown to benefit from using compensatory strategies during prospective memory tasks, which could improve their performance to a high average level. However, compensatory strategies did not help individuals with low average cognition. Basically, compensation can be useful for some, but not all individuals during everyday tasks. Specifically, they may not be useful for individuals who are already suffering from cognitive deficits. Therefore, including the assessment of compensatory strategies into naturalistic assessment of individuals demonstrates the variance between the scores, which could show why people perform differently in the real world. Nevertheless, there is still a need for a specific assessment that can determine these differences.

Chaytor et al. (2006) conducted a study in which they assessed how cognitive processes that occur during everyday tasks affect formal testing scores. The researchers assessed the ecological validity of executive functioning assessments. Specifically, the researchers assessed how environmental demands impact neuropsychological assessment performance and how the

use of compensatory strategies affected performance. The study was conducted on a sample of 46 participants who were recruited from consecutive outpatients referred for neuropsychological assessment at several different medical centers. Participants were diagnosed with TBI, epilepsy, and other diagnoses such as a tumor or MS. Executive functioning tests included the WCST, Trail Making Test (TMT), Stroop, and Controlled Oral Word Association test (COWAT). The Dysexecutive Questionnaire (DEX) and the Brock Adaptive Functioning Questionnaire (BAFQ) were given as subjective measures of executive functioning in everyday life and were completed by significant others. The DEX and BAFQ are completed by significant others in which they use a Likert scale to rate how often they observe specific executive problems in their spouse. Both questionnaires assess similar areas of functioning. Nontraditional cognitive factors such as environmental demands and compensatory strategies were analyzed at the end of the assessment process. Compensatory strategies and environmental demands were also assessed using the DEX and BAFQ measures. The DEX included items that asked significant others whether they notice their partner using any compensatory strategies and if their partner encounters any environmental demands when experiencing executive deficits. The results of the study show a difference in correlations between the executive tasks and the subjective measures. A high correlation was found between the Stroop and the DEX. A high correlation was also found between the BAFQ, Stroop, and Trails B. The WCST and COWAT were not related to the subjective measures. As for environmental demand and compensatory strategies, they were both related to the DEX. This result demonstrates that people in demanding environments may exhibit more executive failures in their daily life and, therefore, are more likely to use compensatory strategies. As some of the executive assessments were related to the subjective measures as well as nontraditional cognitive factors, this shows that these outside variables may be influencing patients' results on the

traditional measures. Therefore, the nontraditional cognitive factors are important to determine everyday cognitive abilities.

Overall, naturalistic and simulated environments can be used to provide an ecologically valid environment for neuropsychological testing, but these environments are not practical for the clinic. Nontraditional cognitive factors were shown to influence the results of traditional paper-pencil measures of executive functioning (Chaytor et al., 2006). Thus, to increase the ecological validity of a formal neuropsychological measure, including these factors may increase the ability to accurately determine a client's cognitive abilities in their everyday life. Neuropsychologists must understand that external factors need to be included in their overall analysis of the client. Hence, increasing the ecological validity of formal neuropsychological assessments is important, as it will allow for external environmental factors to influence the results as they would in the client's daily life, and would result in more accurate depictions of the client's true abilities.

Solutions to Increasing Ecological Validity of Formal Assessments

Adding distractions and stressors to formal neuropsychological measures may aid in increasing the ecological validity of the original assessment. Neuropsychological measures are used to diagnose patients and recommend treatments, but also to provide information about their cognitive functioning. The clinicians must estimate their client's abilities based on the obtained results from the neuropsychological batteries (Lynch, 2008). However, many neuropsychological tests were not normed in a way that is useful for estimating client's abilities in everyday life. The neurocognitive abilities that are demonstrated on standard clinical measures are different than the way these abilities are demonstrated during everyday tasks (Heinrichs, 1990). Expanding the content in which traditional assessments are validated to measure will allow neuropsychologists

to understand the interaction of cognitive abilities and the environment, and how this interaction affects their client's cognitive functioning in everyday life (Heinrichs, 1990).

Adding a distraction to the standardized testing environment is one way that researchers could enhance the ecological validity of formal neuropsychological assessments. Randolph et al. (2017) completed a study using MS patients, on the basis that the difference between objective cognitive measures and subjective cognitive complaints could be due to the highly controlled testing environment, which limits outside distractions that might be present in the patient's natural environment. The research included 28 patients with MS ranging from ages 18 to 65 years. The goal of the study was to determine if an auditory distraction would negatively affect formal test performance in a way that would more accurately resemble subjective cognitive complaints than the traditional standardized test setting. The participants were administered a battery of neuropsychological functioning. The battery included a measure of attention and executive functioning (Symbol Digit Modalities Test-Oral; Paced Auditory Serial Addition Test 3 Second Trial; D-KEFS Trail Making Test-Number Letter Trial and Color-Word Interference - Inhibition Trial) and verbal episodic memory (California Verbal Learning Test-II, Total Learning). They also completed two measures of subjective cognitive dysfunction, the Perceived Deficits Questionnaire and the Fatigue Severity Scale. The background noise that was used as the distraction resembled a busy office, and participants were randomly assigned to either a noisy and then quiet assessment, or a quiet and then noisy assessment. Alternate forms of the measures were used during the second trial. Randolph et al. (2017) found that the performance of the individuals with a higher level of cognitive complaints performed worse during the noisy condition than those with less cognitive complaints. In fact, there was no difference in performance during the quiet condition between those with and without cognitive complaints.

Specifically, the worse performance was seen during the assessment of processing speed. Overall, the auditory distraction had a negative effect on processing speed in participants with subjective cognitive complaints. Therefore, adding an auditory distraction provided information from the environment that appeared to detract from the participants' performances, likely to represent the participants' daily cognitive abilities due to environmental distractions.

Aurtenetxe et al. (2016) proposed that there are two different types of interference when engaging in cognitive tasks- interruption and distraction. Interruption is described as a secondary task that gets in the way of an initial task, and distraction is described as irrelevant stimuli in which an individual must ignore to complete a task. The researchers completed a study in which a group of patients with MCI and a group of healthy participants all performed tasks in non-interference, distraction, and interruption conditions. All participants performed poorly in the interruption and distraction conditions compared to the non-interference condition. Patients with MCI specifically exhibited an alteration in their ability to inhibit irrelevant stimuli and performed worse on all the tasks than the healthy participants. The deficit observed in the MCI patients in the presence of a distraction may relate to the issue of functional cerebral capacity. As patients with MCI have less capacity, they have an increased problem with performing more than one cognitive task at a time. So, when attempting to complete a task while also trying to inhibit a distraction, this causes less cognitive resources to be allocated to the original task and causes a decrease in performance compared to the healthy individuals. As such, when there is a concern about ecological validity of the standardized test, it is important to include distractions during standardized assessments to determine how this type of environmental demand will impact the client's performance. Including distractions in an assessment is especially important because individuals with cognitive impairment may not show significant impairment on a formal

assessment. The idea that some participants perform within normal limits could be because the clinic is too "perfect" of an environment, meaning that it may be free of distractions in comparison to real life (Schnabel & Kydd, 2012). Researchers have been attempting to develop tests for the clinic that more closely relate to the challenges of everyday life.

Schnabel and Kydd (2012) conducted a study in which the goal was to develop a short standardized neuropsychological assessment that included a distraction, which therefore could increase the ecological validity of the test. Specifically, an audio-visual distraction was administered during a primary attention task. The researchers also wanted to use the results of this study to examine what effect this distraction had on mTBI and major depressive (MDE) patients compared to healthy participants. A total of 240 people participated in the study, with 80 participants having mTBI, 80 MDE participants, and 80 healthy individuals. The participants went through 2 conditions. The first condition was a demanding primary attention task in which the client is told to be fully engaged. The second condition uses the same procedure as the first condition, but a competitive environmental stimulation (distraction) interferes with the attention task. For the distraction component, the researchers created two audio-visual recordings, the first shows a woman reading a news item and the second shows the same woman reading random numbers. The WMS LM was given to participants during the news recording, and Digit Span and Letter Number Sequencing (LNS) from the WAIS were given during the random number recording. Important results were found from the study. First, the baseline testing with no distractions showed similar results of all 3 groups, demonstrating that the participants with mTBI and MDE performed at average levels in the standardized and controlled environment. The mTBI group had a significant decrease in performance during the distraction condition, but the MDE patients had a significant increase in performance with the distraction. The MDE group

stated that they were distressed during the distraction condition, so it is unknown as to why their performance improved and it is unknown if their performance would continue to show improvement if the tests lasted longer. Importantly, the control participants did not experience a difference in performance during the distraction condition. Thus, it seems that the healthy participants were able to successfully block out the distraction. Overall, the distraction proved to affect the performance of the patient groups. Even though the MDE patients improved in performance, this still demonstrates a difference from their baseline performance, suggesting that their performance is different in their daily lives because of the distractions that exist. To summarize, adding a distraction to a traditional assessment can increase the ecological validity of the measure, while also protecting the overall standardized environment that is necessary for keeping the measure reliable and valid.

Conclusion

Overall, establishing the ecological validity of formal neuropsychological measures is important. Patients may present to the clinic with memory and cognitive complaints, but their scores on the assessments may provide information which says that they are unimpaired (Caramelli & Beato, 2008; Leavitt & Katz, 2006; Pearman et al., 2014; Randolph et al., 2017). The impairment that the patients are experiencing is likely due to distractions and other stressors that are experienced in their everyday environment (Carrasco et al., 2017; Chaytor et al., 2006; Weakley et al., 2019). To accurately capture the way that patients cognitively function in their daily lives, neuropsychological measures must be ecologically valid.

Neuropsychological tests tend to lack ecological validity (Rzezak et al. 2017; Williams et al., 2021). Even though current research has claimed to have found that formal neuropsychological measures are ecologically valid, these studies were actually assessing the

predictive validity of these assessments, rather than ecological validity. For example, when examining the ecological validity of the CVLT, researchers concluded that the CVLT was ecologically valid based on its ability to predict an outcome of returning to work, to predict job performance, or to predict status of employment (Crean, 2003; Kibby et al., 1998; Williams et al., 2021). These all provide evidence of predictive validity rather than ecological validity.

Several studies have been done to change the way that neuropsychological assessments are conducted so that they would be more ecologically valid, such as using naturalistic environments, or simulated environments such as VR (Giglioli et al., 2019; Weakley & Schmitter-Edgecombe, 2019). However, these types of assessments are not practical to use in the clinic and do not contain norms in which to compare client results. Hence, increasing the ecological validity of formal neuropsychological measures is more practical to attain.

Increasing the ecological validity of traditional assessments can be done by adding a distraction to the formal assessment (Chaytor et al., 2006). Adding a distraction to the traditional assessment may cause reductions in performance, which may more accurately relate to the client's daily cognitive abilities. The idea that distractions and stressors can reduce an individual's cognitive abilities may be related to the idea of functional cerebral capacity, which states that every person has a capacity at which their brain functions. If deficits to these cognitive abilities are present, then cerebral capacity will decrease (Harrison, 2015, p. 406-415). Therefore, an individual with higher FCC will more easily block out distractions (inhibit a response) to complete a task, whereas an individual with lower FCC will have more difficulty with inhibiting a response to a distraction while attempting to complete a task (Haier et al., 1988; Leavitt & Katz, 2006). As distractions are prevalent in our everyday lives, it is important to have assessment procedures that reflect how these distractions may affect people, especially for those

with cognitive impairment. Adding a distraction to a formal neuropsychological measure allows for the standardized environment to be maintained while also creating an environment that more closely resembles a client's real life, where there are distractions and stressors (Aurtenetxe et al., 2016; Schnabel & Kydd, 2012). As stated, research on adding distractions to traditional neuropsychological assessments is limited.

To add to the limitations of the current literature, previous research for assessing the ecological validity of neuropsychological assessments is lacking in several ways. Traditional neuropsychological assessments are not related to subjective memory complaints. So, when a client presents to a neuropsychology clinic with memory complaints, their scores on the objective measures may show opposing results compared to their subjective view. These opposing results also demonstrate that performance on the objective measures is different than the abilities of the client in their everyday environments, where distractions and stressors are present. Some studies have been conducted in which distractions have been added to the standardized testing session, but normative data for formal measures with an added distractor does not currently exist. Additionally, no tests have been developed that are inherently ecologically valid. Existing research has been conducted with a retrospective approach due to the fact that the researchers were assessing the ecological validity of current measures. Therefore, the purpose of this study is to take a prospective approach to the problem of ecological validity.

The main goal of this study is to create a formal neuropsychological memory assessment that is inherently ecologically valid and collect normative data for this assessment. The Hopkins Verbal Learning Test-Revised (HVLT-R) will be used to assess healthy individuals, and an audio distraction will be introduced to the test setting while the participant takes the test. The HVLT-R is an ideal assessment to use for the study as it is quick and easy to administer, has 6 choices of

alternative forms, and does not have a ceiling effect in recall for normally functioning individuals (Brandt, 1991). The HVLT also assesses a wide range of memory functions in those aged 16 years or older, such as total recall, learning, delayed recall, retention, and recognition (Brandt, 1991). Total recall, delayed recall, and retention scores will be assessed for the purpose of this study. Two hypotheses exist for this study. The first hypothesis states participants will perform worse on the HVLT-R when the audio distraction is present. The second prediction is that those with higher scores on the Subjective Memory Questionnaire (SMQ) will have lower scores on the HVLT-R with the audio distraction than those with lower scores on the SMQ.

CHAPTER II: METHOD

Sample

Twelve men and 36 women participated in this study. Participants' ages ranged from 18 to 55 years (M = 22.67, SD = 6.71) (See Table C1). Participants in this study had no history of a serious head injury and no current neurological illness. A serious head injury is defined as a loss of consciousness for more than 2 minutes. Seven participants that were included in this study indicated that they had a concussion at some point in their life but did not lose consciousness for more than 2 minutes. A power analysis was conducted by hand, with the alpha set at .05 and an effect size of .27, that resulted in a sample size of 47 that should be used for the study to achieve a power of .20 (Hays, 1994). The participants consisted of students from the Middle Tennessee State University undergraduate volunteer database. Therefore, the participants were college-aged individuals. The participants were treated in accordance with the ethical principles of the APA.

Instruments

Hopkins Verbal Learning Test-Revised (HVLT-R) (Brandt, 1991)

The HVLT-R, developed by Brandt (1991), is a quick measure of verbal learning and memory that is normed for those who are 13 years old or older. The test has 6 alternate forms, and these forms have been shown to be equivalent to each other (Brandt, 1991). Form 1 and form 3 was used for the purpose of this study. The initial test takes about 5 to 10 minutes to administer. A 20-to-25-minute delay occurs after the initial test, and then the recall and recognition portions of the test occur immediately after the delay. The test is administered to the participant in which the examiner writes down the examinee's responses. The test consists of 12 words that are read at a rate of one word every 2 seconds. The same list of words is read three times, and then the delay occurs. At the end of the delay, the examinee is to recall the same list of words. The last portion

of the test requires the participants to recognize words that were on the list from a new set of words. The dependent variable of the HVLT-R is a memory score, which consists of the examinees total recall, delayed recall, retention, and recognition index scores. Total recall, delayed recall, and retention scores were used for the current study. The highest possible raw score for total recall is 36 points and for delayed recall is 12 points. The possible range in scores for recognition ranges from -12 to 12 points.

The HVLT-R was normed on 1,179 cognitively normal individuals (Brandt & Benedict, 2001). The test has well-established test-retest reliability, with coefficients of .39 for retention, .66 for delayed recall, and .74 for total recall (Benedict et al., 1998). The test also has well-established alternate form reliability, as all the alternative forms were found to be equivalent. Lastly, the test did not show practice effects when alternate forms were used for the assessment (Benedict & Zgaljardic, 1998). As for validity, the HVLT-R was shown to have convergent validity with tests that measured similar constructs, such as the CVLT (Lacritz et al., 2001).

Shipley-2 (Shipley et al., 2009)

The Shipley-2 is a brief measure of cognitive functioning and impairment normed for ages 7 to 89 years. Three different subtests can be administered for this test, which include vocabulary, abstraction, and block patterns. Only two of the subtests need to be used to determine a measure of cognitive abilities, therefore the vocabulary and abstraction subtests will be used for the purpose of this study. The test takes about 20 to 25 minutes to administer when only using two subtests. Each subtest is administered with paper and pencil and the examinee is to follow instructions for each subtest and complete them accordingly within a specific time limit. For vocabulary, the examinee has 10 minutes to answer 40 questions where they must identify a word in a list of words that have the same meaning as a given target word. Abstraction

requires the examinee to fill in blanks to complete a given sequence, and they have 12 minutes to answer 25 questions. The dependent variable for the Shipley-2 is the standard score and the AQ and BQ scores. The vocabulary subtest contributes to both scores, abstraction only contributes to the AQ score, and block patterns contributes only to the BQ score. The standard score provides a brief measure of cognitive ability, and the AQ and BQ scores provide an index of cognitive impairment. The standard score will be used to determine the participants' cognitive abilities, and the AQ score will be used as a supplementary score to determine any evidence of cognitive impairment.

Shipley et al. (2009) states that the Shipley-2 was standardized on a sample of 2,826 individuals from 16 different states. The sample was meant to represent the average U.S. population in regard to sex, education level, and minority representation. The test also has established reliability and validity. For adults, the reliability of the Shipley-2 is estimated to have an internal consistency of .88 to .97 and test-retest reliability of .74 to .94. As for validity, the test is estimated to have good content validity, good construct validity, and good concurrent/discriminant validity.

Subjective Memory Questionnaire (SMQ) (Bennett-Levy & Powell, 1980)

The SMQ is a measure of subjective memory functioning as perceived by the examinee. The questionnaire asks 43 questions about memory and potential problems with memory functioning. The questions can be answered with one of five choices ranging from *very bad to very good*, or *very rarely to very often*. Examples of the questions include, "How good is your memory for names of people," and "Do you often find you are unable to place voices you've heard before." Lower scores indicate that the examinee subjectively experiences a greater amount of memory problems. The dependent variable of this questionnaire is a score to

determine perceived memory functioning. Scores on the SMQ range from 1 to 5, with 1 indicating participants view their memory to be poor and 5 indicating participants view their memory to be good.

Bennett-Levy & Powell (1980) reported that the SMQ was developed on 141 participants that ranged in age from 16 years to 70 years old. Participants consisted of healthy volunteers, postgraduate students in clinical psychology, and chronic alcoholics compared to individuals that drank only rarely. The scores for participants were grouped into age ranges 16-24, 25-34, and 35+ years. As for psychometric properties, the SMQ has been shown to be reliable and valid in measuring subjective memory. Test-retest reliability correlation coefficient was at a high of .86, and the average of the reliability of all the items was a coefficient of .64. For validity, the SMQ has been shown to be correlated with objective memory tests, such as Logical Memory, Partial Figures and Words, Face-Name, William's Object Memory, and Digit Span, with correlation coefficients ranging from .32 to .56.

Medical History Questionnaire (MHQ) (Foster, 2023)

The MHQ is a brief questionnaire given to the participants to assess specific aspects of their medical history. Examples of questions include, "Do you have a history of significant head injury or concussion?" and "Do you have a history of neurological or psychological/psychiatric illness?" The participants are to complete the questionnaire themselves using paper and pencil.

The Center for Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977)

The CES-D is a short self-report scale that consists of statements that measure depression. Participants are asked to answer items about how they have felt during the past week, such as, "I was bothered by things that usually don't bother me," "My sleep was restless," and "I felt that people dislike me." The participants are to respond to the statements with answers

ranging from "*Rarely or None of the Time* (Less than 1 Day)" to "*Most or All of the Time* (5-7 Days)." The dependent variable of this assessment is a measure of the participant's current level of depression.

Radloff (1977) reported that the CES-D was standardized on 4,996 individuals, randomly selected from households in Kansas City, Missouri and Washington County, Maryland and were 18 years old or older. Clinical validation studies were conducted with psychiatric patient samples (N=70). The CES-D has adequate reliability. Internal consistency reliability estimates range from .84 to .90, and test-retest reliability estimates range from .51 to .67. The CES-D has also been shown to have discriminant, convergent, and predictive validity. The scale was able to discriminate between the psychiatric inpatient and general population samples. Lastly, the scale has been shown to predict life events, as those with more negative life events scored higher on the depression scale.

Procedure and Design

The participants were seated in a controlled research room at Middle Tennessee State University and informed consent was initially obtained (See Appendix A). The study was a within subjects design consisting of four different conditions, so the participants were divided into four different conditions to counterbalance the two different forms of the HVLT-R, as each participant completed both forms of the test. Form 1 and form 3 of the HVLT-R were used to control for practice effects. A random number generator was used to determine which participants would be assigned to each condition. The order of the presentation of the tests were counterbalanced to control for sequence effects. The audio distraction was only administered during one version of the HVLT-R, in which the audio was played during the total immediate recall, delayed recall, and recognition trials. The order of the audio distraction was

counterbalanced to control for sequence effects. The audio distraction resembled noises that naturally occur in homes, such as dishes clattering, a television playing, people talking, and a dog barking. The audio recording was made by recording live sounds that were occurring in an apartment, and the audio was set at 80db when played for the participants. The study began by administering either the SMQ or whichever version of the HVLT-R that the participant was assigned.

A short delay must occur after the initial administration of the HVLT-R is administered and before the administration of the second version of the HVLT-R. The MHQ and CES-D were administered during this delay as they are quick assessments and are unrelated to the HVLT-R. A 20-to-25-minute delay must occur after each administration of the HVLT-R before the delayed recall can be administered. Tests of executive functioning and the Shipley-2 were administered during the two delays, and the order in which they were administered were counterbalanced. The tests that were included in the executive functioning battery were the Trail Making Test (TMT), Stroop Color-Word Test (SCWT), Ruff Figural Fluency Test (RFFT), and the Controlled Oral Word Association test (COWAT) and these tests were given in a randomized order. The SMQ was administered either at the beginning of the study or at the end to control for any influence that answering questions about one's own memory may have on results of the questionnaire or the HVLT-R.

CHAPTER III: RESULTS

Initial Analyses

An initial analysis was conducted to determine if any of the variables used in this study were correlated. More specifically, the initial analyses were conducted to determine if the Shipley-2 or CES-D were related to any of the dependent variables, which would indicate if they were potential confounds. A series of Pearson correlations indicated that the Shipley-2 was significantly and positively correlated with several of the variables (See Table C2). There was a positive correlation between the Shipley-2 and the HVLT-R total immediate recall without audio, $r(46)=.43$, $p=.002$, 95% CI [.17, .64]. A positive correlation was found between the Shipley-2 and HVLT-R delayed recall without audio, $r(46)=.46$, $p=.001$, 95% CI [.21, .66]. Lastly, there was a positive correlation between Shipley-2 and HVLT-R retention scores without audio, $r(46)=.32$, $p=.027$, 95% CI [.04, .55]. The Shipley-2 was also positively correlated with the HVLT-R total immediate recall with audio, $r(46)=.40$, $p=.005$, 95% CI [.13, .62], and delayed recall scores with audio, $r(46)=.39$, $p=.005$, 95% CI [.13, .61] (See Table C2). Therefore, the Shipley-2 was used as a covariate in the subsequent analyses. The CES-D was not significantly correlated to the variables of interest, and therefore was not used as a covariate.

Primary Analyses

The first hypothesis states that participants will score significantly lower on the HVLT-R when an auditory distractor is present. The goal of the first hypothesis was to determine if memory performance on a standardized assessment is worse when an auditory distraction that resembles real-life home environments is added to the test. The hypothesis was analyzed by using a series of repeated measures ANCOVAs, with the Shipley-2 score entered as a covariate.

The results of the analysis indicated no significant differences between no noise and noise conditions for any of the dependent variables (See Table C3).

The second hypothesis states that the condition with the audio distraction would be more highly correlated with the SMQ scores than the condition with no auditory distraction. The goal of the second hypothesis was to determine if memory performance due to distractions is more highly correlated to subjective memory complaints than memory performance in a standardized environment (with auditory distractions). A Pearson Correlation was used to analyze the relationship between memory performance in the two distraction conditions as it relates to subjective memory complaints. A difference score was calculated to determine the effects of the distraction and a correlation was conducted between the difference scores and the scores on the SMQ. Additionally, correlation analyses were conducted between scores on the HVLT-R without audio distraction and SMQ score totals, and correlation analyses were conducted between scores on the HVLT-R with audio distraction and SMQ score totals. The results indicated that the SMQ and difference scores on the HVLT-R were not significantly correlated (See Table C4). Furthermore, scores on the HVLT-R with and without an audio distraction were not significantly correlated to the SMQ (See Table C2).

CHAPTER IV: DISCUSSION

The purpose of the study was to attempt to create a standardized environment that resembles patients' real-life environments in which there may be auditory distractions and stressors that may inhibit their memory and cognitive abilities. The first hypothesis of the current study stated that adding an auditory distraction to a standardized neuropsychological assessment would significantly reduce memory performance on the Hopkins Verbal Learning Test – Revised (HVLT-R). We predicted that scores on the HVLT-R would significantly decrease by about 2 to 3 points on the audio distraction, based on previous research on patient populations showing a much steeper decline (Randolph et al., 2017; Schnabel & Kydd 2012). Although the participants' immediate recall and retention scores decreased by about 2 to 3 points on average, these changes in performance were not significant. Similar to the current study, previous studies have added distractions to the standardized testing environment to determine if performance on the tests with the distraction would better resemble real-life functioning. Research conducted on individuals with multiple sclerosis, mild cognitive impairment, and mild traumatic brain injury have shown that performance was worse on testing during the distraction conditions, whereas they performed within normal limits on testing without a distraction (Aurtenetxe et al., 2016; Randolph et al., 2017; Schnabel & Kydd 2012). However, results of the current study indicated that, when assessing healthy college-aged individuals, adding an auditory distraction to the test did not impact their performance.

The fact that the reduction in scores was not significant may be attributed to the fact that this study was conducted on healthy college students with higher levels of functional cerebral capacity. Therefore, they should not have performed significantly worse on the assessment with the distraction. Instead, they should have the ability to attend to the distraction while also

performing similarly on the test as when there was not a distraction, which is what the present results found. These findings may relate to patients presenting to the neuropsychological clinic complaining of memory and cognitive problems but then performing within normal limits on testing. One reason why this might occur could be related to functional cerebral capacity, which is the idea that the brain has limited resources, and using too much of its resources can become demanding, and therefore reduce overall performance (Harrison, 2015). Another reason that the change in scores may not have been significant could have been attributed to the fact that the participants in this study may have exhibited a lack of effort. Twelve participants scored within the impaired range and 5 participants scored within the borderline range on the HVLT-R delayed recall without audio. Given that the participants were healthy college-aged individuals, it is highly unlikely that this is an accurate assessment of their memory functioning. Only two of the participants that scored in the borderline or impaired range on the HVLT-R delayed recall without audio had a history of concussion. Therefore, there does not seem to be a relationship between the history of head injury and low scores on the assessment, indicating that the borderline and impaired scores may be due to a lack of effort.

The second hypothesis stated that the audio distraction condition on the HVLT-R would be more highly correlated with the subjective memory scores on the Subjective Memory Questionnaire (SMQ) than the condition without a distraction. Specifically, we predicted that lower scores on the SMQ, indicating that the participant views their memory to be poor, would be related to a steeper decline in scores on the HVLT-R. However, upon statistical analysis, the correlations were not statistically significant. A similar study conducted by Randolph et al. (2017) with MS patients found that patients with more subjective memory complaints performed worse on an objective memory assessment with an audio distraction than those with less

subjective memory complaints. The study also showed that these participants performed similarly on the objective memory assessment without the audio distraction, regardless of their subjective memory complaints. The results of the current study indicated that there was not a relationship between the subjective memory complaints and objective memory scores. The results of the current study were consistent with previous research that also found that scores on a standardized objective memory assessment were not related to how participants subjectively view their own memory (Caramelli & Beato, 2008; Carrasco et al., 2017).

Importantly, in the current study, scores on the SMQ lacked variability, as did the difference scores on the HVLT-R, and therefore could be the reason that the HVLT-R scores were not shown to be significantly related to the SMQ. The SMQ score can range from 1 to 5, with 1 indicating that participants view their memory to be poor, and 5 indicating that participants view their memory to be good. The average score on the SMQ across all participants was 3.40, demonstrating that the participants viewed their memory to be average. As most participants in this study viewed their memory to be average, there was not much variability to determine if there was a relationship between the HVLT-R and the SMQ scores. Similarly, the HVLT-R scores with and without audio did not differ significantly from each other, so these scores also did not have enough variability to determine if a relationship existed between the two assessments.

The results, though, may still have limited clinical utility. When a clinician is evaluating a college-aged individual, they may expect that the patient will only perform about 2 to 3 points worse with the audio distraction. However, if they have a steeper decline in performance, then the patient may be presenting with a problem of functional cerebral capacity. Before this test can be used in the clinic, though, more research should be done with this version of the HVLT-R with the added audio distraction. As previous research has found that adding a distraction to a

standardized assessment more accurately resembles real-life functioning in patient populations, such as patients with multiple sclerosis, mild cognitive impairment, and mild traumatic brain injury (Aurtenetxe et al., 2016; Randolph et al., 2017; Schnabel & Kydd 2012), replicating the current study on these same patient populations may be beneficial to determine if similar results occur, or if their performance demonstrates a steeper, significant decline. Additionally, older adults could be of interest for future research, as older adults may have lower levels of functional cerebral capacity (Cappell et al., 2010), and should therefore demonstrate a steeper rate of decline during the distraction condition. Conducting research with the entire age span could be beneficial to evaluate the level of functional cerebral capacity on each different age group, based on how they may perform on the HVLT-R with the audio distraction.

Other limitations also exist within this study. First, the current study used a sample of healthy college-aged individuals. As a result, their performance on the assessment with the distraction was not significantly different. As this sample does not allow for much diversity in age, future research may benefit from using a larger sample size with a more diverse age range. Second, the participants in this study may have exhibited a lack of effort, as some of the participants scored within the borderline and impaired range on the HVLT-R delayed recall without audio. Thus, future research may benefit from using a more reliable sample of participants to replicate this study. Another existing limitation is the idea that introducing a distractor to a standardized assessment may not be completely standardized, as everyone's everyday environment differs from one another. Therefore, researchers should determine if the audio distraction that was used in this current study truly resembles what the participants experience in their daily lives. Future research may expand on this limitation by asking participants at the end of the study whether they found the audio to be distracting during the

assessment, and if the noises they heard are sounds that resemble noises that they hear in their own homes. Asking participants these questions could be an indicator of whether this distraction is something that may be affecting their performance based on how distracting it was to the participants, and if it resembles noise that they hear in their homes.

The primary purpose of the current study was to determine if adding a distraction to a test of verbal learning would better resemble a real-life environment in which distractions and stressors are present. Therefore, with the distraction present, the test should be more ecologically valid. Previous research is limited in establishing the ecological validity of neuropsychological tests. Ecologically valid tests are important to determine if a patient's performance on the test is indicative of their memory and cognitive functioning in their daily life. Thus, research has been conducted to attempt to create tests that are ecologically valid. However, current studies have not specifically assessed healthy college-aged individuals to determine if a distraction would affect their performance. The current study sought to use an audio distraction, as has been done in previous research by Randolph et al. (2017), but create an audio distraction that more closely resembles the participants' daily life by using sounds that can be heard in homes. Additionally, the current study looked at healthy participants rather than patient populations to determine if similar results could be obtained as with the patient populations.

In conclusion, the current study sought to determine whether adding a distraction into a standardized testing environment would better resemble the distractions and stressors that are present in daily life. Although the results of the study were not statistically significant, the results were consistent with the idea that healthy college-aged individuals would not perform significantly worse on the version of the HVLT-R with the audio distraction because they have higher levels of functional cerebral capacity. The HVLT-R with the addition of the audio

distraction to better resemble real life distractions and stressors has the potential to be an important tool to evaluate ecological validity, but more research must be done using this tool. Patients may present to the clinic complaining of memory and cognitive problems but perform within normal limits on formal memory assessment. The current version of the HVLT-R with the audio distraction, if investigated further, could be used as a more ecologically valid tool to evaluate the possible problems with functional cerebral capacity to determine why the patient may be experiencing memory and cognitive problems in daily life. Future research would benefit from replicating this study with other participant samples, specifically older-aged individuals that may have lower levels of functional cerebral capacity.

REFERENCES

Akbar, N., & Finlayson, M. (2021). A mixed-methods study of cognitive performance in persons with multiple sclerosis: Association between neuropsychological test performance and interviews about daily cognitive functioning. *Multiple Sclerosis and Related Disorders, 52*, 102911. https://doi.org/10.1016/j.msard.2021.102911

Asbee, J., Kelly, K., McMahan, T., & Parsons, T. D. (2022). Factor analysis of the virtual reality stroop task. *Journal of Clinical and Experimental Neuropsychology*, *44*(8), 604–617. https://doi.org/10.1080/13803395.2022.2150749

Aurtenetxe, S., García-Pacios, J., del Río, D., López, M. E., Pineda-Pardo, J. A., Marcos, A., Delgado Losada, M. L., López-Frutos, J. M., & Maestú, F. (2016). Interference impacts working memory in mild cognitive impairment. *Frontiers in Neuroscience, 10*, 443. https://doi.org/10.3389/fnins.2016.00443

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., Erbaugh, J. (2016). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 53–63. https://doi.org/10.1001/archpsyc.1961.01710120031004

Benedict, R., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins verbal earning test – revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist*, *12*(1), 43–55. https://doi.org/10.1076/clin.12.1.43.1726

Benedict, R. H. B., & Zgaljardic, D. J. (1998). Practice effects during repeated administrations of memory tests with and without alternate forms. *Journal of Clinical and Experimental Neuropsychology*, *20*(3), 339–352. https://doi.org/10.1076/jcen.20.3.339.822

Bennett-Levy, J. & Powell, G. E. (1980). The subjective memory questionnaire (SMQ). An investigation into the self-reporting of ‘real-life’ memory skills. *British Journal of Social*

and Clinical Psychology, 19(2), 177-188. https://doi.org/10.1111/j.2044-8260.1980.tb00946.x

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The clinical neuropsychologist, 5*(2), 125-142. https://doi.org/10.1080/13854049108403297

Brandt, J., & Benedict, R. H. B. (2001). *Hopkins verbal learning test – revised*. Odessa, Fla.: PAR.

Cappell, K. A., Gmeindl, L., & Reuter-Lorenz, P. A. (2010). Age differences in prefrontal recruitment during verbal working memory maintenance depend on memory load. *Cortex, 46*(4), 462–473. https://doi.org/10.1016/j.cortex.2009.11.009

Caramelli, P., & Beato, R. G. (2008). Subjective memory complaints and cognitive performance in a sample of healthy elderly. *Dementia & Neuropsychologia*, *2*(1), 42–45. https://doi.org/10.1590/s1980-57642009dn20100009

Carmona, J. E., Holland, A. K., & Harrison, D. W. (2009). Extending the functional cerebral systems theory of emotion to the vestibular modality: A systematic and integrative approach. *Psychological Bulletin, 135*(2), 286–302. https://doi.org/10.1037/a0014825

Carrasco, M. P., Montenegro-Peña, M., López-Higes, R., Estrada, E., Crespo, D. P., Rubio, C. M., & Azorín, D. G. (2017). Subjective memory complaints in healthy older adults: Fewer complaints associated with depression and perceived health, more complaints also associated with lower memory performance. *Archives of Gerontology and Geriatrics, 70*, 28–37. https://doi.org/10.1016/j.archger.2016.12.007

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology, 21*(3), 217–227. https://doi.org/10.1016/j.acn.2005.12.002

Comer, C. S., Harrison, P. K., & Harrison, D. W. (2015). The dynamic opponent relativity model: An integration and extension of capacity theory and existing theoretical perspectives on the neuropsychology of arousal and emotion. *SpringerPlus, 4*(1), 1–21. https://doi.org/10.1186/s40064-015-1120-6

Crean, R. D. (2003). Ecological validity of neuropsychological performance as a predictor of work outcome following mild traumatic brain injury [ProQuest Information & Learning]. In Dissertation Abstracts International: Section B: *The Sciences and Engineering*, *64*(2B), p. 958.

Davidson, P. S. R., Cooper, L., & Taler, V. (2016). Remembering a visit to the psychology lab: Implications of mild cognitive impairment. *Neuropsychologia*, *90*, 243–250. https://doi.org/10.1016/j.neuropsychologia.2016.07.029

Drozdick, L. W., & Cullum, C. M. (2010). Expanding the ecological validity of WAIS-IV and WMS-IV with the Texas functional living scale. *Assessment, 18*(2), 141–155. https://doi.org/10.1177/1073191110382843

Giglioli, I. C., Ripoll, C., Parra, E., & Raya, M. A. (2019). Are 3D virtual environments better than 2D interfaces in serious games performance? An explorative study for the assessment of executive functions. *Applied Neuropsychology: Adult, 28*(2), 148–157. https://doi.org/10.1080/23279095.2019.1607735

Haier, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., Browning, H. L., & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract

reasoning and attention studied with positron emission tomography. *Intelligence, 12*(2), 199–217. https://doi.org/10.1016/0160-2896(88)90016-5

Harrison, D. W. (2015). *Brain asymmetry and neural systems foundations in clinical neuroscience and neuropsychology*. Springer.

Hays, W. L. (1994). *Statistics* (5th ed.). Cengage Learning.

Heinrichs, R. W. (1990). Current and emergent applications of neuropsychological assessment: Problems of validity and utility. *Professional Psychology: Research and Practice, 21*(3), 171–176. https://doi.org/10.1037/0735-7028.21.3.171

Johansson, B. (2021). Distractibility when suffering from mental fatigue after a mild traumatic brain injury. *Physical Medicine and Rehabilitation - International, 8*(3), 1–8. https://doi.org/10.26420/physmedrehabilint.2021.1184

Kibby, M., Schmitter-Edgecombe, M., & Long, C. J. (1998). Ecological validity of neuropsychological tests focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. *Archives of Clinical Neuropsychology, 13*(6), 523–534. https://doi.org/10.1016/s0887-6177(97)00038-3

Kinsbourne, M. (1980). Mapping a behavioral cerebral space. *International Journal of Neuroscience, 11*(1), 45–50. https://doi.org/10.3109/00207458009147578

Knight, R. G., Nicholls, J., & Titov, N. (2008). The effects of old age and distraction on the assessment of prospective memory in a simulated naturalistic environment. *International Psychogeriatrics, 20*(1), 124–134. https://doi.org/10.1017/s1041610207005923

Lacritz, L. H., Cullum, C. M., Weiner, M. F., & Rosenberg, R. N. (2001). Comparison of the hopkins verbal learning test-revised to the California verbal learning test in Alzheimer's

disease. *Applied Neuropsychology*, *8*(3), 180–184. https://doi.org/10.1207/s15324826an0803_8

Leavitt, F., & Katz, R. S. (2006). Distraction as a key determinant of impaired memory in patients with fibromyalgia. *The Journal of rheumatology*, *33*(1), 127-132.

Lynch, W. J. (2008). Everyday Living Assessment in cognitive evaluations. *Journal of Head Trauma Rehabilitation*, *23*(3), 185–188. https://doi.org/10.1097/01.htr.0000319937.97158.09

Makatura, T. J., Lam, C. S., Leahy, B. J., Castillo, M. T., & Kalpakjian, C. Z. (1999). Standardized memory tests and the appraisal of everyday memory. *Brain Injury*, *13*(5), 355–367. https://doi.org/10.1080/026990599121548

Mendes, L., Reis, L. P., & Barbosa, F. (2021). Effectiveness of a holistic neuropsychological rehabilitation program using virtual reality. *Psychtech & Health Journal*, *5*(1), 20–41. https://doi.org/10.26580/pthj.art34-2021

Park, D. C., & McDonough, I. M. (2013). The dynamic aging mind: Revelations from functional neuroimaging research. *Perspectives on Psychological Science*, *8*(1), 62–67. https://doi.org/10.1177/1745691612469034

Pearman, A., Hertzog, C., & Gerstorf, D. (2014). Little evidence for links between memory complaints and memory performance in very old age: Longitudinal analyses from the Berlin Aging Study. *Psychology and Aging*, *29*(4), 828–842. https://doi.org/10.1037/a0037141

Peeters, D. (2019). Virtual reality: A game-changing method for the Language Sciences. *Psychonomic Bulletin & Review*, *26*(3), 894–900. https://doi.org/10.3758/s13423-019-01571-3

Radloff, L. S. (1977). The CES-D scale. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Randolph, J. J., Randolph, J. S., & Wishart, H. A. (2017). Association between cognitive complaints and vulnerability to environmental distraction in multiple sclerosis. *Archives of Clinical Neuropsychology*, *32*(1), 21–28. https://doi.org/10.1093/arclin/acw096

Robertson, K., Schmitter-Edgecombe, M., Weeks, D., & Pimentel, J. (2017). Naturalistic assessment using a simulated environment: Cognitive correlates and relationship to functional status in individuals with neurologic conditions. *Archives of Clinical Neuropsychology*, *33*(8), 1024–1039. https://doi.org/10.1093/arclin/acx136

Rzezak, P., Lima, E. M., Gargaro, A. C., Coimbra, E., de Vincentiis, S., Velasco, T. R., Leite, J. P., Busatto, G. F., & Valente, K. D. (2017). Everyday memory impairment in patients with temporal lobe epilepsy caused by hippocampal sclerosis. *Epilepsy & Behavior*, *69*, 31–36. https://doi.org/10.1016/j.yebeh.2017.01.008

Schaeffer, M. J., Weerawardhena, H., Becker, S., & Callahan, B. L. (2022). Capturing daily-life executive impairments in adults: Does the choice of neuropsychological tests matter? Applied Neuropsychology: Adult. 1-11. https://doi-org.ezproxy.mtsu.edu/10.1080/23279095.2022.2109970

Schnabel, R., & Kydd, R. (2012). Neuropsychological assessment of distractibility in mild traumatic brain injury and depression. *The Clinical Neuropsychologist*, *26*(5), 769–789. https://doi.org/10.1080/13854046.2012.693541

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. *Shipley-2: Manual*. WPS, 2009.

Weakley, A., & Schmitter-Edgecombe, M. (2019). Naturalistic assessment of task interruption in individuals with mild cognitive impairment. *Neuropsychology, 33*(1), 1–12. https://doi.org/10.1037/neu0000481

Weakley, A., Weakley, A. T., & Schmitter-Edgecombe, M. (2019). Compensatory strategy use improves real-world functional performance in community dwelling older adults. *Neuropsychology, 33*(8), 1121–1135. https://doi.org/10.1037/neu0000591

Williams, M. W., Ulrich, N., & Woods, S. P. (2021). Semantic clustering during verbal list learning is associated with employment status in a community sample. *Perceptual and Motor Skills, 128*(3), 1235–1251. https://doi.org/10.1177/0031512521996875

Youn, J. C., Kim, K. W., Lee, D. Y., Jhoo, J. H., Lee, S. B., Park, J. H., ... & Woo, J. I. (2009). Development of the subjective memory complaints questionnaire. *Dementia and geriatric cognitive disorders, 27(*4), 310-317. https://doi.org/10.1159/000205512

Youn, Y. C., Pyun, J.-M., Ryu, N., Baek, M. J., Jang, J.-W., Park, Y. H., Ahn, S.-W., Shin, H.-W., Park, K.-Y., & Kim, S. (2020). Use of the clock drawing test and the Rey–Osterrieth complex figure test-copy with convolutional neural networks to predict cognitive impairment. *Alzheimer's Research & Therapy*, *13*(85), 1-7. https://doi.org/10.21203/rs.3.rs-125100/v1

APPENDICES

APPENDIX A: INFORMED CONSENT

Office of Research Compliance
2269 Middle Tennessee Blvd.
Sam H. Ingram Bldg., Room 010A
Murfreesboro, TN 37132
(615) 898-2400 compliance@mtsu.edu

INFORMED CONSENT

Study Title: Increasing the Ecological Validity of Formal Neuropsychological Assessments
Protocol Number:
Approval Date:
Principal Investigator: Reanna Trexler
Institution: Middle Tennessee State University

Name of participant: _____ Age: _____

You are being asked to participate in a research project. The following information is provided to inform you about the research project and your participation in it. Please read this form carefully. You will be given an opportunity to ask questions, and your questions will be answered. Also, you will be given a copy of this consent form.

Your participation in this research study is voluntary. You are free to withdraw from this study at any time with no penalty and no loss of benefits already earned. In the event new information becomes available that may affect the risks or benefits associated with this research study or your willingness to participate in it, you will be notified so that you can make an informed decision about whether or not to continue your participation.

Please read the following and sign at the bottom if you wish to enroll in this study.

1. *Purpose of the study:* The purpose of this current study is to evaluate performance on a formal neuropsychological assessment when a distraction is introduced. Additionally, the study looks to determine if memory performance with the distraction is more strongly related to how one views his/her own memory performance in daily life.

2. *Description of procedures to be followed and approximate duration of the study:* The study will begin with an initial administration of a memory test that will require you to memorize a list of words. After the first portion of the test, you will be given either a brief assessment of cognitive functioning, or a short battery of neuropsychological assessments. The second portion of the memory assessment will be administered next in which you will recall the original list of words. After this, you will complete a medical history questionnaire and an assessment for depression. Then, you will complete another memory task which will be similar to the first task but with a different set of words. After the first portion of this test, you will complete either the cognitive assessment or the neuropsychological assessment, whichever you did not receive during the first time. You will then do the second portion of the memory assessment. Additionally, you will complete a questionnaire about how you perceive your own memory functioning which may be

Office of Research Compliance
2269 Middle Tennessee Blvd.
Sam H. Ingram Bldg., Room 010A
Murfreesboro, TN 37132
(615) 898-2400 compliance@mtsu.edu

given to you at the beginning of the study or at the very end. Completion of the entire study should take about 90 minutes.

3. *Description of the discomforts, inconveniences, and/or risks that can be reasonably expected as a result of participation in this study:* Participating in this study is not any more risky than normal life. There are no anticipated risks or discomforts anticipated for participating in this study.

4. *Anticipated benefits from this study:*
a) The potential benefits to science and humankind that may result from this study include an assessment that is useful in clinical practice to assess a patient's cognitive abilities in their everyday life.
b) The potential benefits to you from this study include helping increase the knowledge about standardized assessments and different ways that they can be used in the clinic and for research purposes.

5. *Compensation for participation:* You may receive course credit or extra credit for participating in this study. Course credit will be awarded through the SONA system after completing the study. Extra credit will be awarded by the course instructor after completing the study.

6. *What happens if you choose to withdraw from study participation:* Your participation in this research is voluntary. You may choose to stop participating at any time with no consequence. You may receive partial or no credit if you choose to withdraw.

7. *Contact Information:* If you should have any questions about this research study or possible injury, please contact the Principal Investigator: Reanna Trexler
Contact Information: telephone +1 (586)-553-1901; email rt5a@mtmail.mtsu.edu
You may also wish to contact the Faculty Advisor: Dr. Paul Foster
Contact Information: telephone +1 (615)-898-2007; email paul.foster@mtsu.edu
For additional information about giving consent or your rights as a participant in this study, please contact the Middle Tennessee State University (MTSU) Office of Compliance at 615-494-8918 or via email at irb_information@mtsu.edu. (http://www.mtsu.edu/irb)

8. *Confidentiality:* All efforts, within reason, will be made to keep the personal information in your research record private, but total privacy cannot be promised. Your information may be shared with people at MTSU (such as the MTSU Institutional Review Board) or other agencies (such as the Federal Government Office for Human Research Protection) if you or someone else is in danger or if we are required to do so by law. You will be assigned a participant ID number. Your name and contact information will be kept separately from your participant ID number and your ID will be linked to the data that you provide during the study.

Office of Research Compliance
2269 Middle Tennessee Blvd.
Sam H. Ingram Bldg., Room 010A
Murfreesboro, TN 37132
(615) 898-2400 compliance@mtsu.edu

STATEMENT BY PERSON AGREEING TO PARTICIPATE IN THIS STUDY

I have read this informed consent document and the material contained in it has been explained to me. I understand each part of the document, my questions have been answered, and I freely and voluntarily choose to participate in this study.

Date	Signature of participant

Consent obtained by:

Date	Signature
	Printed name and title

Appendix B: MTSU IRB APPROVAL LETTER

Date: 7-23-2024

IRB #: IRB-FY2024-106
Title: Increasing the Ecological Validity of Traditional Neuropsychological Assessments
Creation Date: 11-8-2023
End Date:
Status: Approved
Principal Investigator: Reanna Trexler
Review Board: MTSU Institutional Review Board
Sponsor:

Study History

Submission Type Initial	**Review Type** Expedited	**Decision** Approved
Submission Type Modification	**Review Type** Unassigned	**Decision**

Key Study Contacts

Member Reanna Trexler	**Role** Principal Investigator	**Contact** rt5a@mtmail.mtsu.edu
Member Reanna Trexler	**Role** Primary Contact	**Contact** rt5a@mtmail.mtsu.edu
Member Paul Foster	**Role** Co-Principal Investigator	**Contact** paul.foster@mtsu.edu

APPENDIX C: TABLES

Table 1

Descriptive Statistics for Sample

Variable	Age	
Gender	M	SD
Male (n=12)	23.92	10.06
Female (n=36)	22.25	5.29

Table 2

Correlations Between Variables of Interest

Variable	SMQ	CES-D	Shipley-2
HVLT-R TIR 1	.06 (.676)	-.15 (.326)	.43** (.002)
HVLT-R DR 1	.14 (.342)	-.01 (.962)	. 46** (.001)
HVLT-R Retention 1	.14 (.335)	.09 (.543)	.32* (.027)
HVLT-R RDI 1	.15 (.322)	.01 (.973)	.152 (.303)
HVLT-R TIR 2	.05 (.723)	-.16 (.283)	.40** (.005)
HVLT-R DR 2	.09 (.532)	-.03 (.835)	.40** (.005)
HVLT-R Retention 2	.09 (.540)	.00 (.979)	.21 (.157)
HVLT-R RDI 2	.05 (.728)	.03 (.859)	.22 (.137)

Note. HVLT-R = Hopkins Verbal Learning Test – Revised; TIR = Total Immediate Recall; DR = Delayed Recall; RDI = Recognition Discrimination Index; SMQ = Subjective Memory Questionnaire; CES-D = The Center for Epidemiologic Studies Depression Scale. HVLT-R 1 is without the audio distraction. HVLT-R 2 is with the audio distraction.

*p<.05, **p<.01

Table 3

Descriptive Statistics and ANCOVA Results

Variable	No Audio (n = 48)		Audio (n = 48)		Results
	M	SD	M	SD	
HVLT-R TIR	26.63	4.01	23.79	5.18	F (1, 46) = .323, p = .573
HVLT-R DR	9.04	2.36	8.04	2.66	F (1, 46) = .010, p = .920
HVLT-R Retention	85.60	19.56	82.69	22.25	F (1, 46) = .155, p = .696
HVLT-R RDI	11.04	1.47	10.48	1.35	F(1, 46) = .063, p = .802

Note. HVLT-R = Hopkins Verbal Learning Test – Revised; TIR = Total Immediate Recall; DR = Delayed Recall; RDI = Recognition Discrimination Index. HVLT-R TIR scores range from 0-36, DR scores range from 0-12, Retention scores range from 0-100, and RDI scores range from -12-12.

*p<.05, **p<.01

Table 4

Correlations Between SMQ and HVLT-R Difference Scores

Variable	HVLT-R TIR Difference Score	HVLT-R DR Difference Score	HVLT-R Retention Difference Score	HVLT-R RDI Difference Score
SMQ	-.01 (.485)	.03 (.413)	.03 (.428)	.08 (.303)

Note. HVLT-R = Hopkins Verbal Learning Test – Revised; TIR = Total Immediate Recall; DR = Delayed Recall; RDI = Recognition Discrimination Index; SMQ = Subjective Memory Questionnaire.

*p<.05, **p<.01