

5-2019

## Validity and Reliability of a 5-minute Web-Camera Based Eye Tracking Test to Assess Visual Memory and Cognition

Emily Bates  
*University of Arkansas, Fayetteville*

Follow this and additional works at: <https://scholarworks.uark.edu/etd>



Part of the [Biomechanics Commons](#), [Cognitive Neuroscience Commons](#), and the [Motor Control Commons](#)

---

### Citation

Bates, E. (2019). Validity and Reliability of a 5-minute Web-Camera Based Eye Tracking Test to Assess Visual Memory and Cognition. *Graduate Theses and Dissertations* Retrieved from <https://scholarworks.uark.edu/etd/3207>

This Thesis is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact [scholar@uark.edu](mailto:scholar@uark.edu).

Validity and Reliability of a 5-minute Web-Camera Based Eye Tracking  
Test to Assess Visual Memory and Cognition

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Science in Kinesiology

by

Emily Bates  
University of Arkansas  
Bachelor of Science in Kinesiology, 2017

May 2019  
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

---

Michelle Gray, Ph.D.  
Thesis Director

---

Nicholas Greene, Ph.D.  
Committee Member

---

Tyrone Washington, Ph.D.  
Committee Member

## Abstract

There are approximately 5.7 million Americans currently living with Alzheimer's Disease (AD). Early detection of cognitive impairment allows for earlier treatment, potentially slowing or halting cognitive decline. A 30-min web-camera eye tracking assessment (30-min VPC) has been validated as a tool to predict AD risk. However, a shorter version would allow for greater scalability and improve user experience. The purpose of this study was to: 1) determine the validity of the 5-minute web-camera based VPC test with the 30-min test, 2) determine the test-retest reliability of the 5-min test, 3) compare the 5-minute test scores of cognitively intact adults (18-39 years of age) to the scores of cognitively intact older adults (>65 years of age), 4) examine the relationship between the 5-min web camera based VPC test and additional cognitive tests. This prospective study included two groups, both with normal cognitive function: 24 young adults ( $26.5 \pm 7.4$  years) and 20 older adults ( $79.3 \pm 6.4$ ). Participants were tested on two separate occasions. Trial 1 included the Montreal Cognitive Assessment (MoCA), Digit Symbol Coding, NIH Toolbox Assessment, 30-min VPC, and 5-min VPC. Trial 2 occurred at least 14 days later and participants completed the 5-min VPC, Digit Symbol Coding, NIH Toolbox, and dual-task walking assessments. The 5-min VPC significantly correlated to the 30-min VPC during the first ( $p = .003$ ) and second ( $p = .001$ ) trial, and showed significant test re-test reliability ( $p < .001$ ). The 5-min VPC mean scores were 83% and 80% for Trial 1 and Trial 2, respectively, with a significant time interaction ( $p = .04$ ). There was a significant relationship between the 5-min VPC and DCCS ( $p = .03$ ), MoCA ( $p = .00$ ), and Digit Symbol Coding ( $p = .00$ ), during Trial 1. As well as Flanker ( $p < .01$ ), PCPST ( $p = .00$ ), PSMT ( $p < .01$ ), MoCA ( $p = .00$ ), and Digit Symbol Coding ( $p = .00$ ), during Trial 2. The results from this study suggest the 5-min VPC test is a valid and reliable tool to assess cognitive function.

## **Acknowledgements**

I would like to thank Dr. Gray for her guidance and support throughout my academic career. I would also like to thank Dr. Washington and Dr. Greene for sitting on my committee. I would like to thank my lab group for collecting data for this study. Lastly, I would like to thank the faculty and staff at the University of Arkansas.

## Table of Contents

<b>Introduction</b>	<b>1</b>
Operational Definitions	3
<b>Literature Review</b>	<b>3</b>
<b>Methodology</b>	<b>8</b>
Research Design	8
Measures	9
Procedures	12
Data Analysis	13
<b>Results</b>	<b>13</b>
Participants Characteristics	13
Relationship between the 30-min VPC and 5-min VPC Tests	15
Reliability of the 5-min VPC Test	15
Comparing 5-min VPC Test Scores Between Young and Older Adults	15
Associations Between the 5-min VPC Test and Other Cognitive Assessments	16
<b>Discussion</b>	<b>18</b>
Threats to Validity	20
Assumptions	21
Limitations	21
Delimitations	21
<b>Conclusion</b>	<b>22</b>
<b>References</b>	<b>23</b>

## Introduction

According to the Alzheimer's Association (2018), Alzheimer's Disease (AD) is the sixth leading cause of death in the United States (US), with approximately 5.7 million Americans currently living with AD, of which 5.5 million are 65 years of age or older. From 2000 to 2015, the number of deaths due to heart disease decreased by 11%, whereas the deaths from AD increased by 123%. This presents the importance of research and education of the general population, in an effort to decrease the rate of deaths and ideally decrease prevalence of AD altogether. In 2018 alone, AD and other dementias were estimated to cost the US \$277 billion dollars, with both costs estimated to rise to \$1.1 trillion and prevalence estimated to increase to 14 million, by the year 2050. With early detection and accurate diagnosis, up to \$7.9 trillion could be saved in medical and care costs. While there is not currently a cure for AD, studies have shown various interventions have the ability to slow the progression of the disease (Nelson & Tabot, 2015; Nganda et al., 2015; Rosenburg et al., 2017). This means early detection of cognitive decline is crucial.

Cognitive disorders are classified on a spectrum dependent on phenomenological, epidemiological, clinical, neuropsychological, and biological variables. Mild Cognitive Impairment (MCI) is the first stage of cognitive decline, characterized as memory loss that deviates from normal aging (Alistair & Zandig 2002; Zola, Manzanares, Clopton, Lah, & Levey, 2012). The development of MCI puts the individual at a greater risk of developing dementia, which is a diagnosis of a group of symptoms associated with a decline in memory and thinking skills, severe enough to interfere with the ability to perform activities of daily living. Furthermore, AD is a specific disease, within the diagnosis of dementia, that causes a decline in memory, thinking, and behavior, specifically inhibiting the ability to remember newly learned

information and affecting mood and behavior (Alzheimer's Association, 2018). MCI does not always progress to dementia and the development of dementia varies tremendously between individuals who do convert (Hayat et al., 2014). However, MCI puts individuals at an increased risk of developing AD when compared to healthy individuals, with the progression estimated at 6% to 25% per year (Crutcher et al., 2009). Furthermore, AD is the most common cause of dementia and makes up 60-80% of total dementia cases (Alzheimer's Association, 2018).

It is common for changes in memory to go undetected for the first 7-10 years. It is important to know the risk of developing dementia and to detect impairments as soon as possible, so preventive interventions can be implemented, in an effort to slow the progression of the disease. This means effective, highly sensitive memory tests are imperative to detect possible deficiencies, as early as possible (Elias et al., 2000). Visual Paired Comparison (VPC) is a validated memory recognition test used to compare the amount of time a subject spends viewing a new/novel image to an image already seen, when displayed side-by-side. Being able to track these eye movements during VPC and use the results as a screening tool for cognitive deficiencies, improves the ability to detect early stages of MCI, predict the risk of developing MCI/dementia, and track cognitive function over time (Crutcher et al., 2009; Daffner, Scinto, Weintraub, Guinessey, & Mesulam, 1992; Lagun, Manzanares, Zola, Buffalo, & Agichtein, 2011).

A commercial grade eye tracker, which consists of a high frame rate camera, capable of capturing a number of visual features, has been validated to track and record eye movement during VPC (Crutcher et al., 2009; Lagun et al., 2011) However, these cameras are expensive, which led to the validation of web-cameras as an alternative. Web-cameras are readily available on most portable devices, including laptops, tablets, and phones, which makes them inexpensive

and accessible to the public (Y. Lin, R. Lin, Y. Lin, & Lee, 2013). A 30-minute VPC eye tracking assessments, using a built-in web-camera, has been validated (Bott et al., 2017). However, with the extensive amount of time required for testing, a shorter version of the same assessment is needed to increase efficiency and improve the ability to use the test as a screening tool.

### **Operational Definitions**

Cognition encompasses the processes of the brain involved in the ability to gain knowledge and comprehend, which includes thinking, knowing, remembering, judging, and problem solving. These functions are used for language, imagination, perception, and the ability to plan and execute certain behaviors (National Institutes of Health, 2017). MCI is described as the stage between normal cognitive ageing and dementia. There are clear signs of impairment to memory when compared to healthy controls, but impairment is less than patients with mild dementia. Dementia is a non-specific diagnosis used to describe a group of symptoms, including the decline of cognition and other skills that may impair a person's ability to perform activities of daily living. Furthermore, AD is a specific disease within the diagnosis of dementia that causes a decline in memory, thinking, and behavior, specifically inhibiting the ability to remember newly learned information (Alzheimer's Association, 2018). VPC task is a method that displays a series of two identical images side-by side, which become the familiar images, followed by a series of slides with a familiar image displayed next to a new/novel image (Crutcher et al., 2009; Daffner et al., 1992).

### **Literature Review**

MCI, dementia, and AD are all diagnoses for varying levels of cognitive impairment. MCI is broad term used to describe memory loss, which has been linked to changes in the medial

temporal lobe, impairing recollection. With MCI, often other cognitive abilities are preserved, but it does have the potential to develop to dementia. The development of dementia is multifactorial, including both genetic and environmental factors. Risk factors may include, but are not limited to; age, alcohol abuse, smoking, hypertension, obesity, dyslipidemia, vascular damage, and neural damage. However, physical activity, education, and cognitive and social activities have shown to protect against the development of dementia. The current method for predicting the risk for developing dementia is the CAIDE (Cardiovascular Risk Factors, Aging and Dementia) Dementia Risk Score which uses age, education, gender, blood pressure, body mass index, total cholesterol, and physical activity. The greater the score, the greater the risk of dementia (Kivipelto et al., 2013).

VPC task recognition aims to detect impairments in memory to predict MCI and AD development. In 1992, researchers developed an objective measure to define cognitive impairment by using an infrared light camera to track the pupils of participants. Curiosity slides, which contain a regular and irregular figure, and incongruous picture slides, a photo containing an irregular stimulus, such as a lion sitting at a desk in a classroom of children, were used. AD patients spent significantly less time looking at both the irregular figures and the irregular portions in the incongruous pictures, when compared to controls who were matched based on age, education, and work experience (Daffner et al., 1992). The ability to measure, and track cognitive impairment over time, allows interventions to be tested to determine ways to further prevent cognitive decline.

Current VPC methods use familiar and novel images to detect impairment. Participants are shown two identical images, side-by-side, during a familiarization period. With varying time between stimuli, the testing period displays old and new images, side-by-side. Using both two-

second delays and two-minute delays between stimuli, Crutcher et al. (2009), compared an MCI group to a matched Parkinson's and cognitively intact control group. Results found there was no significant difference between groups during the two-second delay. However, during the two-minute delay, the MCI group spent approximately 53% of total time viewing the new image, whereas the Parkinson's and cognitively intact groups spent 71% and 74% of their time on the new image, respectively. VPC specifically measures the percentage of time spent looking at novel images, which target declarative memory. These results suggest VPC task has the potential to be used as a screening tool for detecting impairment. Individuals spending less than 65% of time looking at the novel image have an increased risk of developing MCI/AD.

In addition to the infrared eye tracking system, Lagun et al. (2011) validated the use of Support Vector Machines (SVM), which is an automatic classification algorithm. SVM accounts not only the amount of time spent looking at the novel image verse the familiar image, but also the time spent on the "grey area" between images. Most importantly, it detects eye movement patterns, specifically fixations, saccades, and re-fixations. Prior techniques for distinguishing normal subjects from MCI subjects was 67% accurate, 60% sensitive, and 73% specificity. Using SVM performance of accuracy, sensitivity, and specificity increased to 87%, 97%, and 77%, respectively. Potentially, this allows for more sensitive detection of MCI and/or the risk of developing MCI in participants prior to diagnosis.

These findings attest to the validity of VPC and its ability to differentiate between control individuals, who have normal cognitive function, and individuals with MCI. A three-year longitudinal study, using 32 MCI participants and 60 control participants, was used to determine the ability of VPC to predict which individuals in the control group and which MCI individuals would or would not progress to develop MCI and/or AD, respectively. At the conclusion of the

three-year trial, 13 MCI patients developed AD and four control subjects developed MCI, which placed them in the converter group (CONV). The remaining participants were placed in the non-converter group (NONCONV). During the VPC task, the scores associated with the percentage of time looking at the novel image was significantly lower in the CONV group compared to the NONCONV group, with no significant difference between the CON and MCI participants within the CONV group. Additionally, when the same scores were compared to a group of 20 patients with AD, there was no significant difference. Furthermore, both the participants with MCI and the CON participants in the CONV group had significantly lower VPC mean scores than the NONCONV group. Most importantly, out of the 43 subjects with a VPC score  $\geq 67$ , there were zero participants who converted. On the contrary, eight of the nine participants with a VPC score  $\leq 50$  converted over the three-year span. A receiver-operating characteristic curve using scores from all the participants showed an area under the curve of .903, which displays a strong ability for VPC to predict the chances of a participant's diagnosis changing within the next three years (Zola et al., 2012).

With the accuracy, sensitivity, and specificity of VPC increasing using advanced technology, Bott et al. (2017), further investigated the idea of administering VPC using a web-camera, in an effort to make VPC more accessible to the general population. Research eye tracking cameras not only require a significant amount of setup and maintenance, but also cost up to \$50,000. This has confined VPC testing to research labs, making it difficult to test large populations. Web-cameras have become standard hardware in smart devices and a readily available to the general population, without additional costs. Fifty-four clinically normal older adults completed three trials of VPC while simultaneously recording eye tracking using both the standard eye tracking camera, scored automatically, which collects 60 frames per second (FPS),

and the built-in laptop web-camera, scored manually by human coders, collecting three FPS, five FPS, and ten FPS. With a strong correlation between all three-manual web-camera scores, the three FPS scores were compared to the standard eye tracking score. The results showed a strong mean correlation of .76, between the eye tracker camera and the web-camera at three time points. The web-camera based tasks were also more reliable with only 2.3% of tasks excluded due to technical issues compared to 10% of the eye tracker-based tasks. The manual recording not only showed fewer data quality issues, but also showed a strong correlation to the automatic scoring methods previously used. With web cameras being easily accessible to the population, this allows for VPC to be administered with minimal costs, while still producing valid and reliable results.

There is currently no cure, but with early detection, there are interventions aimed at slowing the progression of AD. Dietary interventions, including the Mediterranean diet and the Dietary Approaches to Stop Hypertension (DASH), have been shown to produce greater scores in the Mini-Mental State Examination and improved psychomotor speed, when compared to controls. The Mediterranean-DASH Intervention for Neurodegenerative Delay (MIND) diet has been developed to treat patients with AD. Results showed that high adherence to the previously mentioned diets, along with moderate adherence to the MIND diet, all showed decreased risk of AD (Morris et al., 2015). Furthermore, the FINGER multi-domain approach, including diet, strength and aerobic training, cognitive, and social interventions over a 24-month period was used in a sample of 1,168 older adults, ages 60-77 years, who were at an increased risk for developing dementia. Results showed improved executive function, processing speed, memory, and neurological test scores, compared to the control who received only regular health advice (Ngandu et al., 2015). Using the same intervention, Rosenburg et al. (2017), found that

sociodemographic, socioeconomic status, cognition, cardiovascular factors, and cardiovascular comorbidity did not have an effect on the outcome of the multi-domain intervention. These interventions can be used in conjunction with the web-camera administration of VPC tasks to potentially reduce the number of participants who are at a high risk to convert from normal to MCI and/or MCI to AD.

In order to not only slow the progression of dementia, but to prevent cognitive decline overall, methods used to detect impairment and measures used to calculate risks for developing dementia in the future need to be further developed and perfected. Early detection and prevention are imperative to battle a disease that has no cure.

The purpose of this study was multi-factorial and includes the following: 1) determine the validity of the 5-minute web-camera based VPC test with the 30-min test, 2) determine the test-retest reliability of the 5-min test, 3) compare the 5-minute test scores of cognitively intact adults (18-39 years of age) to the scores of cognitively intact older adults (>65 years of age), 4) examine the relationship between the 5-min web camera based VPC test and additional cognitive tests.

## **Methodology**

### **Research Design**

This was a non-experimental, prospective study. This design compared scores both within the young adult group (18-39 years of age), and between the young adults and the older adults (>60 years of age). The independent variable was the age group cohort. The dependent variables were the 5-min VPC results and the scores of the additional cognitive tests: Montreal Cognitive Assessment (MoCA), Digit Symbol Coding Test, NIH toolbox assessments, and dual-task results.

## Measures

**Demographics.** Height was measured, to the nearest 0.5cm, using a standing stadiometer. Subjects are asked to remove their shoes and stand up as straight as possible. Weight was measured, to the nearest 0.5kg, using a balance-beam scale. Subjects were asked to remove their shoes and heavy clothing (sweaters/jackets/coats) and empty their pockets. Body composition was measured using a dual-energy x-ray absorptiometry (DXA).

**MoCA.** The Montreal Cognitive Assessment (MoCA) is a paper-pencil test used to screen for cognitive impairment and assesses multiple cognitive domains. These cognitive domains include attention and concentration, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. The test takes approximately 10 minutes, with a possible 30 points. If a score  $\geq 26$  was obtained, the participant was categorized as cognitively normal and completed the other assessments. With a score  $< 26$  or diagnosed with AD, the participant was categorized as cognitively impaired. For these individuals, the signature of a guardian was required in order to complete the remaining tests. The ability of the MoCA to identify MCI and dementia has been shown to be 84-90% and 94-100% effective, respectively (Nasreddine et al., 2005; Smith, Gildeh, & Holmes, 2007).

**Digit Symbol Coding Test.** The Digit Symbol Coding Test is a paper-pencil test used to assess cognition, specifically processing speed and executive functioning. Rows of blank squares were matched with randomly assigned numbers from one to nine. Above the rows was a reference key that pairs each number with a different symbol. The reference key was used to pair as many symbols as possible with the assigned number, within 90 seconds. Digit Symbol Coding was found to be the best test for screening cognitive impairment, when compared to various other assessments, with a high test-retest reliability ( $r = .85$ ) (Gonzalez-Blanch et al., 2011).

**NIH Toolbox.** NIH developed an application with a standard set of validated, reliable tests used to assess neurological and behavioral function. The following tests are within the cognitive domain of the NIH Toolbox: Flanker Inhibitory Control and Attention, Dimensional Change Card Sort (DCCS), Pattern Comparison Processing Speed, Picture Sequence Memory. The tests were administered using an iPad, with the order of test administration to always follow the order listed above. Prior to the test trials, it was demonstrated how to complete each task, in addition to a given number of practice sets. If the participant answered at least three out of four correctly during the practice set, the test began. If less than three were correctly answered, two additional practice sets were given. However, the test ended if three answers were not correct by the end of the third practice set (NIH, 2017). When compared to the Gold Standard tests, the NIH Toolbox has been shown to be a valid and reliable assessment, for adults, of six domains of cognition: working memory, executive function, episodic memory, processing speed, language and reading. Each assessment within the NIH toolbox is strongly related to the domain intended, while also being weakly related to the other domains, which attests to the validity of the toolbox, in its entirety. There were no differences found, when compared to gold standard measures, across 20-85 year age groups, allowing for reliable and consistent testing over the course of a lifetime (Mungas et al., 2014).

***Flanker Inhibitory Control and Attention Test.*** The participant was instructed to focus on a particular stimulus while inhibiting attention to the stimuli flanking it. A row of arrows was presented on the screen. Sometimes the arrows were pointing in the same direction and sometimes the arrows were pointing in different directions. The objective was to select the button that matches the way the middle arrow was pointing. This was a method used to measure inhibitory control and attention (NIH, 2017).

***Dimension Change Card Sort Test (DCCS).*** This is a cognitive test used to measure flexibility and attention. Two target pictures were presented that vary along two dimensions (ex., shape and color). Participants were asked to match a series of bivalent test pictures (ex., yellow balls and blue trucks) to the target pictures, first according to one dimension (ex., color) and after numerous trials, according to the other dimension (ex., shape). Instruction on how to match pictures according to shape were given along with the necessary practice trials. Next, it was explained how to match the pictures according to color, again followed by the practice items. The 30-test items are a combination of both matching color and shape dependent on what was indicated in the middle of the screen (NIH, 2017).

***Pattern Comparison Processing Speed Test (PCPST).*** In order to assess processing speed, two pictures were presented on the screen side-by-side. If the pictures were identical, the participants selected the “yes” button and if the pictures were not identical, the participant selected the “no” button. Six practice items were given before the test begins. With a total of 130 items, the test was completed when all 130 items appear or after 85 seconds have elapsed (NIH, 2017).

***Picture Sequence Memory Test (PSMT).*** Sequences of pictured objects and activities were presented in a specific order to test episodic memory. The objective was to reproduce the pictures in the same order they appeared on the screen. It was demonstrated how to drag the picture on the screen prior to the four-step practice sequence. The test consisted of two trials. The first trial was a 15-step sequence and the second trial was an 18-step sequence, which included all 15 pictures from the first trial (NIH, 2017).

***Visual Processing Comparison (VPC).*** The VPC tests (Neurotrack’s 30-min and 5-min tests) were taken on a laptop equipped with a web-camera. The assessments were recorded and

the videos of the participants' face were stored so their eye movements can be tracked. The 5-min VPC test is a condensed version of the 30-minute test. Additionally, the 5-min test is active because it instructs the participant to look at the new image they have not yet seen, whereas the 30-min test is passive and does not instruct the participant what to do while looking at the images (NIH, 2017). The videos of the participants' faces were analyzed by Neurotrack technicians.

**Dual-task.** Cognitive function and the ability to walk are related, meaning cognitive impairment has an effect on gait ability, which can be measured using dual-task assessments. Dual-task has shown to be reliable ( $r \geq .75$ ) when testing individuals with Parkinson's disease, MCI, dementia, AD, and multiple sclerosis (Yang et al., 2017). Dual-task assessment measures the time it takes to walk 10-meters under various instructions. The first measure was walking at a normal pace. The second was walking at a fast pace, without jogging. Two trials at each pace were completed, timed, and recorded. The next measure was at walking at a normal/habitual pace while counting down by threes from a random number assigned to them. The final was the same dual-task but at the maximal walking pace. Different random numbers were assigned to each of the four dual-task trials. Individuals with cognitive impairment had more difficulty completing the dual-task trials compared to healthy adults, which resulted in increased time to walk 10-meters (Doi et al., 2014).

## **Procedures**

Once IRB approval was obtained, adults (18-39 years) and older adults (>60 years) with normal cognitive function, were recruited. A total of 44 participants were used to complete the study (24 young adults and 20 older adults). The young adults reported to the Exercise Science Research Center at the University of Arkansas for two separate trials. The older adults completed the assessments at Butterfield Trail Village. For the first trial, upon arrival, participants

completed the informed consent document, the medical history questionnaire, and the MoCA, to screen for cognitive impairment, prior to the administration of the cognitive assessments (Digit Symbol Coding Test and NIH Toolbox) and the 30-min and 5-min VPC tests.

The order of the cognitive tests remained the same for all trials. However, the participants were randomly assigned to determine the order of test administration for the cognitive tests and the VPC tests. During the first trial, if the participant completed the cognitive assessments before the VPC tests, they were administered the VPC test before the cognitive assessments during the second trial, and vice versa.

The second trial occurred at least 14 days after the first trial. During the session subjects repeated the cognitive tests (Digit Symbol Coding Test and the NIH Toolbox) and the 5-min VPC test. Additionally, dual-task measures and demographics (height, weight, and body composition) were taken during the second trial.

### **Data Analysis**

To determine the validity of the 5-min test compared to the 30-min VPC and the other cognitive assessments, a Pearson correlation was used. To determine the test-retest reliability of the 5-min VPC test, a paired t-test using the day 1 and day 14 data was completed. Additionally, a repeated measure ANOVA was used to determine differences that may exist between age groups. Statistical significance was set at  $\alpha = .05$ . Demographic information was presented as means  $\pm$  SD.

## **Results**

### **Participants Characteristics**

A total of 44 adults completed the study. This included 24 young adults ( $26.5 \pm 7.4$  years) and 20 older adults ( $79.3 \pm 6.4$  years). Overall, the sample was 64% female (28 subjects) and

75% of participants were college or trade school graduates. Not all web-camera VPC data were able to be scored due to technical difficulties, which included poor recording quality or network connectivity, glare from glasses, and/or low light in the testing room. Of the 44 participants who completed the assessment, 98% of the 30-min VPC scores were evaluated, 100% of Trial 1 and 88.5% of the Trial 2 of the 5-min VPC tests were analyzed. On average, the second trial session took place 22.4 days after the first. Table 1 displays the characteristics of the study cohort.

Table 1

*Characteristics of the study cohort (n = 44).*

	Young Adults	Older Adults	Significance ( <i>p</i> )
N	24	20	
Average age (years)	26.5 ± 7.4	79.3 ± 6.4	<i>p</i> < .001
Sex			<i>p</i> = .87
N Female (%)	15 (62.5%)	13 (65.0%)	
Education			<i>p</i> = .08
High School Graduate	4.2%	5.0%	
Some College	29.2%	10.0%	
College Graduates or higher	66.7%	85%	
Race			<i>p</i> = .21
European-American	83.3%	95.0%	
Other	16.7%	5.0%	
Biometric			
Height (cm)	170.1 ± 8.6	166.5 ± 8.6	<i>p</i> = .38
Weight (kg)	74.5 ± 17.8	73.2 kg ± 14.3	<i>p</i> = .96

\**Note.* *p*-values were measured by independent t-tests to determine differences between groups.

### Relationship between the 30-min VPC and 5-min VPC Tests

Among the young adults, analysis displayed a significant positive relationship between the 30-min VPC and 5-min VPC tests. The 30-min VPC was significantly correlated with the 5-min VPC during the first ( $r = .601$ ;  $p = .003$ ) and second ( $r = .722$ ;  $p = .001$ ) trials. This shows the validity of the 5-min VPC test and its ability to measure declarative memory.

### Reliability of the 5-min VPC Test

The results of Trial 1 were positively correlated to the Trial 2 results for the 5-min VPC test ( $r = .81$ ;  $p = .00$ ). This shows the reliability of the assessment. Furthermore, there was no significant difference between Trial 1 and 2 for the 5-min VPC test ( $p = .80$ ). Mean scores for Trial 1 and 2 were .812 and .808, respectively.

### Comparing 5-min VPC Test Scores Between Young and Older Adults

Repeated measures ANOVA revealed no change over time between young and older subjects for the 5-min VPC assessment ( $F = 3.16$ ,  $p = .08$ ; Figure 1). Mean VPC values were 83% (young adults, 81%; older adults, 85%) and 80% (young adults, 81%; older adults, 80%) for Trials 1 and 2, respectively (Figure 2).

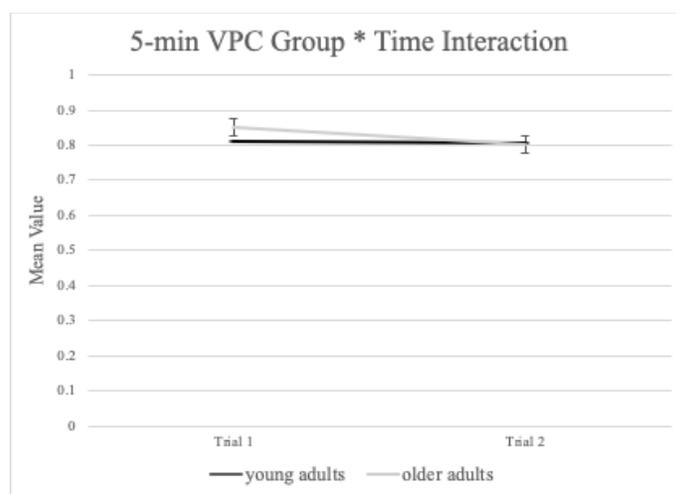


Figure 1. The time effect interaction of 5-min VPC Trial 1 and 2, between the young and older adults.

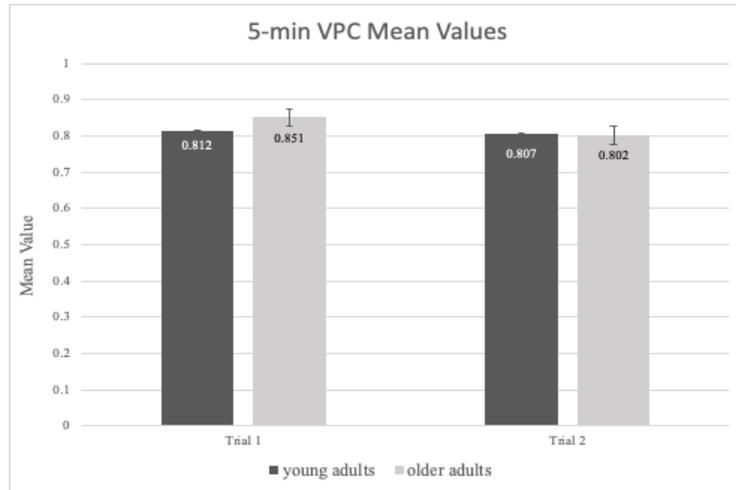


Figure 2. The mean values of the 5-min VPC, Trial 1 and 2, for the young and older adults.

### Associations Between the 5-min VPC Test and Other Cognitive Assessments

Using cognitive function as a specific domain, the correlation between the 5-min VPC test and the NIH Toolbox assessments was examined. The age-based normative percentile scores were used to compare each task to the performance of the 5-min VPC test. During Trial 1 (Table 2), there was a significant relationship between the 5-min VPC and DCCS ( $r = .31$ ;  $p = .03$ ), MoCA ( $r = .50$ ,  $p = .00$ ), and Digit Symbol Coding ( $r = .47$ ,  $p = .00$ ), but not the Flanker ( $r = .14$ ,  $p = .33$ ), PCPST ( $r = .22$ ,  $p = .14$ ) or PSMT ( $r = -.09$ ,  $p = .55$ ). During Trial 2 (Table 3), significant correlations were found between the 5-min VPC test and Flanker ( $r = .38$ ,  $p < .01$ ), PCPST ( $r = .52$ ,  $p = .00$ ), PSMT ( $r = .49$ ,  $p < .01$ ), MoCA ( $r = .65$ ,  $p = .00$ ), and Digit Symbol Coding ( $r = .57$ ,  $p = .00$ ), but not DCCS ( $r = .23$ ,  $p = .12$ ), dual-task habitual speed ( $r = -.24$ ,  $p = .09$ ), or dual-task maximal speed ( $r = .02$ ,  $p = .88$ ).

**Table 2.**

*Correlations between 5-min VPC test data from Trial 1 and cognitive assessments from Trial 1.*

	<u>VPC-5</u>	<u>VPC-30</u>	<u>Flanker</u>	<u>DCCS</u>	<u>PCPST</u>	<u>PSMT</u>	<u>MoCA</u>	<u>DSC</u>
<u>VPC-5</u>	-							
<u>VPC-30</u>	.55**	-						
<u>Flanker</u>	.14	-.09	-					
<u>DCCS</u>	.31*	.00	.62**	-				
<u>PCPST</u>	.21	.20	.61**	.52**	-			
<u>PSMT</u>	.09	.13	.32*	.28*	.35*	-		
<u>MoCA</u>	.50**	.26	.47**	.27*	.57**	.39**	-	
<u>DSC</u>	.47**	.35**	.45**	.47**	.43**	.58**	.51**	-

*Note.* \* $p < .05$ ; \*\* $p < .01$ ; Flanker = Flanker Inhibitory Control Test, DCCS = Dimension Change Card Sort Test, PCPST = Pattern Comparison Processing Speed Test, PSMT = Picture Sequence Memory Test, MoCA = Montreal Cognitive Assessment, DSC = Digit Symbol Code Test

<b>Table 3.</b>										
<i>Correlations between 5-min VPC test data from Trial 2 data and cognitive assessments from Trial 2.</i>										
	<u>VPC-5</u>	<u>VPC-30</u>	<u>Flanker</u>	<u>DCCS</u>	<u>PCPST</u>	<u>PSMT</u>	<u>MoCA</u>	<u>DSC</u>	<u>DT-HS</u>	<u>DT-MS</u>
<u>VPC-5</u>	-									
<u>VPC-30</u>	.49**	-								
<u>Flanker</u>	.38**	.09	-							
<u>DCCS</u>	.23	.03	.60**	-						
<u>PCPST</u>	.52**	.17	.56**	.57**	-					
<u>PSMT</u>	.49**	.27	.09	.25	.33*	-				
<u>MoCA</u>	.65**	.26	.39**	.35*	.58**	.51**	-			
<u>DSC</u>	.57**	.31*	.54**	.40**	.68**	.43**	.57**	-		
<u>DT-HS</u>	-.24	-.17	-.32*	-.20	-.32*	-.37*	-.39**	-.32*	-	
<u>DT-MS</u>	.02	.11	-.20	-.18	-.19	.09	-.02	-.02	.13	-
<p><i>Note.</i> * <math>p &lt; .05</math>; ** <math>p &lt; .01</math>; Flanker = Flanker Inhibitory Control Test, DCCS = Dimension Change Card Sort Test, PCPST = Pattern Comparison Processing Speed Test, PSMT = Picture Sequence Memory Test, MoCA = Montreal Cognitive Assessment, DSC = Digit Symbol Code Test, DT-HS = Dual-task Habitual, DT-MS = Dual-task maximal speed</p>										

### Discussion

The four primary objectives of this study were to: 1) validate the 5-minute web-camera based VPC test using the 30-min VPC test, 2) determine the test-retest reliability of the 5-min

test, 3) compare the 5-min VPC test scores of the young adults (18-39 years of age) to the scores of the older adults (>60 years of age), and 4) examine the relationship between the 5-min VPC test and additional cognitive assessments. Among young adults, the results showed adequate validity between the 5-min VPC and 30-min VPC test. The 5-min VPC scores were also significantly correlated between trials, with no significant differences between Trial 1 and Trial 2. Not only is the 5-min VPC test shorter, but it is an active test (i.e. participants are instructed to look at the new image), which differs from the passivity of the 30-min VPC test. This is important because the 5-min VPC is intended to be taken repeatedly over time, rather than solely being used as a baseline measure. This eliminates the need to conceal what the test is measuring from the participant. Furthermore, the duration of the 5-min VPC lessens the stress on the participant and increases the practicality of test administration.

There was a significant time effect, between trial scores for the 5-min VPC test. However, the differences between the groups over time was not significant. This means the changes in scores, from Trial 1 to Trial 2, did not differ between age groups. Statistically, the group by time interaction is not significant, but amongst the older adults specifically, the interaction is trending toward significance.

The 5-min VPC showed to be significantly related to standard cognitive assessments, which suggests the 5-min VPC may be used as a convenient alternative to the longer cognitive assessments. This supports the results of previous research which has shown a strong relationship between VPC and other cognitive assessments among the older adult population, supporting the specific ability of VPC tests to assess processing speed (PCPST), visual episodic memory (PSMT), and executive function (Digit Symbol Coding) (Bott et al., 2018). Additionally, learning effects associated with the repetitive administration of cognitive assessments has been

shown to potentially cause issues with internal validity (Goldberg, Harvey, Wesnes, Snyder, & Schneider, 2015). This is the case when administering NIH Toolbox assessments and Digit Symbol Coding multiple times, due to the availability of only one version of each assessment (Heaton et al., 2014). The minimum of 14 days between trials aims to alleviate the learning effect. The learning effect in the 5-min VPC was reduced by the use of alternate forms of the test during Trial 1 and 2.

The issue of data quality of the VPC tests serves as a limitation of this study. A total of six VPC tests were not able to be scored due to glare from glasses, poor lighting in the testing room, and/or electronic errors. The tests that were excluded were: the 30-min VPC of one young adult and the 5-min VPC trial 2 assessments of five young adults. The subjects completed both visits at the testing center, so they were still included in the overall analyses. The assessments were completed in a research setting, to ensure high quality data collection, limiting the ecological validity of the results. This study included only individuals with normal cognitive function. Further research including individuals with cognitive impairment, as well as the investigation of varying levels within individuals with cognitive impairment, is needed to further validate VPC and its ability to assess cognition among all populations. Additionally, a longitudinal study tracking participants overtime, exclusively utilizing the 5-min VPC test, would verify the ability of VPC administration in clinical settings, to track cognition throughout a lifetime.

### **Threats to Validity**

The results of the cognitive tests were based off the participants willingness to complete the tests to the best of their abilities. If this was not done, it is a possible threat to the internal validity of the tests. The possible malfunction of technology may pose as a threat, however the

manual analysis of VPC eye tracking ensured greater accuracy. The recruitment for the young adults was done through the university, meaning it is possible a large sum of participants were educated, either pursuing a bachelor's or graduate degree. This may have affected the overall scores for the cognitive tests, which does not represent the general population. This affected the comparison of the 5-min VPC scores to the cognitive test scores. However, this did not affect the results of the test-retest reliability and the validity of the VPC results.

### **Assumptions**

It is assumed the participants completed all cognitive tests as quickly as possible and to the best of their ability. Furthermore, the participants were expected to pay attention throughout the entirety of the VPC tests. It is also assumed participants did not suffer from cognitive impairments, which was screened for using the medical history questionnaire and MoCA.

### **Limitations**

The 30-min VPC test was only administered during the first trial. Even though a crossover design was used, it is possible mental fatigue may have affected the scores of the participants who were required to complete the NIH cognitive assessments after the 30-min test.

### **Delimitations**

The following are delimitations of the current study. A total of 44 subjects participated in the study. Cognitive assessments were completed using both paper-pencil tests and the NIH toolbox on an iPad. Subjects were tested in the Exercise Science Research Center at the University of Arkansas, or at Butterfield Trail Village. The trials were administered at least 14 days apart in order to determine test-retest reliability.

## Conclusion

Among the young adults, the 5-min VPC test demonstrated a moderate convergent validity to the 30-min VPC, along with strong test-retest reliability. Between groups, there was a significant time interaction, with group by time effect trending in significance. The significant relationships between the 5-min VPC and the other cognitive assessments verifies previous research, specifically supporting its ability to assess processing speed, visual episodic memory, and executive function (Bott et al., 2018). Furthermore, the cohort included both older and young adults, generalizing the results to a greater population.

The scalability of the 5-min VPC test may be greater than other cognitive assessments due to its short duration, lessened strain on participants, and convenience of digital delivery. Unlike other cognitive assessments, test administration by a trained professional is not needed (Heaton et al., 2014). The 5-min VPC test can be completed anywhere with an internet connection and web-camera, with minimal instructions required (Bott et al., 2018). This is not a diagnostic test for cognitive impairment, but rather a screening tool used as a baseline and way to monitor cognition over time in healthy individuals. The efficacy of testing individuals with cognitive impairment, as well as testing in remote environments, requires further investigation.

## References

- Alistair, B. & Zaudig, M. (2002). Cognitive impairment in older adults. *The Lancet*, 360, 1963-1965.
- Alzheimer's Association. (2018). 2018 Alzheimer's disease facts and figures. Retrieved from <https://www.alz.org/media/HomeOffice/Facts%20and%20Figures/facts-and-figures.pdf>
- Bott, N., Madero, E. N., Glenn, J., Lange, A., Anderson, J., Newton, D., ... Zola, S. (2018). Device-Embedded Cameras for Eye Tracking–Based Cognitive Assessment: Validation with Paper-Pencil and Computerized Cognitive Composites. *Journal of Medical Internet Research*, 20(7).
- Bott, N.T., Lange, A., Rentz, D., Buffalo, E., Clopton, P., & Zola, S. (2017). Web camera based eye tracking to assess visual memory on a visual paired comparison task. *Frontiers in Neuroscience*, 11(370).
- Crutcher, M.D., Calhoun-Haney, R., Manzanares, C.M., Lah, J.J., Levey, A.I., & Zola, S.M. (2009). Eye tracking during a visual paired comparison task as a predictor of early dementia. *American Journal of Alzheimer's Disease & Other Dementias*, 24(3), 258-266.
- Daffner, K.R., Scinto, L.M., Weintraub, S., Guinessey, J.E., & Mesulam, M.M. (1992). Diminished curiosity in patients with probable Alzheimer's disease as measured by exploratory eye movements. *Neurology*, 42, 320-328.
- Doi, T., Shimada, H., Makizako, H., Tsutsumimoto, K., Uemura, K., Anan, Y., & Suzuki, T. (2014). Cognitive function and gait speed under normal and dual-task walking among adults with mild cognitive impairment. *BMC Neurology*, 14(67).
- Elias, M. F., Beiser, A., Wolf, P. A., Au, R., White, R. F., & D'agostino, R. B. (2000). The preclinical phase of Alzheimer disease: a 22-year prospective study of the Framingham Cohort. *Archives of neurology*, 57(6), 808-813.
- Goldberg, T. E., Harvey, P. D., Wesnes, K. A., Snyder, P. J., & Schneider, L. S. (2015). Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(1), 103–111.
- Gonzalez-Blanch, C., Perez-Iglesias, R., Rodriguez-Sanchez, J.M., Pardo-Garcia, G., Martinez-Garcia, O., Vasquez-Barquero, J.L., & Crespo-Facorro, B. (2011). A digit symbol coding task as a screening instrument for cognitive impairment in first-episode psychosis. *Archive of Clinical Neuropsychology*, 26, 48-58.
- Hayat, S.A., Luben, R., Moore, S., Dalzell, N., Bhaniani, A., Anuj, S., . . . Brayne, C. (2014). Cognitive function in a general population of men and women: A cross sectional study in the European investigation of Cancer-Norfolk cohort (EPIC-Norfolk). *BMC Geriatrics* 14, 142.

- Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., . . . Gershon, R. (2014). Reliability and Validity of Composite Scores from the NIH Toolbox Cognition Battery in Adults. *Journal of the International Neuropsychological Society*, 20(06), 588–598.
- Kivipelto, M., Solomon, A., Ahtiluoto, S., Ngandu, T., Lehtisalo, J., Antikainen, R., . . . Soininen, H. (2013). The finnish geriatric intervention study to prevent cognitive impairment and disability (FINGER): Study design and progress. *Alzheimer's & Dementia*, 9, 657-665.
- Lagun, D., Manzanares, C., Zola, S.M., Buffalo, E.A., & Agichtein, E. (2011). Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. *Journal of Neuroscience Methods*, 201, 196-203.
- Lin, Y. T., Lin, R. Y., Lin, Y. C., & Lee, G. C. (2013). Real-time eye-gaze estimation using a low-resolution webcam. *Multimedia Tools and Applications*, 65(3), 543-568.
- Morris, M.C., Tangney, C.C., Wang, Y., Sacks, F.M., Bennett, D.A., & Aggarwal, N.T. (2015). MIND diet associated with reduced incidence of Alzheimer's disease. *Alzheimer's & Dementia*, 11, 1007-1014.
- Mungas, D., Heaton, R., Tulsky, D., Zelazo, P.D., Slotkin, J., Blitz, D., . . . Gershon, R. (2014). Factor structure, convergent validity, and discriminant validity of the NIH Toolbox Cognitive Health Battery (NIHTB-CHB) in adults. *Journal of the International Neuropsychological Society* 20, 579-587.
- Nasreddine, Z.S., Phillips, N.A., Bédrririan, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of American Geriatrics Society*, 53(4).
- National Institutes of Health. (2017). NIH Toolbox for assessment of neurological and behavioral function: Administrator's manual.
- Nelson, L., & Tabet, N. (2015). Slowing the progression of Alzheimer's disease; what works?. *Ageing research reviews*, 23, 193-209.
- Nganda, T., Lehtisalo, J., Solomon, A., Levalahti, E., Ahtiluoto, S., Antikainen, R., . . . Kivipelto, M. (2015). A 2-year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *Lancet*, 385, 2255-2263.
- Rosenburg, A., Ngandu, T., Rusanen, M., Antikainen, R., Backman, L., Havulinna, S., . . . Kivipelto, M. (2017). Multidomain lifestyle intervention benefits a large elderly population at risk for cognitive decline and dementia regardless of baseline characteristics: The FINGER trial. *Alzheimer's & Dementia*, 1-8.
- Smith, T., Gildeh, N., & Holmes, C. (2007). The montreal cognitive assessment: Validity and utility in a memory clinic setting. *Canadian Journal of Psychiatry*, 52(5), 329-332.

- Yang, L., Lam, F.M., Liao, L.R., Huang, M.Z., He, C.Q., & Pang, M.Y. (2017). Psychometric properties of dual-task balance and walking assessments for individuals with neurological conditions: A systematic review. *Gait & Posture*, 52, 110-123.
- Zola, S.M., Manzanares, C.M., Clopton, P., Lah, J.J., & Levey, A.I. (2012) A behavioral test predicts conversion to mild cognitive impairment and Alzheimer's disease. *American Journal of Alzheimer's Disease & Other Dementias*, 28(2), 179-184.