An Examination of the Effects of Mode of Access on the Computerized Revised Token Test

by

Laura E. Heilman

B.A. in Psychology, Saint Mary's College, 2006

Submitted to the Graduate Faculty of

Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

M.S. in Speech-Language Pathology

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

School of Health and Rehabilitation Sciences

This thesis was presented

by

Laura E. Heilman

It was defended on

July 18, 2008

and approved by

Malcolm McNeil, PhD, Distinguished Service Professor and Chair

Sheila Pratt, PhD, Associate Professor

Katya Hill, PhD, Associate Professor

Thesis Director: Malcolm McNeil, PhD, Distinguished Service Professor and Chair

An Examination of the Effects of Mode of Access on the *Computerized Revised Token Test*

Laura E. Heilman, M.S.

University of Pittsburgh, 2008

Background: The *Computerized Revised Token Test* (CRTT) was recently developed to improve the reliability and accessibility of the *Revised Token Test* (RTT). The CRTT was standardized using a touchscreen monitor; however, for various reasons, clinicians may need to use a mouse for test administration. In general, research suggests that younger individuals who are familiar with computers are more accurate and prefer to use a mouse. However, this may not be the case for brain-damaged persons with physical limitations. Thus, comparable performance when different input devices are used cannot be assumed.

Aims: The purpose was to investigate similarities and differences between participants' performance on subtest and overall scores obtained from touchscreen versus mouse on the CRTT. The study also examined the test-retest reliability of the CRTT when different input devices were used and user preference.

Methods & Procedures: Forty young, healthy adults participated in this study. All participants were native English speakers, and had no history of a speech, language, or learning disability. Participants passed a language screening, the *Story Retell Procedure* (SRP) (McNeil, Doyle, Park, Fossett, & Brodsky, 2002). Each participant took the CRTT with both modes of access, a mouse and a touchscreen, with their non-dominant hand. One-half (20) of the participants were administered both versions of the CRTT a second time. Additionally, all participants answered a preference questionnaire.

Outcomes & Results: The results revealed that touchscreen overall scores were significantly higher than mouse scores. There were also significant differences on six of the ten subtests. The test-retest reliability for both versions was equivalent and not significantly different. The results indicated a significant preference for the touchscreen.

Conclusions: While the touchscreen access method produced significantly higher subtest and overall CRTT scores than the mouse access method, along with equivalent reliability performance in this young normal participant population, it is not clear that it should be used as the preferred access method. If successful algorithms for equating the previously established psychometric data and normative sample derived from the touchscreen access method can be generated, then there will be no need to re-standardize the test.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Several people helped me complete this thesis and deserve a special acknowledgement and many thanks. Neil Szuminsky wrote the computer programs that were used and kept them running throughout the study; he also provided much needed technical advice. Miranda Babiak graciously donated many hours of her time helping me to test participants. Jee Eung Sung patiently helped me analyze the data. And last but certainly not least my husband, Mike, helped by providing much needed support and encouragement, acting as a sounding board for my ideas, helping with data analysis and interpretation, and proofreading several drafts of my paper as well.

1.0 INTRODUCTION

The *Revised Token Test* (RTT) is a diagnostic test used to evaluate the auditory processing abilities of individuals between the age of 5 through the lifespan who have developmental disorders or that have sustained brain damage. The RTT was designed to give the clinician information regarding a client's linguistic processing abilities particularly at the lexical-semantic levels of processing, but also at the syntactic level. It also allows the clinician to deduce information regarding a client's auditory attention, auditory memory, and temporal processing (McNeil & Prescott, 1978).

The RTT consists of 10 subtests that require following commands and indentifying and manipulating plastic objects of standardized shapes, colors and sizes. The objects are placed in front of a patient in specified locations. The clinician then presents an auditory stimulus in the form of an imperative sentence, such as "Touch the black circle" or "Put the big red square in front of the big white circle". The patient responds by touching or manipulating the objects. In six of the ten subtests the appropriate response requires the patient to touch an object and in four of the ten subtests the appropriate response requires the movement of an object (McNeil & Prescott, 1978).

The clinician scores the accuracy of the patient's response by using a 15-point multidimensional scoring system. Each linguistic element in each imperative sentence receives a score ranging from 1-15. For example, each of the ten commands in subtest I consists of three linguistic elements: a verb (direct command), an adjective (color number one), and a noun (shape number one). Thus, each command will receive three separate scores, each score ranges from 1 to 15 (McNeil & Prescott, 1978).

The score represents a description of how the task was performed. For example, a score of 15 means that the patient responded in an accurate and efficient manner and a score of 14 means that the patient vocally or subvocally rehearsed the command but completed the task without delay, which is scored a 13. At the other end of the scale, a score of 1 is described as the patient not responding and a score of 2 is described as the patient omitting the elements of one part of a two-part command, though a cue and a repeat must be given prior to scoring the response as a 1 or 2. The score is meant to be descriptive as well as hierarchical in terms of deficit severity. For example, a score of 9 means the patient needed the command to be repeated and a score of 8 means that the patient needed a cue following a repeat. A score of 9 represents the need for less information in order to perform the task than an 8, and thus represents less impairment on that particular linguistic element or item. See Appendix A for a description of all possible scores (McNeil & Prescott, 1978).

The individual score for each subtest takes into account the responses for each linguistic element as well as the overall responses for each command. The individual scores for each subtest are then used to determine the overall score. The overall test score is the average score of the 580 linguistic elements and the ten commands per subtest across the ten subtests. The results also can be used to develop profiles that further describe the patient's deficits. Thus, the test provides the clinician with substantive information about the client's auditory processing abilities and means to document small amounts of change, which can be useful when documenting treatment effectiveness or progression with additional disease or injury (McNeil & Prescott, 1978).

The complex scoring system can be difficult to use, however, and in order for the RTT to be scored accurately, the clinician must be very familiar with the test, which requires special training in order to meet reliability criteria. As a result, clinicians often prefer to use a simpler method of assessment (Odekar & Hallowell, 2005). In order to increase reliability and reduce training, a computerized version of the RTT has been developed, the *Computerized Revised Token Test* (CRTT).

The CRTT consists of the same ten subtests used in the RTT. The patient still follows commands and manipulates objects equivalent to the RTT with specified and standardized shapes, colors and sizes. However, instead of a clinician presenting the auditory commands from live voice, a computer presents the acoustically controlled commands, and instead of the patient manipulating actual objects, the patient manipulates images on a computer touchscreen monitor.

The objects are represented as images on the touchscreen monitor and are arranged according to the CRTT protocol, following that specified in the RTT (McNeil & Prescott, 1978). Patients manipulate the images by either touching an image on the touchscreen monitor or by touching and dragging an image across the monitor. The same 15-point multidimensional scoring system is used; however, the computer scores the patient's response online. The CRTT is therefore designed to standardize administration and scoring, providing a more reliable and potentially more valid assessment tool. The online capture of responses also provides a wealth of finite temporal information.

The CRTT also provides an *Efficiency Score* (ES), which is a new feature. An ES can be calculated for each individual command, for each subtest, or for the test as a whole. The ES is a measure of how long it takes the patient to respond relative to the score derived. The ES is calculated by multiplying the CRTT score and the length of time (t), in seconds, that it takes to complete the command divided by the maximum time (mt) allowed per command.

As stated above, response requirements of the CRTT involve manually moving images across the monitor. While, the images are representations of actual objects and persons with aphasia or brain damage may have difficulty with abstract concepts, the manual response using a touchscreen is concrete and does not require increased levels of abstraction compared to other methods of access such as the use of a mouse, a keyboard, or other pointing device. Therefore, to limit the amount of abstraction required of the patients, a touchscreen monitor was used in the development of the CRTT. Thus, the only level of abstraction required beyond the original RTT is realizing that the images represent three-dimensional objects.

A mouse is meant to represent a finger and an additional level of abstraction is required to use it. Since the purpose of the CRTT is designed to assess auditory processing and comprehension, it is advantageous to keep other aspects of the test as simple as possible. Because patients taking the CRTT may have difficulty with abstract concepts, the mouse may make the test more difficult, and shift the locus of the deficit from one of auditory language processing or simply add to the auditory processing load. Thus, a decision was made to standardize the test with the use of a touchscreen monitor.

While the use of a touchscreen appears to be a valid (McNeil, et al., 2008b), and reliable (McNeil, et al., 2008a), method of administering and scoring the test, there are limitations to its use. Though computers are almost universally available, touchscreen monitors are not. Many clinics do not have the financial means to purchase a touchscreen monitor. Also, clinicians may need to administer the test from a distance due to the ever increasing demand for telerehabilitation, and many patients will not have access to a touchscreen monitor. There also is the possibility that some individuals might feel more comfortable using one device over another. For example, children and young adults might be more comfortable and perhaps more proficient with a mouse then older individuals.

Telerehabilitation refers to providing services from a distance and is being investigated to determine if it is an effective means of diagnosing and treating individuals with speech and language disorders (Duffy, Werven, & Aronson, 1997; Georgeadis, Brennan, Barker, & Baron, 2004; & Mashima, Birkmire-Peters, Syms, Holtel, Burgess, & Peters, 2003). Telerehabilitation is becoming more common and valuable as individuals in remote locations require more specialized care (Buckwalter, Davis, Wakefield, Kienzle, & Murray, 2002). However, if the remote setting does not have access to a touchscreen monitor, clinicians might be dependent on a standard computer and mouse to administer the test or the treatment.

Clinicians might also elect to use a standard computer and mouse based on client preference and familiarity with computers. Computers are being used more often and earlier by individuals, and young children might prefer to use a mouse rather than a touchscreen. For example, Romeo, Edwards, McNamara, Walker, and Ziguras (2003) conducted a study to investigate the use of touchscreen monitors in early childhood education settings. Ninety-five children between the ages of three and seven served as participants. The children were students in early education classrooms.

The researchers first observed each of the five classrooms to note the use of computers in the classrooms, as well as to gain a sense of the classroom's learning environment. The researchers then observed the children during their free play time and kept an observation journal that documented the children's behavior, their interaction with the touchscreen, and the teacher's behaviors. The classroom teachers also kept an observational journal that documented the children's interactions with the touchscreen as well as reflective comments about their observations. Finally, the researchers conducted semi-structured interviews with the classroom teachers prior to and at the conclusion of the study. The researchers analyzed the data by extracting themes or categories. Three major themes were noted: developmental issues, input device preference, and social interaction and collaboration (Romeo et al., 2003).

Developmental issues referred to the way that the children interacted with the various input devices (mouse vs. touchscreen). Motor skills and positioning of the monitor were accounted for in this theme. The children were reported to have difficulty selecting and dragging an object across a monitor when a touchscreen was used. The more difficult the task became, the more trouble the students had completing the tasks. However, this same difficulty was not observed when a mouse was used. Possible explanations for this difference may be that the mouse is the most appropriate input device or that children are more familiar with and comfortable using a mouse. However, in all of the classrooms, the students had difficulty reaching the touchscreen monitor, which may have accounted for some of the difficulty the children faced when using a touchscreen (Romeo et al., 2003).

Another major theme derived from the data was input device preference. The task requirements and prior experience with computers were noted to have an impact on input device preference. When a task required the children to manipulate smaller objects or had a complex interface, children preferred to use a mouse. The children had been using a mouse throughout the school year, so when this study was conducted towards the end of the school year, the children may have already been more familiar with its use. Nonetheless, the data do indicate that the children preferred to use a mouse over a touchscreen and that the children tended to return to the more familiar input device, the mouse, when possible (Romeo et al. 2003).

The overall conclusions drawn from this study are that children have difficulty manipulating small objects via a touchscreen and that children prefer to use a mouse over a touchscreen, which is likely due to familiarity with a mouse. While the study was descriptive and cannot be used as strong evidence to support the use of a mouse over a touchscreen, a study by Wood et al. (2004) also examined the use of different input devices in early childhood education settings and found similar results.

Wood et al. (2004) investigated the use of input devices by 81 preschoolers ranging in age from 2:10 to 6:6, with a mean age of 4:5, and 43 educators ranging in age from 20 to 44, with a mean age of 29:9. Both the students and educators completed two games using four different input devices: a mouse, an EZ ball, a touch pad, and a touchscreen. One of the games required the participants to select and drag an icon and the other game required the participants to select, drag and release an icon on a different moving image.

Upon completion of the tasks, the researchers examined accuracy with each of the devices, the relationship between cognitive and motor skills with each of the devices, and user preference with each of the devices. Three different types of errors were recorded: icon drops (releasing icon before task completion), initial acquisition failures (unsuccessful initial attempt at acquiring the icon), and reacquisition failures (failing to require the icon after a drop). The results revealed that, for both the children and the educators, using the EZ ball and the mouse resulted in the fewest number of errors while using the touchscreen resulted in the highest number of errors. The mouse also was considered to be the most effective input device for both students and teachers (Wood et al., 2004).

Interestingly though, prior to the study the educators hypothesized that the touchscreen and EZ ball would be the most appropriate input devices for the students. Additionally, upon completing the tasks the teachers rated the mouse as more efficient and more accurate than the other input devices. While the children identified the EZ ball as the most efficient and accurate device, they preferred to use a mouse over the remaining two input devices (Wood et al., 2004). Similarly, the adults also preferred the mouse over the other input devices.

Muahmud and Kurniawan (2005) conducted a study to evaluate both the usefulness of psychometric testing when using different input devices (a mouse, a touchscreen, and a tablet-with-stylus) and the accuracy of these three devices in older individuals. The study included 12 participants who ranged in age from 53 to 75, with 63 being the average age. The participants were all highly educated, in good health, and experienced with using a computer. All the participants had used a computer for at least 2-3 years; however, their experience with the internet was limited, ranging from six months to a year. Prior experience with the three different input devices varied as well. All 12 participants had prior experience using a mouse, with an average of 8.69 years. However, the participants had no prior experience using a tablet-with-stylus and very little prior experience using a touschscreen, with the average number of years being 0.35 (Mahmud & Kurniawan, 2005).

During the study, the participants completed two tasks with each of the input devices. The two tasks, browsing and playing games, were chosen because older individuals often use the computer for these two purposes. Browsing required the participants to point to and click on links and performance was measured by the amount of time it took for the participant to get to the correct page. For the second task, the game solitaire was selected. To play the game participants needed to select, drag, and drop an image, or card. Performance on solitaire was measured by the score obtained after five minutes (Mahmud & Kurniawan, 2005).

The participants also completed non-performance related measures: a questionnaire, a debriefing interview, and psychometric tests. The questionnaire was used to gather general information about the participants, such as age, computer experience, and health. Upon completing the tasks, the participants rated the devices and tasks. The debriefing interview was used to ask participants open-ended questions regarding input devices in general and their experience with using the different input devices on each task. The psychometric tests measured cognitive abilities using the *Mini Mental State Exam* (MMSE), perceptual speed using the *Identical Picture* (IP) test, and motor speed using *Simple Reaction Time* (SRT) tests (Mahmud & Kurniawan, 2005).

The results for the performance related measures revealed that the participants performed significantly better on both tasks when using a mouse than when using a tablet or a touchscreen. Participants performed the worst on the browsing task when using a tablet and on the game task when using a touchscreen. On the non-performance related measures the results indicated that participants performed the best the on SRT test when using a mouse, followed by the tablet and then the touchscreen. Most participants also reported that the mouse was the easiest device to use and that the touchscreen was the most difficult device to use. The results also revealed that the IP test was the best predictor of the score for the browsing task when a mouse was used and for the game task when a tablet was used. When the touchscreen was used for the game task, prior experience was the best predictor of performance (Mahmud & Kurniawan, 2005).

Sears and Shneiderman (1991) also conducted a study to compare the performance and accuracy of three input devices: a stabilized touchscreen, a non-stabilized touchscreen and a mouse. The study consisted of three separate experiments with the first experiment designed to compare the two types of touchscreens with a mouse. The second experiment compared stabilized and non-stabilized touchscreens, and the third experiment compared a stabilized touchscreen to a mouse. The first experiment included 36 college- aged participants with varying amounts of computer experience. Three of the subjects had used a touchscreen once whereas the other thirty-three participants had no experience with a touchscreen. A majority of the participants reported that they used a mouse infrequently. The participants completed four tasks with each of the three devices. The four tasks required the participants to select target items of varying sizes: 1, 4, 16, and 32 pixels. The participants had to select six items for each task. The amount of time it took to select each target and the number of errors made per target was recorded. The participants were also asked which device they preferred.

The results indicated that when the target was 16 pixels, both the stabilized and the non-stabilized touchscreens were significantly faster than the mouse. However, when the target was 1 pixel, the mouse was both faster and more accurate than the touchscreens. When the target was 4 pixels the subjects made significantly more errors using the non-stabilized touchscreens. The participants reported that they preferred to use the mouse and that they preferred to use the stabilized touchscreen over the non-stabilized touchscreen (Sears & Shneiderman, 1991).

The second experiment, designed to compare the stabilized and the non-stabilized touchscreens, included twenty subjects with little to no experience using a touchsceen and a majority of the participants reported using a mouse infrequently. The subjects randomly selected 36 objects of varying sizes (4, 16, and 32 pixels) with each input device (stabilized and non-stabilized touchscreen). The order of presentation also was randomized, so some participants used the stabilized touchscreen first while others used the non-stabilized touchscreen first. As in the first study, the amount of time it took to select each target and the number of errors made per target was recorded. The participants also completed a survey that assessed their preference for using each device (Sears & Shneiderman, 1991).

Analysis of the results revealed that when the target was one pixel, the stabilized touchscreen was significantly faster and more accurate than the non-stabilized touchscreen. However, when the target was larger than one pixel, there were no significant differences in speed or accuracy between the two input devices. The participants also reported that they preferred to use the stabilized touchscreen (Sears & Shneiderman, 1991).

The third and final experiment was the same as the second experiment except that instead of comparing a stabilized touchscreen with a non-stabilized touchscreen, a stabilized touchscreen was compared to a mouse. Twenty individuals participated in the study. Of the twenty, eight had used a touchscreen once and a majority used a mouse infrequently (Sears & Shneiderman, 1991).

The results from this experiment revealed that a mouse was significantly faster and more accurate than the stabilized touchscreen when the target was one pixel. However, when the target was larger than one pixel, there were no significant differences in speed or accuracy between the two input devices. The participants preferred to use the mouse rather than the touchscreen. Thus, it seems that, unless the target was one pixel, mode of test administration did not affect the test scores (Sears & Shneiderman, 1991).

Mode of access also was examined when the *Useful Field of View* (UFOV) test was converted from a touchscreen with a chin rest to a personal computer (PC). Edwards et al. (2005) conducted a study that examined the test-retest reliability of the standard version, the mouse PC version, and the touch PC version. The study also compared the mouse PC version with the touch PC version. Edwards et al. first examined the reliability of the standard version of the test. Sixty-six adults over the age of 50 participated in the study. The mean age was approximately 72 years and education levels varied from sixth grade to Ph.D. The participants took the standard version twice, with an average of 92.58 days between each testing. The results indicated that the standard version of the UFOV test was reliable, with a reliability coefficient of 0.715.

A second and third experiment examined the reliability of the mouse PC version and the touch PC version. Sixty-six participants took the mouse PC version of the test twice, with a 10 day interval. One-hundred and fifty-eight participants took the touch PC version of the test with a 35 day interval. Results for both versions of the test indicated acceptable test-retest reliability with reliability coefficients of 0.884 and 0.735, respectively (Edwards et al, 2005).

A fourth experiment compared the mouse PC version with the touch PC version; although, the experiment also included the standard version of the UFOV test. A total of 364 participants over the age of 55 participated in this experiment. The mean age was 73 years, with the range of 55 to 93 years. The average level of education was 14 years. Participants completed either two or three versions of the test at one sitting, and the order of test administration was counterbalanced across participants. Some of the participants completed two versions of the test, the mouse PC and the touch PC, while other participants completed all three versions. The overall results of the study indicated that the test scores from all three versions were moderately to highly correlated. The percentage of participants who were in the impaired range was roughly equal across the three versions and that individuals who were considered to be borderline often fell into the "impaired" category on both PC versions (Edwards et al., 2005). Thus, similar o he Sears and Shneiderman (1991) study, the mode did not have a substantive impact on UFOV test results.

The emerging trend seems to be that the mouse is the more accurate and preferred input device in typical individuals, though there is also some evidence to suggest that input devices do not play a significant role in scores as well. However, the target population for this test is individuals who have suffered brain damage typically resulting in speech-language and physical disabilities. Therefore, these same trends may not have been seen in a disordered population. In fact, a study by Petheram (1988) found that a mouse was not the most accurate or preferred input device in persons suffering from stroke.

The goal of the Petheram (1988) study was to determine the best input device to use with an at-home computer-based therapy program aimed to supplement direct speech and language treatment in persons with left-hemisphere brain damage resulting in aphasia and physical limitations. In order to do this, the study examined the use of five different input devices on six different language tasks relevant to persons with aphasia. The five input devices were a mouse, a joystick, a touchpad, a tracker ball, and a touchscreen. The six different language tasks were represented by patterns on a screen. The patterns were presented in a similar fashion to the way that actual language activities would have been printed. Thus, the physical requirements necessary to complete both the experimental task and the desired language tasks were the same. Since persons with aphasia have difficulties with language, the tasks were designed to be cognitively simple so that the physical motor movements were not confounded by complex language tasks.

Twelve individuals between the age of 40 and 91 years participated in this study. Nine of the participants suffered a stroke and the remaining three participants served as controls in case the results required a comparison to a control group. Each of the participants completed every task for a total of 3600 tries, or 720 tries per input device. Success rate and time required to complete the task were measured. Success rate was measured by the percentage of correct responses. The study also examined input device preference (Petheram, 1988).

The results indicate that the mouse and the touchscreen were the least successful of the five input devices. Overall, the success rate for both the mouse and the touchscreen was slightly less than 70%, whereas the success rate was over 85% for the joystick, concept keyboard, and tracker ball. Participants completed the tasks most quickly, less than 25 seconds, when the touchscreen and the concept keyboard were used. On the other hand, it took the participants almost 100 seconds to complete the tasks when the mouse, joystick, or the tracker ball was used. The results also indicated that participants preferred the tracker ball, followed by the touch pad, and then the joystick. The mouse and the touchscreen were the two least preferred input devices, though participants did slightly prefer the touchscreen over the mouse (Petheram, 1988).

In general, results from the above studies indicate that either the mouse is the more accurate and preferred input device or that there is no difference in accuracy between the mouse and the touchscreen. However, the Petheram (1988) study did find different results in individuals with aphasia and physical limitations. As a result, comparable performance on tasks when different input devices are used cannot be assumed. Furthermore, the fact that a clinic may not have the financial means to purchase a touchscreen monitor in combination with the ever increasing need for telerehabilitation, there is additional evidence to suggest the need for the CRTT to be administered with clients selecting and manipulating the tokens via mouse instead of via touchscreen. However, since the CRTT was standardized with the use of a touchscreen, comparable performance cannot be assumed if a mouse were to be used instead of the finger for object selection and manipulation. Therefore, with the computerization of the RTT, it is necessary to determine the relationship between CRTT scores obtained when the different response modes (i.e., touchscreen vs. mouse) are used. As such, this study investigated similarities and differences between participants' performance on subtest and overall scores obtained from touchscreen versus mouse on the CRTT. The research questions this study sought to answer were:

1.) Is there a statistically significant ($p \leq .05$) difference in subtest and overall scores derived from the initial administration of the touchscreen and mouse access versions of the CRTT in normal, healthy, young adults?
2.) Are the preference judgments derived from initial administration of the touchscreen access version significantly ($p \leq .05$) higher than those derived from the initial administration of the mouse version?
3.) Is the difference between scores for the overall and subtest scores derived from the test-retest calculations for the touchscreen administration significantly ($p \leq .05$) different from the those same calculations derived from the mouse version?
4.) Are the correlation coefficients for the mean overall scores significantly ($p \leq .05$) higher for the test-retest scores derived from the touchscreen version than those derived from the mouse version?

2.0 METHODS

2.1 PARTICIPANTS

Forty-four participants were recruited from within the University of Pittsburgh community. The participants ages were between 18 and 30 years (*M* = 23.23 years, *SD* = 1.44 years). The participants reported English as their native language, earned a high school diploma or higher (87.5% reported a Bachelor's degree or higher), and were without a history of speech, language, hearing, or learning disorders, as determined by a self-report biographical survey (Appendix B). Participants also reported the same hand as dominant when using a mouse and a touchscreen. One participant reported Attention Deficit Disorder in the past medical history. Four participants were excluded from the study. One was a non-native English speaker; one was older than thirty; one used opposite hands to control a mouse and a touchscreen; and one was excluded due to equipment failure. The data from the remaining forty participants was used for this study.

2.2 MATERIALS

Each participant completed the informed consent form, the biographical survey, and a criterion-referenced language screening. The biographical survey consisted of a series of questions regarding demographic information such as age, self reported hand dominance, past medical history, and speech, language, hearing, and learning history. The participants completed the survey prior to administration of the CRTT. The language screening tool used was the *Story Retell Procedure* (SRP) (McNeil, Doyle, Park, Fossett, & Brodsky, 2002).

Two modes of access were used during administration of the CRTT in this study: mouse and touchscreen. A standard mouse (PC Concepts, Model # MUO2U) and a 17" ELOE touchscreen were used.

2.3 PROCEDURES

The participants completed the survey, the language screening, and both versions of the CRTT (mouse and touchscreen) during one session. The order of administration of the two versions of the CRTT was counterbalanced across participants.

The participants sat in a sound attenuated room during all CRTT data collection. The experimental acoustic stimuli were presented in the soundfield from two speakers, placed approximately 24 inches from the patient, at approximately 45 degrees azmuth. The acoustic stimuli were played through the audiometer to the two speakers and were presented at 75 dB SPL as indicated on the audiometer. The video monitor was calibrated prior to each session and was placed approximately 12 inches in front of each participant.

During administration of the SRP the participants sat in a quiet room. All participants listened to form A of the test and retold each story as accurately as possible immediately thereafter. The participant's responses were recorded with an omnidirectional microphone onto the SRP computer program. The experimenter counted the number of correct information units using standard scoring procedures (McNeil et al., 2002). The experimenter listened to the recordings and scored their responses. All participants received a passing score of 24.42 or higher (McNeil et al., 2002) on the SRP and were subsequently administered both versions of the CRTT. See Appendix C for SRP scores from each participant.

During administration of the CRTT the participants used their self-identified non-dominant hand to complete each task. The non-dominant hand was used to circumvent limb paresis or paralysis that is common in persons with aphasia, a population relevant to this tests eventual application. Thirty-nine of the forty participants reported right-hand dominance. One participant reported a general left-hand dominance but a right-hand dominance when using a mouse and a touchscreen; she used her left hand to control both input devices. The participants were instructed to place their non-dominant hand on the table prior to and upon completion of each test item for the touchscreen administration. Participants were instructed to keep their non-dominant hand on the mouse throughout the mouse access administration. None of the participants reported a problem with their non-dominant hand or wrist. One participant reported tendinitis in her right dominant wrist, which was not used for participation in this study.

Each participant successfully completed the CRTT pretest before completing each version of the CRTT. Successful completion of the pretest ensured that the participants' vision, hearing, and motor skills were adequate to complete the tasks and that they were able to process the linguistic stimuli in the experimental tasks at a level that fundamental perceptual, motor, and lexical knowledge and ability could be assumed.

The auditory stimuli were presented and the participants were instructed to respond to the command accordingly. The computer recorded and scored each response on-line. The CRTT program occasionally had to be restarted, at which time the program was set to begin on the test item on which it had failed. The experimenter was able to see the data on a second computer monitor that was not visible to the experimental participants. Upon completion of the first version of the test, the participants were given the option of a five minute break before beginning the second administration.

Upon completion of both versions of the CRTT, participants answered one question assessing their response mode preference (Appendix D). The participants indicated a preference for the mouse, the touchscreen, or no preference.

Between 4 and 11 days after completing the experiment, one-half (20) of the participants were administered both versions of the CRTT a second time. The order of administration of the two retest versions of the CRTT was again counterbalanced. The same procedures used during the initial test were used for the retest. However, participants did not have to complete another biographical survey, pass another language screening, or indicate their preference for mode of access.

2.4 DATA ANALYSIS

Data were analyzed to report results for the following conditions: 1) mode of access 2) access mode preference, and 3) test-retest reliability. A two-way repeated measures ANOVA was used to determine significant ($p \leq .05$) differences in subtest scores derived from the touchscreen and mouse access versions of the CRTT in these normal, healthy, young adults. Paired sample *t*-tests were used to locate the significant contrasts for any significant interactions. A dependent *t*-test was used to determine significant ($p \leq .05$) differences in overall scores between the touchscreen and mouse access versions. Pearson correlation coefficients were computed to determine if there was a high ($\geq$ .70) correlation coefficient for overall and subtest scores between mouse and touchscreen administered versions of the CRTT. A chi-square test was used to determine if there was a significant ($p \leq .05$) preference for either the mouse or touchscreen administered versions of the CRTT. The expected probability under the null hypothesis for each response was set at 1/3 since there were three possible responses.

Overall and subtest scores from those participants performing both access versions of the test administered a second time were compared to the group performance from the first administration. A three-way ANOVA was used to determine significant ($p \leq .05$) differences in scores obtained on the test and the retest, differences between mode of access, and differences between subtest and overall scores. Paired sample *t*-tests were used to follow up on significant interactions. Pearson correlation coefficients were computed to determine if there was a high ($\geq$ .70) correlation coefficient for test and retest overall and subtest scores between mouse and touchscreen administered versions of the CRTT. The Test of Significance of the Difference between Two Correlation Coefficients was used to determine a significant ($p \leq .05$) difference between the test and retest correlation coefficients for the mouse and touchscreen overall scores.

3.0 RESULTS

3.1 MODE OF ACCESS

Group means and standard deviations for both access versions, for subtest and overall scores, are presented in both Figure 1 and Table 1. Every subtest and the overall mean score for the mouse condition was lower than the mean for the touchscreen condition, except for subtest five. Furthermore, the standard deviation for the mouse was higher than the standard deviation for the touchscreen, except for subtests nine and ten.

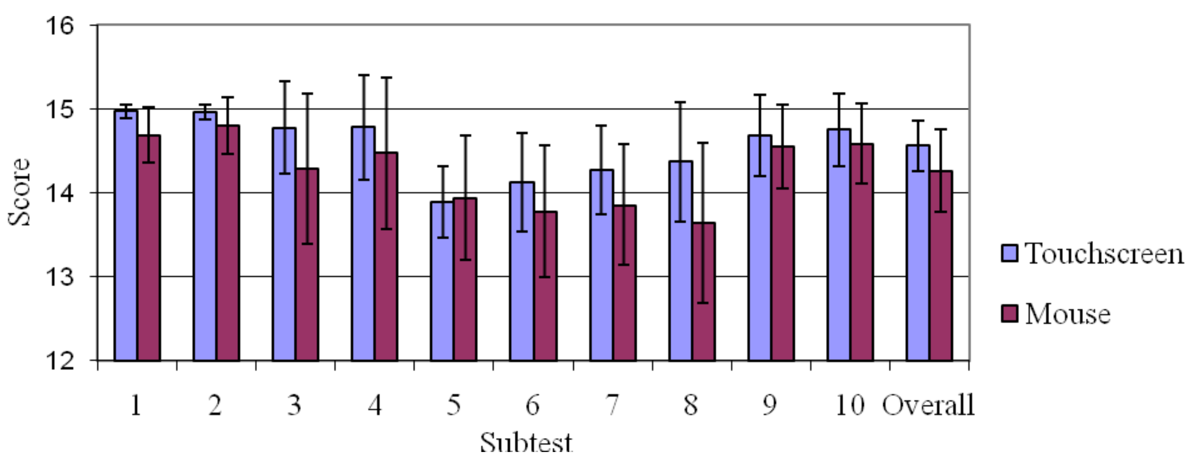

Figure 1. Mean subtest and overall scores on the CRTT for Touchscreen and Mouse access modes. Error bars indicate +/- 1 standard deviation.

Table 1. Subtest and Overall Means and Standard Deviations on the CRTT for the Touchscreen and Mouse Experimental Conditions

Variable	M	SD
Touch Subtest 1	14.97	.08
Touch Subtest 2	14.96	.08
Touch Subtest 3	14.77	.55
Touch Subtest 4	14.78	.62
Touch Subtest 5	13.89	.42
Touch Subtest 6	14.13	.59
Touch Subtest 7	14.68	.53
Touch Subtest 8	14.75	.71
Touch Subtest 9	14.68	.48
Touch Subtest 10	14.75	.43
Touch Overall	14.56	.30
Mouse Subtest 1	14.68	.33
Mouse Subtest 2	14.80	.34
Mouse Subtest 3	14.29	.89
Mouse Subtest 4	14.47	.90
Mouse Subtest 5	13.94	.74
Mouse Subtest 6	13.78	.79
Mouse Subtest 7	13.85	.72
Mouse Subtest 8	13.63	.98
Mouse Subtest 9	14.55	.50
Mouse Subtest 10	14.58	.48
Mouse Overall	14.26	.50

A two-way analysis of variance (ANOVA) was conducted to determine if there were statistically significant differences between subtest scores derived from the touchscreen and mouse access versions of the CRTT. The *alpha* was set at .05. The Wilks' lambda (Λ) value was used as the multivariate criterion value. A significant main effect for subtest (Λ=.13, $F(9, 31)=24.13$, $p<.01$) was found. There was also a significant main effect for access mode (Λ=.58, $F(1, 39)=27.86$, $p<.01$), with higher scores obtained with the touchscreen than with the mouse. In addition, the interaction between subtest and mode of access was significant $\Lambda=.39$, $F(9, 31)=5.4$, $p<.01$.

Pairwise *t*-test comparisons were computed to locate the source of the significant interaction between access mode and subtests. A Bonferroni-corrected alpha of .005 was used. This analysis revealed significantly higher scores on subtests one ($t(39) = 5.15$, $p < .005$), three ($t(39) = 4.29$, $p < .005$), four ($t(39) = 3.41$, $p < .005$), seven ($t(39) = 3.72$, $p < .005$), eight ($t(39) = 6.27$, $p < .005$), and ten ($t(39) = 3.09$, $p < .005$) when a touchscreen was used. See Table 2 for means, standard deviations, and confidence intervals.

Table 2. Average Differences, Means, Standard Deviations, and Confidence Intervals between Mouse and Touchscreen Subtest Scores on the CRTT in the Experimental Condition

			95% Confidence Interval of the Difference	
	Touchscreen-Mouse Mean Difference	SD of the Difference	Lower	Upper
Subtest 1*	.29	.35	.17	.40
Subtest 2	.16	.35	.05	.28
Subtest 3*	.49	.72	.26	.72
Subtest 4*	.31	.58	.13	.50
Subtest 5	-.05	.77	-.29	.20
Subtest 6	.35	.85	.08	.62
Subtest 7*	.42	.71	.19	.64
Subtest 8*	.73	.74	.50	.97
Subtest 9	.14	.51	-.02	.30
Subtest 10*	.17	.34	.06	.28

Note. * denotes a significant difference between mouse and touchscreen access modes ($p<.05$).

A dependent *t*-test was used to determine if there were significant ($p < .05$) differences in overall scores derived from the touchscreen and mouse access versions of the CRTT. The results indicate that the touchscreen access version overall mean score (*M* = 14.56, *SD* = .30) was significantly (t (39) = 5.28, $p < .01$) higher than the overall mean score from the mouse access version (*M* = 14.26, *SD* = .49). The 95% confidence interval for the overall mean difference between the two access versions was .19 to .42. A moderately high correlation coefficient between the overall scores from the touch access version and mouse access versions ($r = .684$, $p < .01$) was found. When the mouse was used participants received a score of twelve 14.85% of the time (594 of 4000 commands) whereas when the touchscreen was used participants received a score of twelve 4.65% of the time (186 of 4000 commands).

Pearson correlation coefficients were computed for overall and subtest scores between mouse and touchscreen administered versions of the CRTT. The results from the correlation analyses can be found in Tables 3-5. The average correlation coefficient between the mouse and touchscreen access modes was .390.

Table 3. Correlation of Touchscreen Subtest Scores on the CRTT in the Experimental Condition

Subtest	T 1	T 2	T 3	T 4	T 5	T 6	T 7	T 8	T 9	T 10
T 1	-	-.034	.025	.009	.398*	.553**	.311	.168	.013	.045
T 2		-	.237	-.086	-.135	-.020	.049	.134	.042	.073
T 3			-	.603**	-.107	.282	.222	.750**	.410**	.720**
T 4				-	.055	.211	.159	.624**	.117	.615**
T 5					-	.502**	.407**	.102	.107	-.021
T 6						-	.563**	.571**	.298	.328*
T 7							-	.576**	.210	.156
T 8								-	.325*	.605**
T 9									-	.626**
T 10										-

Note. * significant $p<.05$ level, ** significant $p<.01$ level. T = touchscreen.

Table 4. Correlation of Mouse Subtest Scores on the CRTT in the Experimental Condition

Subtest	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10
M 1	-	.123	-.233	-.198	-.131	-.133	-.230	-.078	-.130	-.257
M 2		-	.142	-.099	.303	.135	.302	.239	.321*	.112
M 3			-	.808**	.723**	.484**	.691**	.594**	.386*	.716**
M 4				-	.623**	.569**	.702**	.621**	.389*	.692**
M 5					-	.679**	.757**	.735**	.451**	.595**
M 6						-	.646**	.607**	.361*	.642**
M 7							-	.715**	.526**	.649**
M 8								-	.396*	.574**
M 9									-	.613**
M 10										-

Note. * significant *p*<.05 level, ** significant *p*<.01 level. M = mouse.

Table 5. Correlation of Mouse versus Touchscreen Subtest Scores on the CRTT in the Experimental Condition

Subtest	T 1	T 2	T 3	T 4	T 5	T 6	T 7	T 8	T 9	T 10
M 1	-.129	-.135	-.219	-.292	.076	-.123	.049	-.196	.058	-.250
M 2	-.121	-.042	-.032	-.145	.182	-.084	.158	.063	.282	-.024
M 3	.055	.152	.588**	.611**	.011	-.103	.291	.612**	.148	.561**
M 4	.107	-.006	.672**	.769**	.109	.342*	.299	.690**	.227	.689**
M 5	.032	-.033	.360*	.387*	.220	.255	.497**	.560**	.179	.366*
M 6	.081	.000	.294	.381*	.002	.271	.170	.426**	.194	.427**
M 7	.108	-.030	.398*	.553**	.142	.295	.388*	.587**	.278	.557**
M 8	-.128	-.133	.455**	.419**	.081	.262	.317*	.646**	.269	.463**
M 9	.196	.141	.187	.267	.184	.293	.399*	.419**	.464	. 314*
M 10	.095	.060	.528**	.655**	-.024	.247	.233	.682**	.446*	.720**

Note. * significant *p*<.05 level, ** significant *p*<.01 level. T = touchscreen, M = mouse.

3.2 ACCESS MODE PREFERENCE

A chi-square test was used to determine if there was a significant ($p < .05$) preference for access mode. The results indicate a significant preference for touchscreen ($\chi 2$ (2, N =40) = 31.40, $p < .001$). The number of participants who preferred the touchscreen, thirty, was significantly greater than the expected number of thirteen. Six participants preferred the mouse and four participants had no preference.

3.3 TEST-RETEST RELIABILITY

Means, standard deviations, and change scores of touchscreen and mouse test and retest administrations are presented in Table 6 and Figures 2 and 3. Every retest mean and subtest score was lower in the touchscreen condition with the exception of subtests one, five, and nine. In the mouse condition, the retest scores on subtests one, three, seven, and nine were higher than the test scores; the remaining subtests and the overall score were lower in the retest.

Table 6. Test and Retest Means, Standard Deviations, and Change Scores on the CRTT for the Touchscreen and Mouse Access Conditions

Test	M	SD	Retest	M	SD	Change Score
T Subtest 1	14.97	.07	T Subtest 1	14.97	.07	.00
T Subtest 2	14.97	.07	T Subtest 2	14.94	.12	-.03
T Subtest 3	14.84	.48	T Subtest 3	14.71	.70	-.13
T Subtest 4	14.69	.83	T Subtest 4	14.56	.88	-.13
T Subtest 5	13.91	.35	T Subtest 5	14.07	.55	.16
T Subtest 6	14.06	.59	T Subtest 6	14.00	.71	-.06
T Subtest 7	14.30	.38	T Subtest 7	14.26	.63	-.04
T Subtest 8	14.36	.76	T Subtest 8	14.22	.82	-.14
T Subtest 9	14.76	.37	T Subtest 9	14.76	.31	.00
T Subtest 10	14.73	.37	T Subtest 10	14.54	.67	-.19
T Overall	14.56	.27	T Overall	14.50	.37	-.06
M Subtest 1	14.69	.36	M Subtest 1	14.70	.44	.01
M Subtest 2	14.82	.38	M Subtest 2	14.82	.27	.00
M Subtest 3	14.24	.92	M Subtest 3	14.26	.96	.02
M Subtest 4	14.40	1.00	M Subtest 4	14.34	.87	-.06
M Subtest 5	13.92	.75	M Subtest 5	13.81	.68	-.11
M Subtest 6	13.77	.80	M Subtest 6	13.77	.79	.00
M Subtest 7	13.75	.84	M Subtest 7	13.83	.86	.08
M Subtest 8	13.61	.98	M Subtest 8	13.56	.99	-.05
M Subtest 9	14.55	.49	M Subtest 9	14.66	.47	.11
M Subtest 10	14.53	.54	M Subtest 10	14.43	.59	-.10
M Overall	14.23	.52	M Overall	14.22	.49	-.01

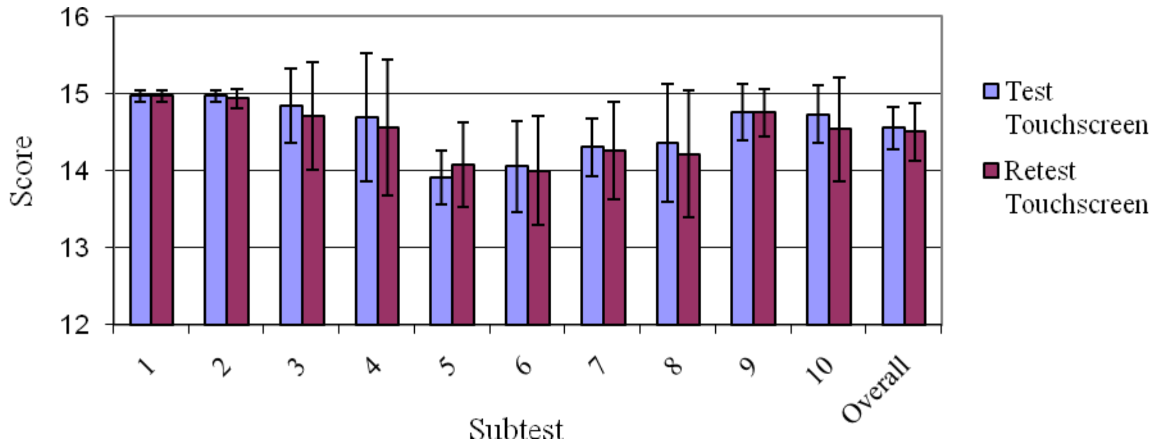

Figure 2. Mean test and retest, subtest and overall scores on the CRTT for the touchscreen access method. Error bars indicate +/- 1 standard deviation.

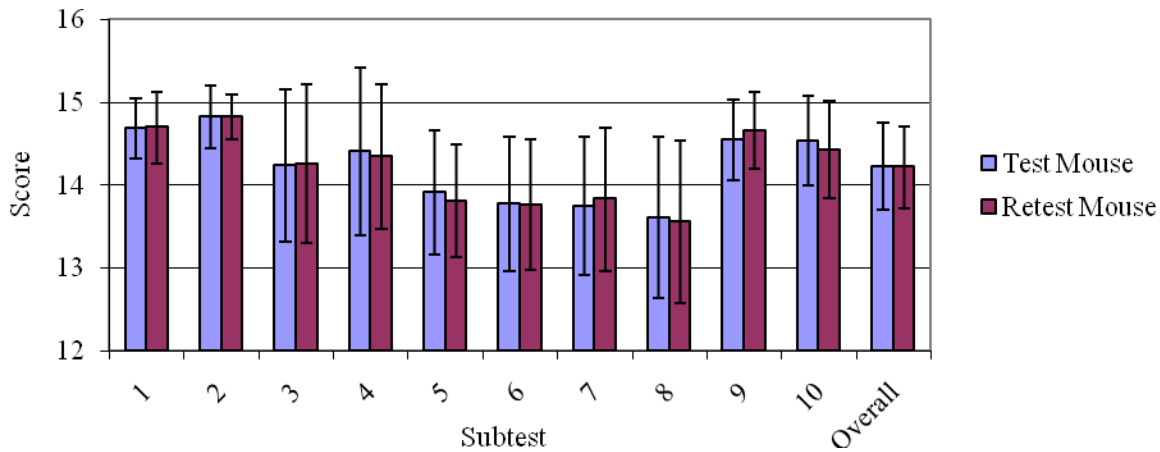

Figure 3. Mean test and retest, subtest and overall scores on the CRTT for the mouse access method. Error bars indicate +/- 1 standard deviation.

A three-way analysis of variance (ANOVA) was computed to determine differences in scores obtained on the test and the retest, differences between mode of access, and differences between subtest and overall scores. The *alpha* was set at .05. The Wilks' lambda (Λ) value was used as the multivariate criterion value.

Results from the ANOVA indicated a significant main effect of access mode (mouse vs. touchscreen), $F(1,19) = 23.03$, $p < .01$. Scores obtained on the touchscreen access mode were significantly higher than scores obtained when the mouse access mode was used. When the mouse was used in the retest condition participants obtained a score of twelve (immediacy) 19% of the time (380 of 2000 commands), and when the touchscreen was used participants received a score of twelve 9% of the time (173 of 2000 commands). A significant main effect for subtest ($F(9, 11) = 10.75$, $p <. 01$) was found. There was no main effect for test-retest ($F(1,19) = .36$, $p = .56$), indicating that there was no significant difference between test and retest scores for either the touchscreen or mouse access modes. Results also indicated that there were no significant interactions between test-retest and access mode ($F(1,19) = .35$, $p = .56$) or between test-retest and subtest ($F(9, 11) = 1.32$, $p = .33$). However, a significant interaction between access mode and subtest ($F(9,11) = 4.59$, $p < .05$) was present. There was no significant three-way interaction between test-retest, access mode, and subtest, $F(9, 11) = .51$, $p = .84$.

Test and retest conditions were averaged together since no significant main effect of test-retest scores was found. For example, the average score for subtest one on the test and the average score for the subtest one on the retest were averaged. Paired Sample *t*-tests were again computed to examine the significant interaction between access mode and subtest. A Bonferroni-corrected alpha of .005 was used. The results revealed significant differences between subtests one ($t(19) = 3.78$, $p < .005$), three ($t(19) = 4.57$, $p < .005$), seven ($t(19) = 3.31$, $p < .005$), and eight ($t(19) = 5.25$, $p < .005$).

Pearson correlation coefficients were computed for overall mouse and touchscreen subtest scores between test and retest. See tables 7 and 8 for correlation coefficients for mouse and touchscreen test and retest subtest scores. Figures 4 and 5 present a scatter plot depicting test and retest scores for both access versions.

Table 7. Correlation of Test versus Retest Subtest Scores on the CRTT in the Mouse Condition

Variable	Retest 1	Retest 2	Retest 3	Retest 4	Retest 5	Retest 6	Retest 7	Retest 8	Retest 9	Retest 10
Test 1	.261	.126	-.185	-.203	-.084	-.296	-.095	-.287	-.314	-.193
Test 2	.411	.531*	-.277	-.126	-.077	.220	.103	.167	.058	.166
Test 3	.413	.431	.559*	.631**	.302	.414	.319	.319	.546*	.565**
Test 4	.349	.146	.634**	.797**	.354	.226	.230	.251	.566**	.532*
Test 5	.476*	.516*	.465*	.512*	.360	.376	.345	.335	.301	.364
Test 6	.156	.251	.418	.614**	.454*	.485*	.336	.367	.436	.610**
Test 7	.327	.313	.421	.666**	.295	.532*	.526	.497*	.598**	.662**
Test 8	.303	.331	.507*	.685**	.491*	.476*	.429*	.529*	.320	.486*
Test 9	.433	.268	.015	.449*	.214	.392	.283	.361	.490*	.426
Test 10	.158	.114	.562**	.744**	.350	.604**	.331	.527*	.612**	.628**

Note. * significant $p<.05$ level, ** significant $p<.01$ level.

Table 8. Correlation of Test versus Retest Subtest Scores on the CRTT in the Touchscreen Condition

Variable	Retest 1	Retest 2	Retest 3	Retest 4	Retest 5	Retest 6	Retest 7	Retest 8	Retest 9	Retest 10
Test 1	.216	-.229	-.090	-.102	.396	.246	-.086	.415	.219	.095
Test 2	-.176	-.229	-.056	-.199	-.232	-.358	-.316	-.315	-.252	-.265
Test 3	-.101	.193	.818**	.683**	-.056	-.099	.714	.396	-.211	.424
Test 4	-.057	.101	.974**	.873**	-.062	-.051	.045	.232	-.016	.712**
Test 5	.321	.215	-.056	-.135	.216	-.033	-.138	-.126	.083	-.042
Test 6	.297	.197	.048	.072	.484*	.401	.281	.457*	.244	.181
Test 7	.539*	.295	.214	.332	.588**	.395	.538*	.665**	.033	.053
Test 8	.291	.330	.687**	.759**	.253	.353	.474*	.581**	.042	.490*
Test 9	.460*	.405	-.047	-.026	.196	.273	.235	.274	.301	.123
Test 10	.271	.087	.692**	.697**	.338	.243	.206	.397	.490*	.794**

Note. * significant $p<.05$ level, ** significant $p<.01$ level.

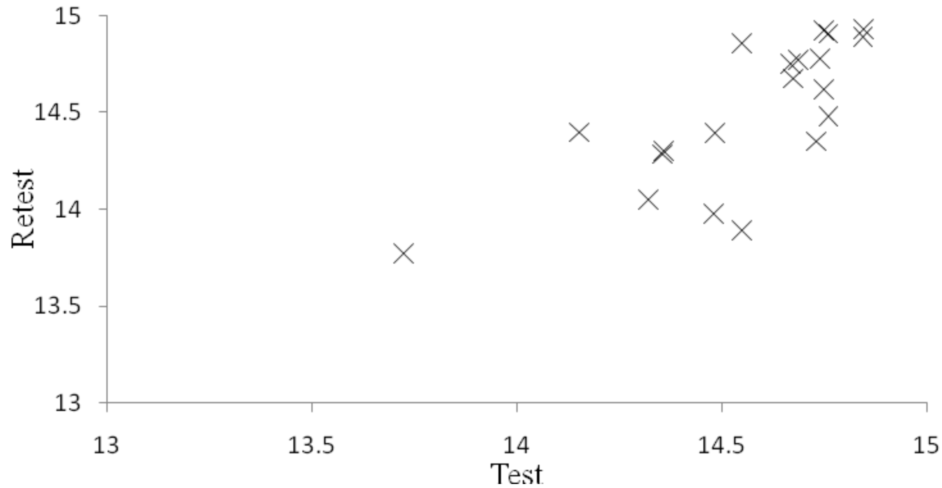

Figure 4. Scatter plot depicting test and retest scores on the CRTT for the touchscreen access version.

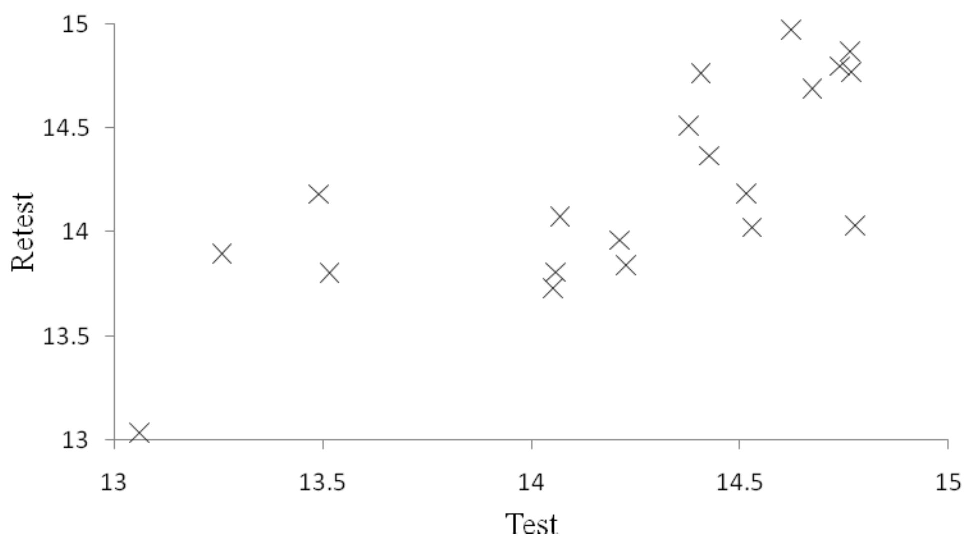

Figure 5. Scatter plot depicting test and retest scores on the CRTT for the mouse access version.

Table 9 summarizes the correlation coefficients for the overall scores for test and retest for both access conditions. The correlation coefficients indicate a moderate relationship between test and retest, regardless of the mode of access. Fifty-four percent of the variance ($R^2 = .54$) in the overall score was obtained between test and retest for touchscreen, and 55% ($R^2 = .55$) of the variance was obtained between tests for the overall mouse scores. Thus, both the touchscreen and the mouse access versions were about equally reliable from test to retest. Likewise, 53% (R = .73) of the variance was shared between access modes for the tests and 64% (R = .80) was shared between tests on the retest.

Table 9. Correlation of Test-Retest Overall Scores on the CRTT for both Access Conditions

Variable	Test Touch	Retest Touch	Test Mouse	Retest Mouse
Test Touch	-	**.736****	.730**	.581**
Retest Touch		-	.715**	.800**
Test Mouse			-	**.742****
Retest Mouse				-

Note. * significant $p<.05$ level, ** significant $p<.01$ level.

The Test of the Significance of the Difference between Two Correlation Coefficients was used to determine if the difference between the correlation coefficients was significant. Results indicated that there was no significant difference between the test and retest correlation coefficients (z = .04, $p = .968$).

4.0 DISCUSSION

Limited research exists comparing the effects of access mode on the scores obtained on computer-based language cognitive assessments, and even less research is available examining these effects in pathological populations. Thus, the purpose of this study was to compare the use of two access modes, a touchscreen and a mouse, on scores obtained on the *Computerized Revised Token Test* (CRTT); a test designed for use with pathological populations. The results revealed that one access mode, the touchscreen, produced significantly higher scores than the mouse access method for this test. Overall scores obtained when a touchscreen was used were .3 points higher than scores obtained when a mouse was used, resulting from significantly higher scores on six of the ten subtests. In addition, scores tended to be less variable when the touchscreen was used by about .25 for the test and .12 for the retest conditions.

The difference between the mouse access mode and the touchscreen access mode was likely due to the higher number of twelves (immediacy) scored when the mouse was used. Immediacy means that the participant responded too fast by initiating their response before the command was finished. This is a strategy that persons with aphasia or linguistic processing disorder may use in order to compensate for their deficits. For example, if the person is asked to follow a two-step command but is unable to do so, they may follow the first step as it is being told and then follow the second step, essentially breaking the two-step command into two one-step commands. However, in non-brain damaged individuals, a score of 12 does not necessarily mean that the person is compensating for deficits. The participants likely received a score of 12 because they wanted to finish the task as quickly as possible because they were bored or in a hurry; perhaps it was because they wanted to demonstrate that the task was not particularly challenging and that they were able to complete it rapidly. The participants were instructed to leave their hand on the mouse in between each item, which may have made it very easy for the participants to start responding before the command was over in the mouse version. Thus, this tendency to move too quickly most likely resulted in the lower scores obtained on the mouse.

Another possible explanation for the difference between mouse and touchscreen scores may have been due to the fact that when using the touchscreen participants occasionally dropped the token and would have to reacquire it before they could continue the move. The Wood et al. (2004) study found that the mouse was the better input device for this reason. However, this explanation does not seem likely for this study because touchscreen scores were actually higher than mouse scores. Participants did occasionally drop the token, but this did not appear to affect their score. The difference between the Wood et al. study and this study is that they were not measuring the same variables. The Wood et al. study was measuring the number of drops, initial acquisition failures, and reacquisition of an icon whereas this CRTT study was measuring the effect that different input devices have on the accuracy and timing of sentence comprehension. In other words, this study went one step beyond the Wood et al. study and examined the affect that dropping an icon, or token, would have on a complex cognitive process. Thus, it seems that the same conclusions cannot be drawn from these studies, and that differences found in the Wood et al. study are not as relevant to test scores obtained on the CRTT.

While further research is needed to determine the exact reason why the differences between mouse and touchscreen scores occurred, the data from the present study do indicate significant differences in performance between the two access methods. As a result, if the mouse is to be used as the input device, adjustments to the scores would need to be made in order for the mouse scores to be comparable to the touchscreen scores, the access method on which the test was standardized. Alternatively, the test will require re-standardization and re-norming using the mouse as the access device.

The study also examined participant preference for the access method and found that the touchscreen was the preferred input device, with three-fourths of the participants preferring the touchscreen. This was somewhat surprising because the majority of the previous studies had found that the mouse was the preferred input device (Romeo et al., 2003; Wood et al., 4004; Sears & Shneiderman, 1991). Even the Petheram (1988) study which found that the touchscreen was preferred over the mouse found that neither of these two input devices was preferred relative to other input devices. However, unlike the previous studies, the participants in this study completed the CRTT with their non-dominant hand, which likely contributed to the difference. The participants in this study were college students, all of whom use a computer and a mouse on a daily basis. However, this practice and familiarity with the mouse was accomplished with their dominant hand, and perhaps did not generalize to the non-dominant hand. Participants often made unsolicited comments that moving the images with the mouse was more difficult; however, they did not frequently report a difference in the ease of the task when required to touch a token. Thus, the difference may be that in the present study the participants had to use their non-dominant hand which they may have felt less comfortable using or were less skilled motorically, particularly when using the mouse.

This study also examined the test-retest reliability of the CRTT. Past research examining the test-retest reliability of a computer-based assessment found that the computer-based assessment was slightly more reliable than the standard version (Edwards et al., 2005). Similarly, both the RTT and the CRTT are also highly reliable in persons with aphasia (McNeil & Prescott, 1978; McNeil, et al., 2008ab). While there was no significant difference between test and retest for either access mode, both the mouse and the touchscreen versions of the CRTT have only moderately-high reliability correlation coefficients. Because normal participants tend to perform within a very narrow range, test-retest reliability as measured by correlation coefficients tends to be quite low (McNeil, et al., 2008ab); however, no significant differences were found between test and retest across any relevant measure. Scores obtained during the mouse access condition were slightly higher than the touchscreen access condition though.

The fact that the reliability of the CRTT was not as strong when given to normal individuals was not surprising because the test was designed to measure disordered populations and such tests tend to be less reliable when they are administered to normal populations. As mentioned previously, this is because the distribution of scores is smaller and correlations tend to be lower under such circumstances. That is the CRTT was not intended to distinguish among normal test takers, but rather it was designed to indentify persons with a disorder and to provide descriptive information about the nature of the disorder. As it stands, there is low variance in the scores from the present study which makes it more difficult to discern a pattern, or difference, between the variables. Nonetheless, the test-retest reliability was moderately-high given the population.

While the touchscreen access method produced significantly higher subtest and overall CRTT scores than the mouse access method, along with equivalent reliability performance in this young normal participant population, it is not clear that it should be used as the preferred access method. If successful algorithms for equating the previously established psychometric data and normative sample derived from the touchscreen access method can be generated, then there will be no need to re-standardize the test. If such an algorithm cannot be generated, it still may be worth the expense to develop new validity, reliability and normative data for a mouse version. Indeed, a mouse accessible CRTT would be more applicable to a variety of settings because they will be more readily available in clinics, laboratories and telerehabilitation settings than touchscreens.

The fact that the CRTT needs to be accessed and developed with non-dominant hand use because of the high incidence of dominant hand/arm hemiparesis could argue for the sole use of the touchscreen access method so as to minimize the confounds of motoric limitations on the assessment of language comprehension. Future research should examine the effects of hand dominance relative to mouse versus touchscreen access method on CRTT performance and preference. Until this is accomplished, it is difficult to recommend the use of one access method over the other in spite of the result that the mouse produced significantly lower overall scores on the CRTT than the touchscreen access method.

APPENDIX A

RTT/CRTT SCALE SCORE DESCRIPTIONS

Score	Description of Response
15	Complete
14	Rehearsal
13	Delay
12	Immediacy
11	Self-Correction
10	Reversal
9	Repeat
8	Cue
7	Error
6	Perseveration
5	Intelligible but incorrect response; rejection
4	Unintelligible (differentiated)
3	Unintelligible (perseverated)
2	Omission
1	No response

APPENDIX B

SUBJECT HISTORY

Subject #______________

Birth date: ______________ Age: _______________

Is English your native language? Yes No

If no, what is the primary language spoken in your home? _________________

Do you wear glasses? Yes No

Do you have difficulty hearing? Yes No

If yes, do you wear a hearing aid? Bilateral/ Right / Left / NA

Have you ever had any kind of speech, language or learning problem? Yes No

If yes, explain:

Did you ever have speech or language treatment? Yes No

If yes, explain:

Have you had any medical, psychological, or other conditions that might affect your ability to communicate or participate in the study (e.g., Stroke, Parkinson's disease, Alzheimer's disease, alcoholism, depression, etc.)? Yes No

If yes, explain:

Race: Caucasian African-American Asian Native-American Other

What is the highest level of education you completed? ____________________________

What is your occupation? (If retired, etc., indicate last occupation): _________________

Which is your dominant hand? Left Right

Which hand do you use a mouse with? Left Right

Which hand do you use a touchscreen with? Left Right

Do you have any problems with your hand or wrist (e.g., carpal tunnel syndrome, arthritis)? Yes No

If yes, what is the problem? ____

APPENDIX C

PARTICIPANTS' SCORES ON THE *STORY RETELL PROCEDURE*

Table 10. Participants' Scores on the SRP

Participant	SRP Score	Participant	SRP Score
1	53.67	21	43.42
2	49.78	22	60.03
3	29.41	23	40.11
4	35.46	24	43.87
5	56.65	25	47.08
6	48.48	26	71.82
7	35.92	27	56.86
8	44.29	28	51.63
9	49.00	29	48.42
10	43.89	30	50.92
11	45.38	31	40.03
12	48.91	32	52.91
13	44.38	33	46.42
14	52.51	34	44.52
15	63.97	35	48.36
16	27.92	36	44.28
17	46.91	37	35.21
18	56.47	38	35.47
19	52.39	39	43.73
20	55.36	40	41.56

APPENDIX D

PREFERENCE QUESTIONNAIRE

Please select the preferred test version:

a. touchscreen

b. mouse

c. no preference

REFERENCES

Buckwalter, K.C., Davis, L.L., Wakefield, B.J., Kienzle, M.G., & Murray, M.A. (2002). Telehealth for elders and their caregivers in rural communities. *Family and Community Health, 25(3),* 31-40.

Duffy J.R., Werven, G.W., & Aronson, A.E. (1997). Telemedicine and the diagnosis of speech and language disorders. *Mayo Clinic Proceedings,* 72, 1116-1122.

Edwards, J.D., Vance, D.E., Wadley, V.G., Cissell, G.M., Roenker, D.L., & Ball, K.K. (2005). Reliability and validity of useful field of view test scores as administered by personal computer. *Journal of Clinical and Experimental Neuropsychology, 27*, 529-543.

Georgeadis, A.C., Brennan, D.M., Barker, L.M., & Baron, C.R. (2004). Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders. *Aphasiology,18*, 639-652.

Mahmud, M. & Kurniawan, S.H. (2005). Involving psychometric tests for input device evaluation with older people. *Proceedings of OZCHI 2005, Canberra, Australia,* 1-10.

Mashima, P.A., Birkmire, Peters, D.P., Syms, M.J., Holtel, M.R., Burgess, L. P.A., & Peters, L.J. (2003). Telehealth: Voice therapy using telecommunications technology. *American Journal of Speech-Language Pathology, 12,* 432-439.

McNeil, M.R., Doyle, P.J., Park, G.H., Fossett, T.R.D., & Brodsky, M.B. (2002). Increasing the sensitivity of the story retell procedure for the discrimination of normal elderly subjects from person with aphasia. *Aphasiology, 16(8),* 815-822.

McNeil, M.R., Pratt, S., Sung, J.E, Szuminsky, N., Ventura, M., Kim, A., Fossett, T.R.D., Doyle, P.J., & Musson, N. (2008a, June). *Test-retest reliability of the auditory computerized revised token test (crtt) and three experimental reading (crtt-r) versions in normal elderly individuals and persons with aphasia.* Poster presented at the Clinical Aphasiology Conference, Jackson Hole, WY.

McNeil, M., & Prescott, T. (1978). *Revised Token Test*. Austin, ProEd.

McNeil, M.R., Sung, J.E., Pratt, S.R., Szuminsky, N., Kim, A., Ventura, M., Fassbinder, W., Fossett, T.R.D., Doyle, P.J., & Musson, N. (2008b, June). *Concurrent and construct validation of the computerized revised token test (crtt) and three experimental reading versions (crtt-r) in normal elderly individuals and person with aphasia.* Poster presented at the Clinical Aphasiology Conference, Jackson Hole, WY.

Odekar, A. & Hallowell, B. (2005). Comparison of alternatives to multidimensional scoring in the assessment of language comprehension in aphasia. *American Journal of Speech-Language Pathology*, 14, 337-345.

Petheram, Brain (1988). Enabling stroke victims to interact with a minicomputer - comparison of input devices. *International Disabilities Studies, 10(2),* 73-80.

Romeo, G., Edwards, S., McNamara, S., Walker, I. & Ziguras, C. (2003). Touching the screen: Issues related to the use of touchscreen technology in early childhood education. *British Journal of Educational Technology, 34(3),* 329-339.

Sears, A., & Shneiderman, B., (1991). High precision touchscreens: Design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies, 34(4)*, 593-613.

Wood, E., Willoughby, T., Schmidt, A., Porter, L., Specht, J., & Gilbert, J. (2004). Assessing the use of input devices for teachers and children in early childhood education programs. *Information Technology in Childhood Education Annual*, 261-280.