A Novel Virtual Reality Executive Function Assessment

Christine Parsons

University of Central Florida
A NOVEL VIRTUAL REALITY EXECUTIVE FUNCTION ASSESSMENT

by

CHRISTINE A. PARSONS
B.S. UNIVERSITY OF CENTRAL FLORIDA 2017

A thesis submitted in partial fulfillment of the requirements
for the degree Master of Sciences
in the Department of Interdisciplinary Studies
in the College of Graduate Studies
at the University of Central Florida
Orlando, Florida

Spring Term
2022
ABSTRACT

Executive Function (EF) is related to cognitive ability and includes such constructs as working memory, inhibition control, and cognitive flexibility (i.e., set shifting). The individual constructs work together to allow a person to set and achieve goals. Student success and achievement has been linked to satisfactory EF skills. Research indicates the testing methods for executive function are diverse and may lack sensitivity. Currently, the NIH Toolbox-Cognitive Battery (NTCB) assessment is a nationally normed standardized battery used to measure individual constructs of executive function in isolation. However, this assessment does not allow measurement of EF as the individual constructs interoperate. Virtual reality (VR) allows the assessment of a combination of individual constructs in real-world settings providing increased ecological validity. This study evaluated the Virtual Reality Grocery Store (VEGS) as an assessment for executive function skills to determine the existing barriers for the use of virtual reality as an assessment for executive function. Although there is a possibility for using VEGS for EF assessment, there is still much work to be done to create an ecologically valid assessment.
ACKNOWLEDGMENTS

To my husband Rick, and my children, Tom, Tim, Alyssa, Stephen, and Sara. I would not be who I am if not for all of you. I cherish the times we spend together and how each of you challenges me every day to be a better wife, mother, friend, and overall human being. Thank you for pushing me to be a better person. I love you all.

To my committee chair, Dr. Matthew Marino, and Dr. Trey Vasquez, thank you for giving me so many amazing opportunities and trusting me to complete your crazy ideas. I could never have imagined the opportunities for education research and personal growth you both have provided me. I am eternally grateful.

To Dr. Taub and Dr. Cruz-Neira thank you for serving on my committee and supporting this work. You are both so inspirational to me and I greatly appreciate your support and guidance.

To Dr. Thomas Parsons, thank you for sharing your work and expertise throughout this process. Your guidance made much of this work possible.
# TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................ vi

INTRODUCTION .................................................................................................................. 1

LITERATURE REVIEW ......................................................................................................... 6
   Executive Function Assessments ....................................................................................... 6
   NIH-Toolbox Assessment ................................................................................................. 7
   Simulations ....................................................................................................................... 8

METHODOLOGY .................................................................................................................. 12
   Participants ..................................................................................................................... 12
   Inclusion and exclusion criteria ...................................................................................... 12
   Materials ......................................................................................................................... 13
   Procedure ......................................................................................................................... 15
      Virtual environment grocery store ............................................................................... 17

RESULTS ............................................................................................................................. 20

DISCUSSION ......................................................................................................................... 23

APPENDIX A CONSENT FORM .......................................................................................... 28

APPENDIX B STUDY SCREENING FORM ......................................................................... 31

APPENDIX C FAMILARIZATION LIST OF ITEMS ............................................................... 35

APPENDIX D POST ASSESSMENT QUESTIONS ................................................................. 37

REFERENCES ....................................................................................................................... 39
LIST OF FIGURES

Figure 1: HTC Vive Set up including head mounted display and hand controllers. .................. 14

Figure 2: An example of the Virtual Environment Grocery Store simulation.......................... 15

Figure 3: Map, shopping list, shopping cart ........................................................................ 16

Figure 4: Shopping Cart with time stamp, money, and basket total........................................ 17

Figure 5: Map of the grocery store including paths of participant. ........................................ 17

Figure 6: Comparison of Maps for participants who finished the VEGS evaluation. .............. 21
INTRODUCTION

A high school student wants to make a birthday cake for their best friend’s birthday. They set a goal to make them a cake and present it in one week. In addition to knowing how to bake, there are other skills required to reach the goal of presenting the birthday cake. Planning is required to determine how long it will take to make the cake and have it ready for the birthday. Are all the ingredients and tools available before baking? While making the cake, the details of the recipe must be followed exactly. Have the eggs been added, the sugar? How long does it need to bake? The ability and skills required to make decisions, plan, and control behavior to reach a goal (make a cake) has often been referred to as executive function skills.

The specific definition of Executive Function (EF) can vary depending on the discipline or researcher. Different researchers have focused on a single EF definition, which includes multiple subsets, while others aim for independent definitions under the umbrella of EF (e.g., Barkley, 2014; Church et al., 2019; Diamond, 2012; Johnston et al., 2019; Kassai, et al., 2019; Miciak et al., 2019; Parsons, et al., 2017). Most agree EF’s main components include such cognitive skills as working memory, inhibition control, self-regulation, and cognitive flexibility (Barkley, 2021; Kassai et al., 2019; Miyake & Friedman, 2012). Others include a variation of the definition including subsets such as information processing, self-monitoring, or self-regulation (Barkley & Murphy, 2011; Boelema et al., 2014; Kassai et al., 2019). Alternative definitions may also include other constructs or other subsets of constructs. According to Rapport, et al., the primary constructs should include working memory, inhibition control and cognitive flexibility, also called set shifting (Rapport, et al., 2013). The constructs of executive function work together to solve problems or attain a specific goal.
An individual’s working memory processes information over time (Diamond, 2012; Miyake & Friedman, 2012) and operates separately from short term memory. This process temporarily holds information while managing and performing cognitive tasks. It also allows an individual to hold onto a set of instructions while processing and implementing those instructions. For example, remembering the ingredients needed at the store to bake a cake or solving a mathematics problem while using the order of operations. Each of these tasks requires working memory essential for goal-oriented tasks.

Inhibition control enables an individual to select where attention should be directed and how to behave (Diamond, 2012). For example, inhibition control allows an individual to carry out a set of instructions or not. Inhibition control works alongside working memory to make decisions to act, or not act, based on the instructions held in working memory. If an individual has difficulty accessing the instructions from working memory, especially multi-step instructions, they will not be able to effectively carry them out. If an individual goes to the grocery store for cake ingredients and purchases a smoothie instead, they will not fulfill their goal of making a cake for their friend’s birthday. If an individual cannot hold onto the rules for the order of operations when solving a math problem, the math may be correct, but the definitive answer for the calculation will be incorrect. Effective inhibition control can enhance one’s ability to sustain the attention needed to complete complex tasks.

Cognitive flexibility, mental flexibility or set shifting, according to Diamond (2012), is the ability to be flexible with thoughts and view different perspectives and adjust outlooks when necessary. If the store is out of eggs, what can be done? Give up on the cake? In the mathematics example, cognitive flexibility allows an individual to be creative with the instructions and persist when challenging instructional barriers are present. Cognitive flexibility allows an individual to
be creative in learning. For example, an individual would use the acronym PEMDAS (i.e., parenthesis, exponents, multiplication, division, addition, then subtraction for the order of operations) to solve a math problem correctly. Cognitive flexibility works in conjunction with each construct of EF to complete tasks, solve problems, and use creativity to be successful.

Researchers agree executive functioning skills are a key element to learning and academic achievement (Sulla, 2018). Academic achievement is strongly related to positive socioeconomic development; therefore, it is imperative to have an understanding and reliable assessments of EF skills in students (Steinmayer, 2015; Sulla, 2018). A decrease in executive function has been associated in children with specific learning disabilities such as autism, attention deficit hyperactivity disorder, and traumatic brain injury (Barkley & Murphy, 2011; Maki et al., 2015; Miciak et al., 2019; Sulla, 2018). Also, public schools in the United States are mandated by the Individuals with Disabilities Education Act (IDEA) to provide support services to students with disabilities, including learning disabilities (2004). Correctly identifying students with learning disabilities has far-reaching implications for students, states, and local schools, including timely interventions for improved learning, complying with the law, and federal funding (Fuchs, et al., 2005; Riggs et al., 2006).

Assessments of EF can be just as convoluted as its definition. Without a standardized definition of executive function, assessments are difficult to standardize and may not provide the necessary information for clinicians make a diagnosis of learning disabilities (Koziol, 2014). Specific tests of the individual constructs under the EF “umbrella” are often utilized and may not provide an accurate representation of EF as a unified construct (Barkley, 2012; 2014, Kamradt, 2013). Tests for executive functions are often used to establish the presence of learning disabilities based on cognitive weaknesses (Maki et al., 2015). According to Kamradt et al.
“only 50% of children with ADHD demonstrate impairment on one or more EF task” (p. 1095). Self-reporting evaluations such as the Barkley Deficits in Executive Functioning Scale have limitations, such as bias from parent or teacher perspectives (Kamrat, 2013). Traditional paper and pencil assessments do not include real-world stresses and therefore may not be as ecologically valid as other assessments such as an immersive virtual reality (VR) assessment (Davison, 2018). Providing tests for EF can be time-consuming, expensive, and focus only on individual EF constructs. Yet, research has shown, with the proper support, deficits in executive function can be improved (Diamond, 2012; Riggs et al. 2006). It would be valuable to have an assessment for executive function that provided information about deficits that occur during real-world activities, including stresses, which are also time-efficient and inexpensive. Virtual Reality provides an opportunity to simulate real-world activities to assess executive function skills.

This evaluation will review the Virtual Environment Grocery Store (VEGS) developed by Parsons et al., (2017) as an assessment for EF. According to Parsons et al. (2017) ecological and construct validity were established in a small sample, through a comparison with the California Verbal Learning Test – Second Edition (CVLT – II), for working memory and the Color-Word Interference Test from the Delis – Kaplan Executive Function System (DKEFS Inhibition/Switching) for assessing inhibitory control (Parsons et al., 2017). Demonstrating construct validity between virtual reality assessments and standardized measures is a step forward in creating a simulated assessment for executive function that includes real-world experiences and stresses. The purpose of this study is to evaluate the barriers affecting the usability and feasibility of a virtual reality assessment for executive function skills. The usability includes the ability for the user to interact with the virtual environment, navigate through the virtual space, and complete assigned tasks. The feasibility includes the ability for the evaluator to
conduct an assessment based on standardized measures and receive informative data from the user experience.
LITERATURE REVIEW

Executive Function Assessments

There is an extensive battery of tests, which measure constructs or subsets of constructs for EF. A systematic literature review was conducted across two library databases over a ten-year period (2011 to 2021). The databases included ERIC((EbscoHost), and Health and Psychosocial Instruments (EBSCOHost). The searches were limited to peer-reviewed, scholarly, English language journals. The search terms included (a) executive function, (b) executive function assessment, and (c) virtual reality executive function assessment. Inclusion criteria consisted of articles from the United States with a sample population age range of 5 to 21 years old and included studies of executive function assessments. Qualitative research studies, intervention studies, reports, reviews, and literature reviews were not included. Journal titles and abstracts were reviewed to determine if the studies met the age and language criteria and evaluated executive function assessments or virtual reality assessments for executive function.

A variety of tests used in research and clinical settings designed to evaluate working memory include the California Verbal Learning Test – Second Edition (CVLT-II) Learning Test (Thiruvelvam & Hoelzle, 2020), Memory Search (Boelema et al., 2014), the Automated Working Memory Assessment (AWMA) battery (Packiam, et al., 2009). Assessments usually challenge the participant to recall a list of viewed elements or words.

In the explanation of inhibitory control by Diamond (2012), there was a list of seven different assessments for inhibition, which she stated, “Not everyone agrees that these tasks require inhibitor control” (p. 139). Assessments of inhibition control include the Stroop color naming task (Stroop, 1935), or variations of this test such as the Color-Word Interference Test from the Delis-Kaplan Executive Function System (D-KEFS) Inhibition/ Switching (Crawford, et
The Color-Word Interference Test presents, for example, the word RED in green letters. The participant is told to name the word RED or green. Different combinations are presented in a set time.

Other tools used by neurologists include Wechsler Adult Intelligence Scale (WAIS-III) digit Span, Arithmetic and Letter/Number Sequencing (Biederman et al., 2011), and Behavior Rating Inventory of Executive Function (BRIEF) parents and teacher report (Krieger & Amador-Campos, 2018). Executive Functioning rating scales described by Barkley have demonstrated to be a more effective tool than traditional single construct assessments for determining executive dysfunction (Barkley, 2014; Krieger & Amador-Campos, 2018).

Assessments for cognitive flexibility include the Wisconsin Card Sorting Task and the Shape Trail Test (Chan et al., 2018; Hommel, et al., 2022). These tests require participants to sort cards with a set of rules, which change as the evaluation progresses. The ability to switch with the changes determines cognitive flexibility.

**NIH-Toolbox Assessment**

The National Institutes of Health Toolbox was created to fulfil the need for standardized methods of assessments for executive function, and provide valid, easy to use, quick, computerized assessments (Gershon et al., 2010). The NIH Toolbox-Cognition Battery (NTCB) website provides a list of the individual measures and summary scores for cognition, including working memory, flexibility, inhibitory control, and other constructs such as processing speed, reading, and language skills ([https://www.healthmeasures.net/explore-measurement-systems/nih-toolbox](https://www.healthmeasures.net/explore-measurement-systems/nih-toolbox)). These measurements can be completed in thirty minutes with minimal training, and on a computer tablet with minimal materials. However, there are still problems with testing constructs in isolation and not assessing real-world assessments for EF (Kamradt, 2013).
Although the NIH-Toolbox employs technology by using computer tablets, it is still based on traditional paper and pencil assessments.

McCoy (2019) explains the need for ecological validity for EF assessments, which she describes as “a measure’s ability to capture processes that are relevant to real-world behaviors and outcomes” (p. 67). Although the NTCB provides a standardized test for individual constructs of EF and total EF score, it does not provide real-world stresses associated with increased ecological validity (Davidson 2018). Performance on paper and pencil tests may not serve as a representation of total executive function skills (Parsons & McMahan, 2017). Although they are used extensively, these assessments can be expensive and time consuming (Parsons & Barnett, 2017). Some measures cannot examine developmental changes across age groups (Church et al., 2019; McCoy, 2019; Miciak et al., 2019; Zelazo et al., 2013) and others lack real-world associations to executive function skills (Kamradt, 2014). Individual constructs of EF work together to solve problems or reach a goal, like solving a math equation or remembering required tasks for a project and focusing to complete them. Evaluating each component in isolation may not provide a comprehensive or accurate picture of EF functioning in a real-world scenario.

Simulations

The Link trainer was a flight simulator and is considered the first commercially available simulator, which made its appearance during the early days of World War II (Jeon, 2015). Today, the benefits of using flight simulators for learning, practicing, and evaluating flying instruction are unparalleled. While the use of simulated flight trainers provides a variety of affordances such as safety and the ability to learn from mistakes during a variety of real-world scenarios, simulations can also apply to other experiences or skills such as manufacturing, military exercises, and medicine (Lateef, 2010). Simulations can model a diverse array of real-
world scenarios in a variety of disciplines including mechanics, construction, heavy machinery training, military exercises, and education. Virtual reality can simulate cognitively difficult or complex tasks in a safe, low-cost environment, while providing data about EF skills, such as time to complete tasks.

Virtual reality (VR) can be defined as a synthetic medium to showcase a simulation driven by rules and behavior (Sherman, 2018). A virtual reality simulation can mimic aspects of the real world that users can interact with. The environment can be viewed in 3D or 2D modalities. A 3D immersive experience may use a head mounted display to provide a full, 360° view and interactivities of the virtual world. According to Sherman, (2018) when developing virtual reality simulations, there are four key elements encompassing the experience: 1) Participants, 2) Creators, 3) Virtual World interface, 4) and Interactivity (2019).

1) Participants use their senses to experience the VR environment. The creator of a VR world must consider the participant including the interactions with elements of the game or virtual experience and the “space” the user is occupying within the game space (Sherman, 2018). According to Sherman, (2018), the more realistic the experience for the participant, the more authentic the experience becomes.

2) Individual or a team of creators are a key element in the game because they provide the elements the participant interacts with. The creator is the storyteller of how the game should proceed and designs the sequence of events.

3) The virtual world interface is the access point the participant uses to access the game (Sherman, 2018).

4) The interactivity of the game provides the authenticity of the environment.
Creators must understand the unique collaboration among the participants, the interface, and the interactivity to provide an authentic experience (Sherman, 2018). Without a seamless interaction between the participant and the interface, the experience will be diluted and not represent authentic behaviors from the participant reactions.

Virtual reality may provide the real-world environment missing from executive function assessments. Lalonde et al. (2013) found VR was a better predictor of EF performance in everyday behaviors than traditional survey methods. Recently, Parsons et al., (2017) described the potential for the use of virtual reality as an ecologically valid assessment for EF based on a real-world scenario. Assessments for EF using VR can be accomplished using 2D (non-immersive) or 3D (immersive) simulations with data collection occurring in real-time. These assessments have the potential to measure individual constructs while also including real-world scenarios missing in current measurements (Parsons & Barnett 2017). Simulations using a head mounted device, such as the HTC Vive or Oculus Rift, allows the participant to become immersed in a real-world scenario. Their movements and actions are recorded by the VR software, data can be collected and later analyzed. The Virtual Environment Grocery store (VEGS) (2012) was designed to allow such assessments of EF.

The VEGs simulation was built using a Unity platform as the virtual world interface. Unity provides a game engine development platform, which allows users to “build, manage and grow” the game. It includes coding using C#, graphics, and drop functionality (Unity game engineguide: How to get started with the most popular game engine out there, 2018). The use of Unity provides a game engine for experienced programmers to create a VR world in a variety of platforms such as mobile, 2D, or 3D VR. Unity provides the creator with a variety of texture compressions and resolution settings and provides supports such as hotlines, user groups,
consulting services and technical support. Analytics are also provided, which can support the use
of VR assessments with required data collection for EF measurement. The development of
simulations today is straightforward and can provide real time information about performance.
The questions become what current issues need to be overcome to move forward.
METHODOLOGY

This study was approved by the university’s Institutional Review Board. The purpose of this study was to discover any barriers present for the use of virtual reality as an assessment for executive function. The hypothesis is there are barriers that will need to be overcome before virtual reality can become an ecologically valid measure of executive function with an environment that a user can interact and complete tasks, and an evaluator can collect pertinent data.

Participants

The population under investigation consists of undergraduate students in a large university in the southeastern United States. The total population of undergraduate students was 59,483 for the 2019-2020 school year. The study included five healthy undergraduate students ($n = 5$): one senior, two juniors, and two freshmen. Recruitment of students was through social media, distributed flyers, and email. Limited face-to-face experiences during the study resulted in a small sample size. Participants were provided with a brief overview of the study. Once the participants read an explanation of their participation requirements, they signed a consent form, which served as their agreement to participate (Appendix A). The group included three males and two females. Participants answered questions concerning basic demographics, computer usage (e.g., How many hours per week do you spend on the computer?), and their perceived level of computer skill (1 - not at all to 5 – very skilled).

Inclusion and exclusion criteria

Students were screened for history of visual impairment not corrected by glasses to ensure visual acuity in the virtual world (Appendix B). Also, those with epilepsy or history of
seizures, and psychiatric disorders were excluded since such conditions may alter executive function. Participants were not screened for motion sickness.

Materials

The HTC Vive Pro 2 consists of an adjustable head mounted video display with two 1440 x 1600 pixels per eye high resolution AMOLED displays, hi-res certified headphones, and 90 Hz refresh rate, with 110° field of view with stereoscopic vision. Also included in the system are two wireless handheld controllers with multifunction trackpad, grip buttons, dual-stage trigger, system button, and menu button. The hand-held devices are used to interact with objects and are visible in the virtual environment. In the game participants could view the name and price of items by pointing the handheld device to the rows of items in the store and click on the desired item they wanted to add to their cart.

Two base stations mounted opposite from each other in the corners of the testing room can support 16’5” x 16’5” space. Optics in the base station allow the optics in the headset to track position in the play area. As participants move their head, the environment changed to allow complete visualization. The head mounted device is wired to an Alienware M15 R3 laptop computer, which allows the researcher to view the environment and the participants’ movements in the virtual world. Figure 1 shows the HTC Vive Pro 2 set up including the headset device and controllers.
Figure 1: HTC Vive Set up including head mounted display and hand controllers.

The VEGS software was generously provided to the researcher by the developers. The interface allows users to immerse themselves in the computer-generated environment. The environment was simulated by taking pictures of an actual store near the developer labs. Sounds played during low and high distractions in the simulated environment were obtained by recording the sounds in the same store as the obtained pictures. The design of the virtual grocery store experience was based on the multiple-errands test (MET) (Law et al., 2006) and consists of a shopping exercise. Participants enter the VEGS and are given the task of going to the pharmacy to drop off a prescription and listen for their number to be called, buy items from a provided list, keep track of how much money they spend, and get a coupon after 5 minutes. Each individual task is related to different EF constructs and require the use of all constructs working together to accomplish the tasks successfully. The participant needs to remember their prescription number, keep track of time, stay focused on what items to purchase, and manage how much money they must spend. An example of the environment can be found in Figure 2.
Figure 2: An example of the Virtual Environment Grocery Store simulation

Procedure

Sessions were conducted over one 45-minute period. The study room was free of distracting noises and interruptions. All participants were given scripted instructions according to the VEGS user manual (Parsons, 2012). The instructions also included explanations of the handheld controllers; the function of the track pads on each controller; and the location and function of the trigger. The participants were seated during the simulation to reduce cyber sickness. They had full 360° mobility. The VEGS can be set to no distractions, low distractions, and high distractions. Distractions in the VEGS include background noise, the number of non-player characters (avatars) walking around, and the number of overhead announcements. In this study all participants had low distractions to assess the use of distraction elements within the VEGS. Participants were instructed to use two controllers to navigate the VEGS, movement was always controlled with their non-dominant hand and the map, shopping list and cart were controlled with their dominant hand for consistency. The participants were allowed to practice navigating the virtual environment by using the track pad on the handheld controller in their non-dominant hand to move forward, backward, and turning. For consistency, all participants used their non-dominant hand to move within the virtual environment. Before entering the virtual
world, the researcher read a list of sixteen items on the shopping list and participants were instructed to recall as many items as possible in any order (Appendix C), (Parsons, 2012). This activity was a learning task to assess recall, become familiar with the list of items to purchase within the activity, and was repeated for five trials. The participant’s recall was recorded.

Instructions were given for navigating the virtual environment. Clicking on the top button of the trackpad on the controller in their dominant hand the participant could view the shopping list including all 16 items (Figure 3).

**Figure 3: Map, shopping list, shopping cart**

Pressing the right button on the trackpad allowed participants to view the shopping cart, items added to the cart, time, amount of money to spend, and the total cost of each item including the total spent (Figure 4).
Figure 4: Shopping Cart with time stamp, money, and basket total.
Pressing the left trackpad will display the map of the store and the path taken by the participant and includes distinct colors based on the activity within the session. (Figure 5).

Figure 5: Map of the grocery store including paths of participant.
Clicking the pharmacist with the trigger begins the timer and data collection within the VEGS.

Once the participants felt comfortable moving in the environment the researcher asked if they had any questions. Then, they were instructed to go to the back of the store to “drop off” a prescription. They were also reminded to go to the coupon machine 5 minutes after dropping off their prescription. They also needed to go to the ATM machine if they needed money. If after 5 minutes the participant did not get coupons, the participants were prompted to check the time on the clock-based timer in the shopping cart window.

Virtual environment grocery store

The participant placed the HTC Vive head mounted device on their head and adjusted the device for comfortable fit. The trigger is located on the underside of the hand-held device and is used to activate the data collection by clicking on the pharmacist. Once they ‘dropped off’ their
prescription and the virtual pharmacist gave them a prescription number, they could move about the store and collect items from the shopping list.

According to Parsons (2017) the VEGS measures episodic and prospective memory. Several data points are collected during the simulation. The VEGS List Items Score = the number of items in their cart. Intrusion Score = number of items in their cart not on the original shopping list. During the time-based task, the participant must remember to visit the coupon machine after 5 minutes which is a prospective memory measure. From dropping off their prescription at the pharmacy. The Time-based score = go to the coupon machine after 5 minutes and was coded 1 if they remembered and 2 if they needed prompting. Only the time from dropping off the prescription and going to the coupon machine was evaluated based on Parsons previous work (Parsons & Barnett, 2017), although the program does collect time-based data throughout the simulation and may provide more information.

Once the pharmacist is clicked the participant begins shopping for items on their shopping list. The event-based task includes remembering their prescription number and picking up their prescription when the number is called. The pharmacist announced their number after 10 minutes. Participants were not told the amount of time it would take for their number to be called. The event based prospective memory was calculated as amount of time it took for participants to return to the pharmacist. This measure however could be affected by where the participant was in the simulation. Although the simulation should move consistently for each participant, it should be noted the speed of movement for each participant varied slightly. This issue must be standardized within the simulation program to allow valid measurements.

After the headsets was removed, the participants were asked to recall aspects of the experience (e.g., “describe the people you saw, the elements of the environment, what you did in
the store” (Appendix D). After each participant left the room, the head mounted device and hand-held controllers were wiped down along the straps, faceplate, and the outside of the devices with disinfectant cloths.
RESULTS

This goal of this study was to discover barriers to using the Virtual Environment Grocery Store (2012) as an ecologically valid assessment of executive function. Undergraduate students participated in the virtual reality assessment to measure prospective and episodic memory. This study used prospective and episodic memory as a measure of executive function based on Parsons & McMahon (2017) evaluation of VEGS. Participants included three male and two female undergraduate students from a large southeastern university. All participants reported comfort with computers and related their competency as very skilled. All had used a virtual reality headset at least once previously. For each participant, distractions were set to low. Each participant noticed the background noise. Participant 3 reported it was too loud and not representative of an actual shopping experience. Of the five participants two completed the entire scenario from dropping off the prescription to picking it up when their number was called. The VEGS program navigation was not working correctly for Participant 2. The navigation would only work travelling backwards. The game was stopped and restarted twice. After 45 minutes of troubleshooting the participant could not continue. Troubleshooting included restarting the computer and the game. Resetting did fix the navigation for the next participant.

Two participants needed to stop the assessment due to cyber sickness. Cyber sickness is motion sickness in a virtual world (Martirosov & Kopecek, 2018). Both blamed the slow movements through the environment with causing the sickness. They both also stated they rarely had issues previously in a virtual environment. One participant used VR games regularly and one had only tried them once.

Only Participant 1 and 5 provided data to assess EF. Participant 1 did not remember to check their starting time but did remember to go to the coupon machine. Participant 1 estimated
the time to be about 5 minutes and the actual time was 4 minutes. Participant 5 did complete the coupon acquisition after five minutes but could not remember the pharmacy number. They did go to the pharmacy after about 10 minutes to pick up the prescription, which did stop the simulation at the announcement of the correct number being called. Although each participant had lapses with the task, they both were able to complete the tasks. Their EF skills worked together to complete the goal. Future research should focus on not only the individual constructs, but the ability to utilize those constructs together to complete the task at hand. As explained previously, constructs work together, but without a clear definition of what executive function skills are, it will be difficult to measure those skills.

The maps for each participant were also evaluated. There are three paths represented by the colors red, green, and yellow. Red represents the path after clicking on the pharmacist to begin the session. Green represents the path after hearing their number called. Yellow represents the path from starting to going to the pharmacist. Each participant was able to view the map and were also shown where essential game components (e.g., the coupon machine, the ATM machine, and the pharmacist) were located. Figure 6 shows a comparison of the maps for the two participants that were able to complete the evaluation.

*Figure 6: Comparison of Maps for participants who finished the VEGS evaluation.*
The map on the right shows the methodical movements of the participant while shopping. It is challenging to see, but the participant was near the pharmacist when the number was called. Therefore, limited green color representing movement through the store after the prescription number was called is present. If participants were in the front of the store, there would be more time required to navigate to the back of the store. This small detail will need to be addressed if using time to pick up prescription as a measurement of the ability of the participant to complete the task successfully. This participant also neglected to go to the ATM machine or the coupon machine in the required time. The participant was focused on shopping for items and listening for the prescription number. The map on the left shows the movements of the second participant and their progress through the simulation. This participant did go to the coupon machine and the ATM but had difficulty interacting with the user interface in the simulation, which took time away from collecting items in the shopping cart. Although there is not a standard measure for reviewing the map after participation, there may be value in analyzing the paths of participants during each phase of the assessment.

All participants experienced a variety of problems interacting with the simulation. The frame rate was often unsteady which may have been the major cause of cybersickness. The interaction between selecting an item and having the item show up in the cart was delayed which caused the participant to add them more than one. Participants would need to remove the item to stay within budget. This caused a delay in the time they had to collect items. These types of delays will need to be considered when using time as a measure within the simulation. If the measurements for executive function include evaluating the items in the cart, the simulation will need to be seamless in selecting and adding items to the cart.
DISCUSSION

This study aimed to evaluate the Virtual Environment Grocery Store (2012) to determine usability and feasibility barriers present during the simulation. Due to the small sample size and participant cyber sickness, the evaluation can only describe the use of VEGS across 2 participants, with discussion of possible uses for further research. Larger sample sizes are required to establish diagnostic utility of the VEGS assessment in the future. Compared to standardized assessments for working memory such as the NIH Toolbox, the VEGS took longer to complete the assessment (VEGS: 45 minutes, NIH Toolbox: 10 minutes). The actual assessment time for the VEGS was ten minutes but set up and practice time added to the length of the total evaluation. There is also a paper component to the VEGS that added time to the evaluation. The paper and pencil component scores can be added to the data collected in the simulation and may provide added information for executive functioning. What those measures mean will require more extensive research. Although the overall time was longer, it must be noted every added construct to the NIH Toolbox adds time to the overall evaluation. It is possible the VEGS will be able to measure added constructs within the current ten-minute time frame.

Although the allure to measure EF skills in real-life scenarios is intriguing as an ecologically valid measure, current measures only include individual constructs. More research is required to understand the outcomes of evaluating EF skills in virtual reality as an ecologically valid measure (Kamradt, 2013). Participant 5 struggled remembering their prescription number but was still able to pick up the prescription in 10 minutes. The data would appear to demonstrate success of the task in the correct amount of time, but the dialog from the participant indicated a lack of success. During the evaluation, the researcher was able to make observations about the
participant progress that was not recorded by the simulation. The post assessment questions found in Appendix D could provide additional information for the evaluator to assess the total executive functioning skills not observed during the simulation. Future research will be required to evaluate the data collected from the simulation and observations of participants by the evaluator. It is possible this participant would need further evaluation for working memory. If this participant has a deficit in one construct of EF, working memory, yet they can complete the task in the real-world scenario, how does this affect a measure of EF skills? This may indicate adaptations by the individual to overcome deficits. If this tool is to be used as an assessment there will need to be processes in place to limit an individual’s ability to learn the parameters of successfully completing the measured tasks. Adaptations of individuals who struggle with executive function constructs is an area for future research to determine how this could be measured. Future research will need to include the possibility of normal EF skills with deficits in individual constructs and determine what that means for individuals’ ability to successfully meet goals.

Future assessments will require a standardized set of graphic and programming protocols for assessment distribution and accession to create authentic experiences. There is great disparity in the quality of the simulations for 2D and 3D experiences. If the 3D experience increases the authenticity of the experience, it may or may not increase the information provided to evaluators. However, if the 3D experience is fraught with frame rate inconsistencies or interactive problems a 2D simulation may be effective. Future research should include the use of 2D versus 3D virtual experiences to determine what is gained by the 3D experience. If there is more information to be gained from a 3D experience, the quality of that experience will need to be standardized.
Cyber sickness can be a problem for many when experiencing VR. As in motion sickness, cyber sickness is due to mixed signals sent to the brain from a person’s senses (Martirosov & Kopecek, 2018). The key aspects affecting cyber sickness are frame rate, textures, and scale. Each of these components send signals to the brain about the environment. In virtual reality a person can see movement but not feel the movement in their muscles, sending mixed signals to the brain. When the signals are not aligned, the brain can be confused and send sickness messages. Some areas of research have looked at frame rates, field of view, time in virtual environments, and even possibly sleep deprivation on the effects of VR and cybersickness (Martirosov, 2018). A cyber sickness self-reporting questionnaire (SSQ) developed by Kennedy et al., (1993) is still often used today and could be included in future studies to determine the possibility of participant sickness. Reducing or eliminating cyber sickness is imperative to the future of virtual reality as EF assessments.

The researchers setting up the VEGS before participant trials experienced various levels of cyber sickness. To eliminate issues the ratio aspect was adjusted. This did improve the cyber sickness for some, but not all. Some cyber sickness was reduced by sitting in chairs, therefore all participants sat in a swivel chair. Other means of eliminating cyber sickness include improving the frame rate or using teleportation as a means of movement (Choros & Nippe, 2019; Kemeny et al, 2020). Frame rate is the number of images displayed per second. Reducing the level of detail could increase the quality of the frame rate for the VEGS and may decrease the levels of cyber sickness (Choros & Nippe, 2019). The consistency of the frame rate needs to be steady throughout the simulation to eliminate cybersickness. Teleportation would allow the user to move through the environment without the slow movements of the VEGS, which should eliminate cyber sickness, but this may reduce the authenticity of the experience (Kemeny et al.,
Teleportation allows the user to point to a place they want to go and pressing the trigger to automatically move to the selected area. Since the movement is instantaneous, the visual movement cues to the brain are reduced, however this may also reduce the sense of immersion in the virtual environment.

Probable future uses of VR environments can be to support students to improve areas of EF deficits. Simulated environments have demonstrated the ability to allow students to rehearse using their EF skills and transfer those new skills to outside the simulation (Donehower, et al. 2020). More research is required to assess the outcomes of virtual reality in improving EF skills. The VEGS may provide an opportunity for students to practice going to the store, managing time and money, and learning to plan what to do and when in the store. Other environments could be added such as going to the mall or sitting in a large classroom.

A current limitation of the Virtual Environment Grocery Store (2012) includes the use of head mounted devices on young children. The use of VR is not recommended for younger children under the age of 12 (Zelazo et. al., 2014). Because of this the use of VR in schools for younger students will be limited to students over 12 years old if using a head mounted device. Technological advancements will be required to evaluate the use of virtual reality in individuals under 12. Computer screens using 2D simulation could replace the virtual environment for younger children, but that eliminates the immersive real-world experience. Much work is still needed to evaluate the use of virtual reality in executive function assessments in younger children and its use in schools as a diagnostic tool for executive function deficits.

The implications for the use of real-world scenarios to evaluate executive function skills are considerable. Considering current measures include individual constructs of executive function, it may be that a new measurement of total executive function skills with a standardized
definition and ecologically valid results are yet to be identified. Investigations comparing VR results to current standards may not measure the same functions. Other possibilities of virtual reality include remote assessments using the metaverse. Someday it may be possible for students to meet with evaluators in a virtual examination space. Families will have the ability to meet with examiners in their home or school using a virtual simulation meeting space. Examiners can interact with the examinees and place different stresses or distractions within the space based on the individuals’ needs and concerns. Within a brief period, the researcher can observe the interaction of the examinee with the virtual environment and determine the executive function benchmarks for the appropriate age group. There continue to be barriers to reaching such an evaluation. Researchers continue to make progress in the area of executive function assessments, however there is still much work to be accomplished.
APPENDIX A
CONSENT FORM
EXPLANATION OF RESEARCH

Title of Project: A Comparison of the NIH Toolbox-CB Assessment for Executive Function with a Novel Virtual Reality Executive Function Assessment

Principal Investigator:

Christine Parsons

Faculty Supervisor:

Dr. Matthew Marino

You are being invited to take part in a research study. Whether you take part is up to you. The purpose of this research is to compare a standard assessment for executive function, the NIH Toolbox-CB, with a new assessment based on a virtual reality game in a grocery store (VEGS). If you participate in this study, you will be asked to take both assessments. It will take approximately 60 minutes to complete both assessments. The NIH Toolbox is a computer-based assessment that takes approximately 30 minutes to complete. The VEGS is a virtual reality assessment and takes about 30 minutes to complete. In each assessment you will complete a number of tasks that involve memorizing and organizing. You will be provided with instructions on how to complete the tasks. You will be asked questions before you begin any tasks. You will also have an opportunity to ask questions before each assessment to ensure your understanding. The assessments will take place in the Teaching Academy in room 103-106 on the UCF campus.
Your participation in this study is voluntary. You are free to withdraw your consent and discontinue participation in this study at any time without prejudice or penalty. Your decision to participate or not participate in this study will in no way affect your relationship with UCF, including continued enrollment, grades, employment, or your relationship with the individuals who may have an interest in this study.

To be eligible to participate in this study you must be an undergraduate student over 18 years of age. You may not participate if you have a visual impairment not corrected by glasses or contacts, have epilepsy or a history of seizures, or have been diagnosed with a psychiatric disorder.

**Study contact for questions about the study or to report a problem:** If you have questions, concerns, or complaints contact Christine Parsons, Graduate Student Interdisciplinary Studies Program, College of Interdisciplinary Studies, 407-823-6705, or by email christine.parsons@ucf.edu; or Dr. Matthew Marino, Faculty Supervisor, College of Community Education and Innovation, 407-509-592-3760, or by email matthew.marino@ucf.edu.

**IRB contact about your rights in this study or to report a complaint:** If you have questions about your rights as a research participant, or have concerns about the conduct of this study, please contact Institutional Review Board (IRB), University of Central Florida, Office of Research, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901, or email irb@ucf.edu.
APPENDIX B
STUDY SCREENING FORM
Thank you for your interest in participating in the Executive Function Assessment research study. To determine your eligibility for this study please answer the following questions.

Are you currently and undergraduate student?

Yes
No

What is your age
18 - 25
25 - 40
41 or over
Under age 18

Do you have a visual impairment (that is not corrected by contacts or glasses)?

Yes
No

Do you have epilepsy or a history of seizures?
Yes
No

Have you ever been diagnosed with a psychiatric disorder?

Yes
No
No

What is your major?

Thank you for completing the Executive Function Assessment research study survey. You meet the requirements for participation in this research study. Use this link to schedule a time to participate. Please arrive 5-10 minutes before your appointment time. The assessments should take approximately 60 minutes to complete.

Thank you for completing the Executive Function Assessment research study survey. You do not meet the requirements to participate in this research study. If you would like more information, please contact Christine Parsons at christine.parsons@ucf.edu. Again, thank you for completing the survey.
APPENDIX C
FAMILARIZATION LIST OF ITEMS
APPENDIX A

Record all responses verbatim, in the order recalled. Prompt only once (e.g., Anything else?) at the end of each free and cued recall trial (i.e., after 15 seconds with no response or when the examinee says he/she cannot remember more words).

- Participant was reminded to pick up prescription at the end: Y  N
- Participant was reminded to go to the coupon machine: Y  N

<table>
<thead>
<tr>
<th>LIST Item</th>
<th>Immediate Free Recall</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
<th>Trial 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any Nutriene</td>
<td>Let's suppose that you</td>
<td>Okay,</td>
<td>Okay,</td>
<td>Okay,</td>
<td>Okay,</td>
</tr>
<tr>
<td>coloring</td>
<td>needed to go shopping</td>
<td>now I</td>
<td>now I</td>
<td>now I</td>
<td>now I</td>
</tr>
<tr>
<td>Men's Health</td>
<td>for some everyday</td>
<td>am</td>
<td>am</td>
<td>am</td>
<td>am</td>
</tr>
<tr>
<td>Magazine &quot;Build new</td>
<td>items, I'm going to</td>
<td>going</td>
<td>going</td>
<td>going</td>
<td>going</td>
</tr>
<tr>
<td>Muscle Issue</td>
<td>read a list of items</td>
<td>to you</td>
<td>to you</td>
<td>to you</td>
<td>to you</td>
</tr>
<tr>
<td>M&amp;M's from mark</td>
<td>for you to buy.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>down bin</td>
<td>Please listen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glad Press and</td>
<td>carefully as I read</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seal</td>
<td>them. When I am</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batteries</td>
<td>through, I want you</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin D</td>
<td>to say back as many</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Robin bran</td>
<td>as many items as you</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fritos Chips</td>
<td>can, in any</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Popcorn</td>
<td>order. Be sure to say</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bounty paper</td>
<td>the items on the list</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>towels</td>
<td>that you told me the</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light Bulbs</td>
<td>first time.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tylenol cold</td>
<td>Read the List at an</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Vite Shampoo</td>
<td>even pace, taking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Listerine</td>
<td>slightly longer than</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Ives Soup</td>
<td>one second per word,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band Aid brand</td>
<td>so the entire list</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>band aids</td>
<td>takes 15 to 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>seconds. Then say:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Go ahead.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10 
11 
12 
13 
14 
15 
16 
17 
18 
19 

Total Correct:  
Total Repetitions:  
Total Intrusions:  

14
APPENDIX D
POST ASSESSMENT QUESTIONS
APPENDIX D: Long Delay Description of VEGS:

Please describe (in detail) everything that you remember experiencing in the virtual environment. This should be done in the order that you experienced them in the virtual environment. Here are a few things to keep in mind when you describe all the elements you saw in the grocery store.

**ACTIONS:** Try to describe where all you went in the environment and describe what you did. (e.g., When I went to drop off my prescription, I turned...)

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

WHO, WHAT, WHEN, WHERE:

“Try to remember all the people and elements you saw in the grocery store” (e.g., a boy, the Minute Clinic, the Pharmacy, etc). Be sure to give as many perceptual details about the elements as possible (e.g., the boy was wearing a red shirt and blue shorts). Also, please situate the elements in time: were they at the beginning, the middle, or at the end of the grocery store. Finally, when you are describing people and elements, please situate the people and elements of the scene where they were in relation to each other” (e.g., the boy was to the left of the entrance)

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________
REFERENCES


Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18,


Unity game engine guide: How to get started with the most popular game engine out there. (2018, 2 13). Retrieved 2022, from freeCodeCamp: www.freecodecamp.org
