2010

Examining the Errors and Self-Corrections on the Stroop Test

Ashley K. Miller
Cleveland State University

EXAMINING THE ERRORS AND SELF-CORRECTIONS ON THE STROOP TEST

ASHLEY K. MILLER

Bachelor of Arts in Psychology

Miami University

May, 2008

Submitted in partial fulfillment of requirements for the degree

MASTER OF ARTS IN PSYCHOLOGY

at the

CLEVELAND STATE UNIVERSITY

May, 2010

This thesis has been approved

for the Department of Psychology

and the College of Graduate Studies by

Thesis Committee Chairperson, Dr. Amir Poreh, Ph.D.

Department & Date

Committee Member, Dr. Boaz Kahana, Ph.D.

Department & Date

Committee Member, Dr. Stephen Slane, Ph.D.

Department & Date

ACKNOWLEDGEMENTS

Thank you to my family and friends for their encouragement and support during this thesis, and especially to my mother who continues to be a source of inspiration to me.

EXAMINING THE ERRORS AND SELF-CORRECTIONS ON THE STROOP TEST

ASHLEY K. MILLER

ABSTRACT

The purpose of this study was to collect normative data for a computer-assisted version of the Comalli Stroop Test, a commonly used neuropsychological measure. Additionally, the study was aimed at investigating the self-corrected errors on the Stroop Test, which have not previously been accounted for on the traditional test versions. Participants included one hundred and seventy two individuals from Cleveland State University and the community. Participants were administered computer-assisted versions of the Comalli Stroop Test and Trail Making Test. Participants were also asked to rate their agreement to four statements on a 5-level Likert scale to assess self-perceptions of testing. Errors, self-corrected errors, and time of completion for both tasks were recorded. Answers to the Self-Monitoring Scale were scored and recorded. The results of this study show that age and education both affected the quantity and location of errors and self-corrected errors on the Stroop Test. The Trail Making Test, which was used to validate the errors on the Stroop Test, showed a similar pattern of location of errors to the Stroop Test. Errors were frequently made in the middle to later portions of these tests, whereas self-corrections were made in the earlier portions. This pattern is partially due to participants' limited cognitive and attention resources as the tests progress. The results of this study suggest that self-corrections are measuring a separate construct than errors on the Stroop Test. The ability to self-correct on the Stroop Test is a sign of mental health,

flexibility, and ability to self-monitor. Utilizing the self-corrected errors on the Stroop Test gives test administrators an additional tool in detecting control, and higher mental processes. Also, the results demonstrate that errors are measuring a separate construct than time of completion. The traditional approach to neuropsychological testing examines the total number of errors and time of completion for the entire task, rather than examining the critical parts of each task separately (the middle to latter portions). When only examining composite scores, significant increases in errors or time of completion from more difficult portions of the test are being averaged with better performance from the easier portions, yielding a score within normal limits. The results of this study support the process approach to neuropsychological testing.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER I

INTRODUCTION

The purpose of this study was to collect normative data and to assess a computer-assisted version of the Comalli Stroop Test. The Stroop Test consists of three separate conditions in which a participant reads color words printed in black ink, names the color ink that blocks of XXXXs are printed in, and names the color ink in which incongruent color words are printed (i.e. the word red printed in green ink). The Stroop Test is a commonly used neuropsychological measure, which is believed to measure selective attention, cognitive flexibility, processing speed, and more. It is routinely used in the evaluation of executive functions.

There are some differences between the Comalli et al. (1962) version and the computer-assisted version, mainly being that the practice items appear on separate pages. This is said to help improve the administration process. Additionally, self-corrected errors can be recorded and accounted for, whereas in the original version they could not. Furthermore, one can investigate for which word-color pairings there were more errors or self-corrections or took longer to complete.

The computer-assisted version represents a process approach to neuropsychological assessment, which differs from the traditional fixed approaches in the scoring (Poreh, 2006). The tests are not scored or administered in binary fashion (right or wrong). With this approach, qualitative aspects of behavior are quantified and used in the statistical analyses, such as self-corrections. In contrast, the original version only examined the number of errors and time of completion for each condition.

For these two reasons the present study is very important. While there are many versions of the Stroop Test, this will be the first computer-assisted version of the Comalli Stroop Test. Making available a computerized version of this widely used task, will help make for an easier administration and faster and more reliable scoring. Additionally, as stated above, researchers will be able to analyze the data qualitatively. Not only will the number of errors be analyzed, but it will be possible to see where participants are making the most errors, and how many are self-corrected. These qualitative data could be used in future research to diagnose certain illnesses. For example, how errors are distributed through the task may be diagnostic for people with Attention Deficit Hyperactivity Disorder (ADHD). We could potentially find that people with ADHD start the task fine, and make their errors later on, due to limited attention resources.

CHAPTER II

LITERATURE REVIEW

2.1 Statement of Problem

Today's society relies heavily on computers. They are an integral part of almost every business, school, interpersonal communication and more. It only makes sense that neuropsychological tests would follow suit and begin to rely on their usage as well. The current study is aimed at collecting normative data for a computer-assisted version of a commonly used neuropsychological test- the Comalli Stroop Task. While the original task has been shown to be both valid and reliable (Comalli et al., 1962), more data is needed for the computer-assisted version in order to claim that they are similar. Once found to be reliable and valid, the computer-assisted version has high potential to be used over the original version merely out of convenience as well as due to the qualitative approach in its nature. The computer-assisted version will allow researchers to look more into the process that participants are taking rather than just analyzing the end results. Looking more closely at the errors, self-corrected errors, and the quantity of occurrences will be able to give information related to attention, compulsitivity, and more.

2.2 History of the Stroop Test

Jaensch (1929) demonstrated that subjects, when presented with the name of a color printed in the ink of another color and were asked to name the ink color, read the name of the word instead; this was referred to as the interference effect (Jaensch, 1929). Jaensch's work did not receive much attention.

Later Stroop (1935) published *Observations on the interference phenomenon* in the Journal of Experimental Psychology. Stroop conducted several studies dealing with the interference effect. He found that color names printed in non-matching colored ink were not read as quickly as when they were printed in black ink. In a second study, he found that naming the color of square patches was accomplished much more quickly then naming the color of the ink of the non-matching color names. Following the publication of this article, many further studies were conducted on this phenomenon. The general method was referred to as the Stroop task and the interference effect as the Stroop effect.

2.3 Psychodynamics of the Stroop Test

Perret (1974) used the Stroop task to study patients with localized brain injuries. His work demonstrated that this task was an executive process, mediated by the left hemisphere frontal lobes (Perret, 1974). These results have been supported to show that patients with lateral prefrontal lobe lesions commit more errors on the Stroop Test than individuals from the non-clinical population. Because the task has heavy reliance on frontal lobe functions, the Stroop task is useful for studying executive processes, both typical and atypical.

The theory of parallel processing of relevant and irrelevant information can be used to explain the Stroop Effect. With this particular model, it is thought that processing occurs through activation moving along various pathways, each of different strength. It is thought that if two pathways are active simultaneously and produce conflicting activations, then facilitation to the stronger pathway is the result.

Golden (1975) suggested that the Stroop Test actually measures creativity, because it requires the participant to quickly and accurately devise new ways of handling and responding to novel situations. Golden assessed the creativity of 450 high school students (Matchstick Test, Improvements Test, or teacher ratings), and each student completed the Stroop Test. The results showed there was a positive correlation between performance on the Stroop incongruent condition and scores on the creativity measure. This was strongest and most significant for the teacher ratings condition ($r=.42$, $p<.001$).

2.4 Impacts on Performance on the Stroop Test

Similar to many neuropsychological tests, performance on the Stroop Test has been found to be related to numerous demographic factors, including age and education level (Seo et al., 2008). Most normative data on the Stroop Test comes from highly educated, young, and healthy individuals. One study employed 564 non-clinical individuals aged 60-90 years old to further study the performance of elderly and educationally diverse people on the Stroop Test. The results showed that a lower educational level and an advanced age were associated with lower Stroop performance (Seo et al., 2008). This illustrates that information processing speeds and executive function decline with age and this decline is slower in individuals with higher education. This finding supports the theory that individuals with higher education have a greater reserve capacity, which is based on more efficient utilization of brain networks or of ability to recruit alternate brain networks as needed, referred to as the cognitive reserve hypothesis (Stern, 2002; Nagandu et al., 2007).

Additionally, sex was significantly related to Stroop Test performance in this particular study. Women performed better than men in all three subtests. These results have been seen in earlier studies, which have found women to perform better at verbally based tests (Lee et al., 2004). This has been explained as being due to women having a greater facility in verbal reactions, and being more accustomed to responding to color stimuli than men.

2.5 The Golden Version

There are multiple versions of the Stroop Test that, for the most part vary slightly from one another. The Golden version of the Stroop Test also involves three subtests. In the first task, the participant is asked to read words (of color names) printed in black ink. The words appear in five columns of 20 words. The participant is asked to read as many words as possible in 45 seconds. In the second task, the participant is asked to name as many colors (red, green, blue) of blocks of XXXXs as possible in 45 seconds. In the third task, the participant is asked to name the color of the ink in which color words are printed. The same words and colors are used from the previous subtests; all word-ink pairings are incongruent (e.g. the word red printed in blue ink). The participant reads as many as possible in 45 seconds. The number of correct responses and errors are recorded for each subtest.

2.6 Comalli Version

Most of the Stroop Tests used by psychologists are derived from John Ridley's original Stroop task, also referred to as the Comalli et al. (1962) version. In this version, three white cards are used, each with 100 stimuli arranged in a 10 X 10 grid and an additional row of 10 practice items at the top. The first (word-reading) card is made up of color words that are printed in black ink. The second (color naming) card consists of rectangles of the same colors. The third task (interference) consists of color names printed in incongruent colored ink. In this task the colors were arranged in order to avoid any regularity of occurrence so that each color would only appear twice in each column and each row. The time needed to complete all 100 items and the number of errors made on each task (card) is recorded.

2.7 Aims of the Present Study

The present study was aimed at: 1. Collecting normative data for this method of computer-assisted analysis. 2. Examining the self-corrected errors on the Stroop Test, specifically, where they occurred and if they were distanced equally. Demographic variables such as age and education were also accounted for when examining the self-corrected errors. 3. Validating the self-corrected errors by correlating them with errors on another neuropsychological measure- the Trail Making Test.

CHAPTER III

METHOD

3.1 Measures and Hypotheses

3.1.1 Computer-Assisted Software

The computer-assisted versions of the Stroop Test and Trail Making Test were developed by Dr. Amir Poreh and Quantified Process Scoring Systems (QPSS Inc.). The software provides easy, real-time recording and scoring of the entire test process on a PC. Standard instructions are available at each stage of the task and were presented via the computer sound system in order to ensure standardization for all participants.

3.1.2 The Computer-Assisted Stroop Test

The computer-assisted version of the Stroop Test is based on the original Comalli version. There are some differences between the Comalli et al. (1962) version and the computer-assisted version. In the Comalli version practice items appear at the top of the page for each subtest, whereas in the computer-assisted version the practice items appear on separate pages. This is said to help improve the administration process. Additionally, self-corrected errors can be recorded and accounted for, whereas in the original version they could not. Furthermore, one can investigate which word-color pairings, if any, might cause more errors or self-corrections, or took longer to complete.

Individuals are provided with sheets of paper on which the items are printed. In each of the three conditions the participant is to read or name items printed in rows on a sheet of paper. In each condition, a one-row practice trial preceded a ten-row test. For the first task, color reading, the participant is to name colored blocks line by line, until he or she finishes the page. The participant is allowed as much time as needed for the task. For the second task, the participant reads color words printed in black ink. The third task is the incongruent condition; the participant is to read color words printed in incongruent ink. (i.e. the word blue printed in red ink).

While the participant read the colors or words, the examiner followed along on the computer screen, which displayed what was on the paper the participant held. The examiner recorded misses by using a mouse to click once on the item number, and recorded self-corrections by clicking twice on the item number. The examiner recorded total time for each line by clicking a button located at the end of each row.

Prediction for the Stroop Test

It is predicted that participants in this sample will perform much like the published data on a normative sample. This is primarily because we will not be collecting from a clinical population, and the majority of participants will be young, healthy, and well educated. It is predicted that young adults will make relatively fewer errors than older adults. Also, it is predicted that younger adults will self-correct more than older adults.

3.1.3 Correlation with Neuropsychological Measures

3.1.3.1 The Computer-Assisted Trail Making Test

The Trail Making Test (TMT) is a neuropsychological test of divided attention and executive functioning. Its current form, which consists of two parts, A and B, was first published as part of the Army Individual Test Battery (1944). In Part A, individuals are required to connect 25 numbered circles in numerical order that are spread across a sheet of paper. Part B is similar, but the sheet contains circles with numbers and letters. In this part, individuals must alternate between numbers and letters (i.e., 1-A-2-B-3-C, etc.). The score is derived from the difference in time of completion of each part.

If participants commit an error on the TMT, for example, by connecting 1 to 2 rather than A in Part B of the test, the examiner tells the participant they made an error, stops him or her, and has him or her return to 1 or to the last item they connected correctly, and to connect it to the correct item in sequence.

In the present study, the participants complete the TMT with a paper and pencil, while the examiner follows along on a computer screen identical to the participant's paper. The cursor automatically starts on item 1 and once the examiner clicks the mouse button, the cursor moves to the next test item in sequence (i.e., item A for Trails B). In this way, it is easier for the examiner to follow along with the participant's responses. If the participant goes out of sequence the examiners only needs to manually move the cursor with the mouse and click on the same item number. The computer will then say "you skipped a circle" and the examiner will also tell the participant to stop and return to the last test item. There is a button at the end of the task that the examiner clicks to stop the clock and record the time of completion. The participant also completes a sample for both Trails A and Trail B, which consists of only a few items in the sequence. Once the participant completes the sample and demonstrates they understand the task, they are permitted to begin the test items.

Prediction for the Trail Making Test

It is predicted that participants in this sample will perform much like the published data on a normative sample. This is primarily because we will not be collecting from a clinical population, and the majority of participants will be young, healthy, and well educated. It is believed that the location of errors on the TMT will significantly correlate with errors on the Stroop Test. Furthermore, it is predicted that the location of these errors will be similar for both tasks, in the middle to later portions of the test, as mental resources as becoming taxed.

3.1.4 Self-Monitoring Scale

A self-monitoring scale was developed in order to assess participants' self-perception of performance at the end of testing after completion of the Stroop Test and

Trail Making Test. A brief questionnaire included four statements, and asked for the participants to specify their level of agreement for each. A 5-level Likert scale was used on which 1 indicated “strongly disagree”, 2 indicated “disagree”, 3 indicated “neither agree nor disagree”, 4 indicated “agree”, and 5 indicated “strongly agree”. The statements were: “I did well on the tasks”, “I did better on the tasks than most people my age”, “I made fewer errors than a typical person would make”, and “The time it took me to complete the tasks was less than one would typically expect”. Responses were summed and could range from four to 20. A high score indicates that a participant felt he or she did well or better than average on the two tasks. A low score, such as 4, indicates that a participant felt he or she did not perform well or did worse than average on the tasks.

Prediction for the Self-Monitoring Scale

It is hypothesized that participants who make more self-corrections on the Stroop incongruent condition, make more errors on the Trail Making Test Parts A and B, or have a longer time of completion for the two tasks will score lower on the Self-Monitoring Scale, indicating that they are rating their performance as below average. It is hypothesized that errors on the Stroop Test are made unknowingly, or else participants would self-correct, and thus does not negatively impact self-perception of performance.

3.2 Participants

Participants included 172 individuals from the greater Cleveland area. The average age was 34.5 years (SD=14), ranging from 18 to 78 years old. There were 108 females and 64 males in the sample. The majority was right handed. The mean years of education was 14.69 years (SD=2). Some of the participants signed up as part of an extra credit opportunity for a class; otherwise no compensation was received for participation in the study.

The participants filled out an informed consent form prior to taking part in the study. A copy was kept for the examiner's records and an additional copy was provided to the participant so they would be provided with contact information for the key investigator (Ashley Miller). Prior to testing, each participant was to provide information about his or her age, sex, hand preference, education level, and whether English was his or her first language. Any person whose first language was not English was ineligible for the Stroop Test. There were approximately seven participants who fell into this category. In addition, it was stressed to the participants that this study aimed to collect normative data and any personal history of neurological illnesses, such as Alzheimer's, Parkinson's, dementia, traumatic brain injury, etc. would make him or her ineligible for the study.

3.3 Procedure

Every participant was given the same instructions and test battery, with the exception of the Self-Monitoring Scale. Every participant first filled out the informed consent and was asked demographic questions. Next, he or she completed the Stroop Test, and then the Trail Making Test; lastly, a subset of all participants completed the Self-Monitoring Scale. The Self-Monitoring Scale was developed and decided to be included in the study once data collection had begun. Entire time of administration for the informed consent, two measures, and questionnaire was approximately ten minutes per participant.

3.4 Data Analysis

Descriptive statistics were conducted on the Self-Monitoring Scale, and indices of the Trail Making Test Part B and Stroop Test incongruent condition; specifically, time of completion, number of errors, and self-corrected errors on the Stroop incongruent condition. Pearson’s R was used to assess the association between participants’ scores on the Self-Monitoring Scale, and indices of the Trail Making Test Part B and Stroop Test incongruent condition. Scatterplots were used to examine participant age against time of completion for the Trail Making Test Part B, and participant age against time of completion for the Stroop incongruent condition. Descriptive statistics were calculated for mean years of education, time of completion, number of errors, and number of self-corrections on the Stroop Test for the younger age group (18-45 years) and the older age group (46-80 years). Pearson’s R was used to assess the association between age and the number of errors, as well as self-corrections on the Stroop incongruent condition by line number. Line graphs were used to examine quantity of self-corrections and errors on the Stroop Test incongruent condition by line number and age group. A stepwise regression analysis was used to determine which section from the Trail Making Test Part B was the best predictor of number of errors on the Stroop incongruent condition. A stepwise regression analysis was used to determine which section from the Trail Making Test Part B was the best predictor of number of errors on the Trail Making Test Part B.

CHAPTER IV

RESULTS

Table I shows descriptive statistics which were calculated for indices of the Trail Making Test Part B, Stroop Test incongruent condition, and the Self-Monitoring Scale. The mean time of completion for TMT Part B was 49.08 seconds, and the mean time of completion for the Stroop incongruent condition was 98.61 seconds. There was an average of .2 errors on the TMT Part B and .93 errors on the Stroop incongruent condition. The mean self-corrections for the Stroop incongruent condition was 1.39. The mean score on the Self-Monitoring Scale was 14.6. Once again, a higher score (closer to 20) indicated the participant rated his or her performance as better than average. The minimum score was 6 and maximum score was 20. 93 out of the 172 participants completed the Self-Monitoring Scale.

Table I.

Descriptive Statistics on Indices of the Trail Making Test Part B, the Stroop Incongruent Condition, and the Self-Monitoring Scale

Index	Minimum	Maximum	Mean	Std. Deviation
Time of Completion TMT Part B (s)	22	126	49.08	18.986
TMT Part B Errors	0	4	.20	.618
Time of Completion Stroop Incongruent (s)	35.1	160.5	98.613	22.4645
Stroop Incongruent Errors	0	21	.93	2.441
Stroop Incongruent Self-Corrections	0	7	1.39	1.464
Self-Monitoring Score	6	20	14.60	2.655

Table II shows Pearson's correlations between age, education, score on the Self-Monitoring Scale, and indices of the Stroop Test and Trail Making Test, specifically, time of completion, errors, and self-corrections. Significant correlations were found between age of participant and number of errors on TMT Part B ($r=.182, p<.05$), age of participant and time of completion for TMT Part B ($r=.377, p<.01$), age of participant and the Stroop incongruent condition time of completion ($r=.413, p<.01$), and age of participant and number of errors on the Stroop incongruent condition ($r=.268, p<.01$). Significant correlations were found between education level of the participant and time of completion for the Stroop incongruent condition ($r=-.2$, $p<.01$), and education level and number of errors on the Stroop incongruent condition ($r=-.187$, $p<.05$). Significant correlations were found between the number of errors made on TMT Part B and time of completion of TMT Part B ($r=.451$, $p<.001$), and the number of errors on TMT Part B and the number of errors on the Stroop incongruent condition ($r=.220$, $p<.05$). Significant correlations were found between time of completion for the Stroop incongruent condition and number of errors on the Stroop incongruent condition ($r=.224$, $p<.001$), and time of completion for the Stroop incongruent condition and time of completion for the TMT Part B ($r=.300$, $p=.001$). A significant correlation was found between score on the Self-Monitoring Scale and participant education level ($r=.235$, $p<.05$).

Table II.

Pearson's Correlations between Demographic Information and Neuropsychological Indices

	Self-Monitoring Score	Age	Education	Time Stroop Inter.	Errors Stroop Inter.	Self-Corrections Stroop Inter.	Errors TMT Part B	Time TMT Part B
Self-Monitoring Score	1	-.070	.235(*)	-.198	-.145	-.099	-.040	-.176
Age	-.070	1	.312(**)	.413(**)	.268(**)	-.025	.182(*)	.377(**)
Education	.235(*)	.312(**)	1	-.200(**)	-.187(*)	-.021	.038	-.060
Time Stroop Inter.	-.198	.413(**)	-.200(**)	1	.224(**)	.106	.091	.300(**)
Errors Stroop Inter.	-.145	.268(**)	-.187(*)	.224(**)	1	.030	.220(*)	.225(*)
Self-Corrections Stroop Inter.	-.099	-.025	-.021	.106	.030	1	-.091	-.084
Errors TMT Part B	-.040	.182(*)	.038	.091	.220(*)	-.091	1	.451(**)
Time TMT Part B	-.176	.377(**)	-.060	.300(**)	.225(*)	-.084	.451(**)	1

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Figure 1 shows a scatterplot of participant age against time of completion for the Trail Making Test Part B. Figure 2 shows a scatterplot of participant age against time of completion for the Stroop incongruent condition. They show similar patterns, as participant age increases the time of completion for the Trail Making Test Part B and Stroop incongruent condition increase.

Figure 1.

Effect of Age of Participant on Time of Completion for the TMT Part B

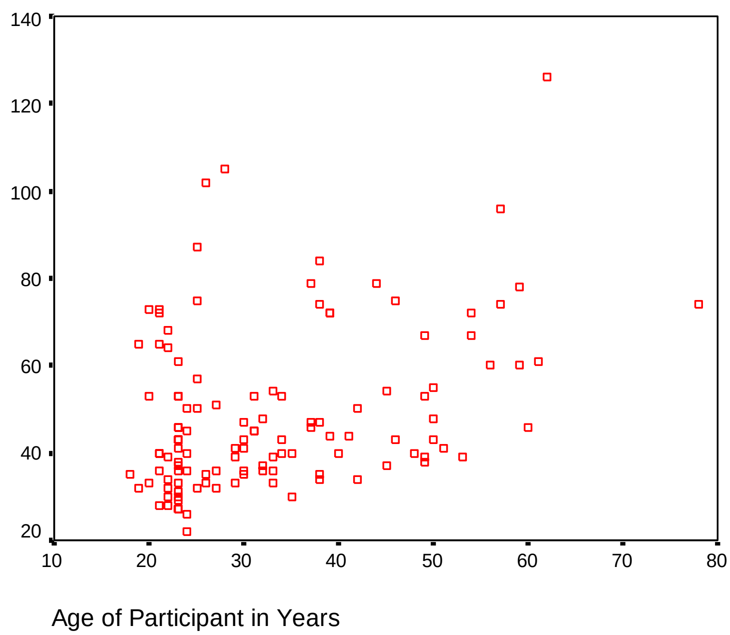
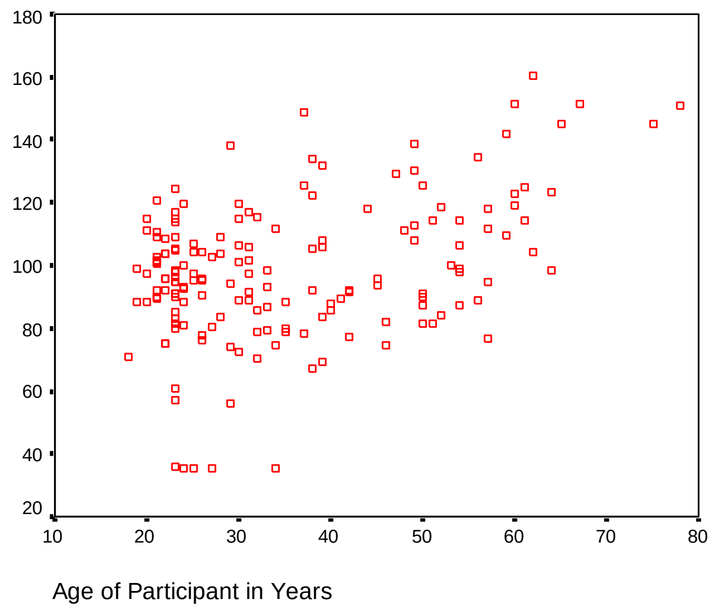


Figure 2.

Effect of Age of Participant on Time of Completion for the Stroop Incongruent Condition



Descriptive statistics were calculated for the two age groups on the mean time of completion, number of errors, and self-corrections for the Stroop incongruent condition. Table III shows the older age group (46-80 years) had a mean time of completion of 112.56 seconds for the Stroop incongruent condition, whereas the younger age group (18-45 years) had a mean time of 93.62 seconds (Table IV). The older age group had a mean of 1.5 errors for the incongruent condition, whereas the younger age group had a mean or .72 errors. In addition, the older age group had a mean of 1.23 self-corrections, while the younger age group had a mean of 1.45.

Table III.

Descriptive Statistics for Older Adults (46-80 years) on Indices of the Stroop Incongruent Condition

	Minimum	Maximum	Mean	Std. Deviation
Time of Completion (s)	74.8	160.5	112.566	22.9110
Number of Errors	0	21	1.50	3.909
Number of Self-Corrections	0	7	1.23	1.523

Table IV.

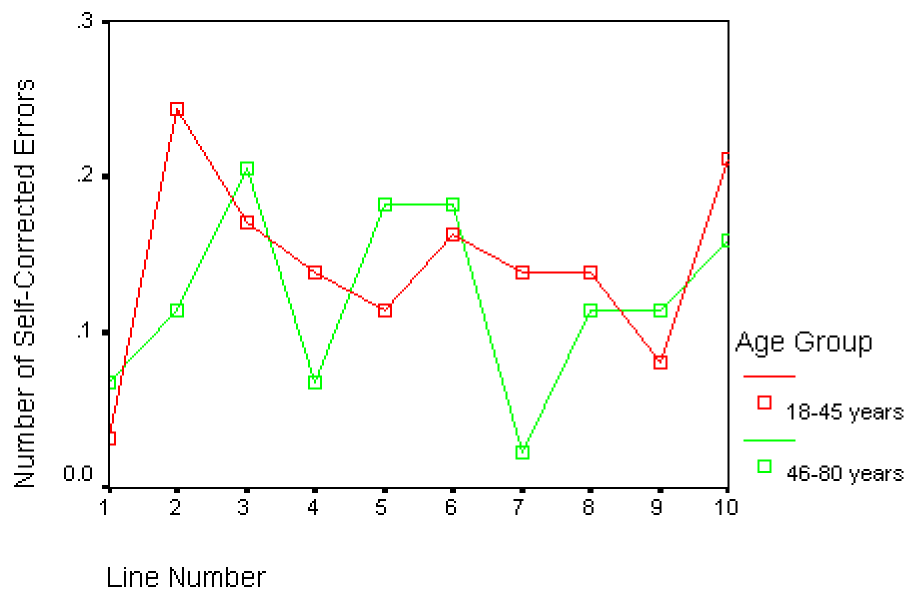
Descriptive Statistics for Younger Adults (18-45 years) on Indices of the Stroop Incongruent Condition

	Minimum	Maximum	Mean	Std. Deviation
Time of Completion (s)	35.1	148.5	93.622	20.1573
Number of Errors	0	11	.72	1.601
Number of Self-Corrections	0	7	1.45	1.444

Age was found to be significantly correlated with the location of self-corrected errors on the Stroop incongruent condition. Specifically, it was found that age of the participant was correlated with the number of self-corrected errors in line two ($r=-.163$, $p<.05$). Figure 3 shows the pattern of self-corrected errors per line broken down by age group. The younger age group (18-45 years) had more self-corrected errors at the beginning of the task and relatively few until line ten. The older age group (46-80 years) continued to make self-corrected errors throughout the task with no improvement.

Figure 3.

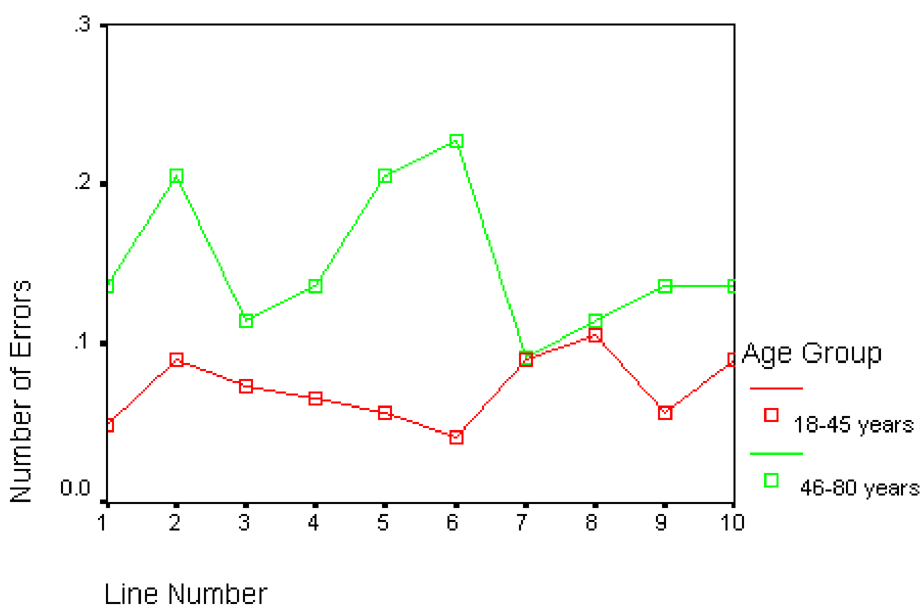
Total Self-Corrected Errors on the Stroop Incongruent Condition by Age Group and Line



Age was found to significantly correlate with the location of errors on the Stroop incongruent condition. Specifically, it was found that age of the participant was correlated with number of errors in line one ($r=.202$, $p<.01$), number or errors in line two ($r=.231$, $p<.01$), number of errors in line three ($r=.197$, $p<.05$), number of errors in line four ($r=.210$, $p<.01$), number of errors in line five ($r=.217$, $p<.01$), and number of errors in line six (r=.346, p<.01). Figure 4 shows the pattern of errors per line broken down by age group. The older age group made more errors throughout the task than the younger age group. Additionally, the older age group made errors consistently throughout the test, while the younger age group showed a peak at the beginning of the test (line 2) and end of the test (line 8).

Figure 4.

Total Errors on the Stroop Incongruent Condition by Age Group and Line



Descriptive statistics were calculated for mean years of education for each age group. The younger age group (18-45 years) had an average of 15.15 years of education

(SD=1.833). The older age group (46-80 years) had an average of 13.41 years of education (SD=1.945).

Table V shows the results of a stepwise regression analysis, which revealed that the second to last section on the Trail Making Test Part B (items 16-20) was the best predictor of number of errors on the Stroop Test incongruent condition (F=10.773, p=.001). Table VI shows the results of a stepwise regression analysis, which revealed that the second to last section on the Trail Making Test Part B (items 16-20) was also the best predictor of number of errors on the Trail Making Test Part B (F=29.356, p=.001).

Table V.

Stepwise Regression of Number of Errors on the Stroop Test onto Section Four of the Trail Making Test Part B

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.287(a)	.082	.075	2.037

a Predictors: (Constant), TMT PB #4 (items16-20)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.703	1	44.703	10.773	.001(a)
	Residual	497.928	120	4.149		
	Total	542.631	121			

a Predictors: (Constant), TMT PB # 4 (items 16-20)
b Dependent Variable: Number of errors on the Stroop Test incongruent condition

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.271	.391		-.694	.489
	PB 16-20	.110	.034	.287	3.282	.001

a Dependent Variable: Number of errors on the Stroop Test incongruent condition

Table VI.

Stepwise Regression of Number of Errors on the Trail Making Test Part B onto Section Four of the Trail Making Test Part B

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.436(a)	.190	.184	.558
2	.501(b)	.251	.239	.539

a Predictors: (Constant), TMT PB #4 (items 16-20)
b Predictors: (Constant), TMT PB #4 (items 16-20), TMT PB #5 (items 21-25)

ANOVA(c)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.144	1	9.144	29.356	.000(a)
	Residual	38.935	125	.311		
	Total	48.079	126			
2	Regression	12.063	2	6.031	20.765	.000(b)
	Residual	36.016	124	.290		
	Total	48.079	126			

a Predictors: (Constant), TMT PB #4 (items 16-20)
b Predictors: (Constant), TMT PB #4 (items16-200, TMT PB #5 (items 21-25)
c Dependent Variable: Number of errors on the TMT PB

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.313	.106		-2.942	.004
	PB 16-20	.049	.009	.436	5.418	.000
2	(Constant)	-.419	.108		-3.878	.000
	PB 16-20	.035	.010	.312	3.585	.000
	PB 21-25	.025	.008	.276	3.170	.002

a Dependent Variable: Number of errors on the TMT PB

CHAPTER V

DISCUSSION

The mean time of completion for the Stroop Test was 98.61 seconds, whereas the Trail Making Test Part B had a mean time of completion of close to half the time, 49.08 seconds. The TMT Part B, on average had fewer errors than the Stroop Test incongruent condition, which was close to one per participant. Additionally, the mean number of self-corrections was 1.39. The maximum errors on the Stroop Test incongruent condition was 21, whereas it was only 4 on the TMT Part B. The maximum number of self-corrections on the Stroop incongruent condition was 7.

This information could mean that participants found the Stroop Test to be more difficult and taxing than the Trail Making Test, based on time of completion, number of errors, and self-corrections. More research would be needed to determine if this is the case.

The mean score on the Self-Monitoring Scale was 14.6 out of 20. The minimum score was six and the maximum score was 20. This indicates that most participants thought they did well, or at least better than average on the tasks. It was unexpected that not a single participant felt his or her performance was poor enough to yield a rating of four on the scale.

Education level was significantly correlated with the time of completion ($p<.01$) and the number of errors ($p<.05$) on the Stroop incongruent condition. As participants' education level increased the amount of time for completion and number of errors on the Stroop incongruent condition both decreased. This gives support to the cognitive reserve hypothesis, which states that individuals with higher education have a greater reserve capacity. Age was significantly correlated with number of errors on Part B TMT ($p<.05$), Part B time of completion ($p<.01$), incongruent condition time of completion ($p<.01$), and number of errors on the incongruent condition ($p<.01$). As participant age increases the number of errors made on Part B TMT and the Stroop incongruent condition both increase. Additionally, as participant age increases the time of completion for the TMT Part B and Stroop incongruent condition both increase.

As participants made more errors on Part B TMT or the Stroop incongruent condition, time of completion for these tasks increased ($p<.01$). The number of errors made on Part B TMT was significantly correlated with the number of errors made on the Stroop incongruent condition ($p<.05$). As participants made more errors on the TMT Part B, the number of errors on the Stroop incongruent condition was also increased.

As participants' education level increased the score on the self-monitoring scale increased ($p<.05$). This indicates that the more educated an individual is, the higher he or she rates his or her performance on the two tasks, regardless of a shorter time of completion or fewer errors. However, it should be noted that only 93 out of the 172 participants completed the Self-Monitoring Scale. Perhaps if more participants had completed the questionnaire there would be additional significant correlations.

The results showed that participant age predicted the location of errors and self-corrected errors on the Stroop incongruent condition. The older age group continues to make self-corrected errors throughout the task, whereas the younger age group makes a few at the beginning of the task and then relatively few until line ten. Additionally, the older age group made more errors, less self-corrections, and had a longer time of completion than the younger age group.

Furthermore, it was shown that the younger age group and older age group did not drastically differ in terms of education level. The younger age group had a mean of approximately 15 years of education, whereas the older age group had a mean of approximately 13.5 years of education. This illustrates that age is the variable responsible for influencing the quantity and location of errors and self-corrections.

Section four of the Trail Making Test Part B (items 16-20) was the best predictor of number of errors on the Trail Making Test Part B and number of errors on the Stroop Test incongruent condition. This shows that as participants complete the TMT Part B it gets more challenging as it progresses, and requires more cognitive resources. The reason section four is correlated with more errors, and considered more complicated is because there are still plenty of circles left to connect and these items are further along in the number sequence and alphabet. Also, the pattern requires the participant to look both backward and forward at responses, whereas the last few items in section five (items 21-25) only require the participant to look forward.

The pattern we see for the location of errors on the Trial Making Test Part B is similar to what is seen for the Stroop Test incongruent condition. Both tasks show an increase in errors towards the later portions, which could be caused by fatigue, limited attention resources, and inability to ignore the interfering stimuli. Additionally, this shows that as the tasks progress they become more challenging, requiring more cognitive resources. This illustrates how similar the required resources are for both tasks. The higher incidence of self-corrections at the beginning of the Stroop incongruent condition, with a steady decline until line ten shows the steady fatigue of participants and depletion of cognitive resources; as much more attention is needed to catch a mistake and self-correct. This is further supported by the finding that older adults made fewer self-corrections than younger adults. Self-correcting may be a sign of good mental health. Utilizing the self-corrected errors on the Stroop Test gives test administrators an additional tool in detecting control, and higher mental processes.

These results demonstrate that errors are measuring a separate construct than time of completion. The traditional approach to neuropsychological testing examines the total number of errors and time of completion for the entire task, rather than examining the critical parts of each task separately (the middle to latter portions). When only examining composite scores, significant increases in errors or time of completion from more difficult portions of the test are being averaged with better performance from the easier portions. This can often yield a score within normal limits when it is not truly deserved. The results of this study support the process approach to neuropsychological testing, where tests are not scored or administered in binary fashion, but rather qualitative aspects of behavior are quantified and used in the statistical analyses.

Additionally, the results of this study suggest that the number of errors made on either the Trail Making Part B or the Stroop Test incongruent condition measure a separate construct than the self-corrected errors on these tasks. Errors may occur unknowingly, whereas self-corrections demonstrate the participant's awareness of an error and enough mental flexibility to self-correct. This gives support that self-corrections are a sign of good mental health and higher mental processes.

This finding is compatible with recent studies, which show that patients with particular circumscribed frontal damage exhibit an increase in self-monitoring errors while patients with damage to other frontal regions do not exhibit this phenomenon. The belief is that the ability to perform the incongruent condition successfully requires consistent activation of the intended response mode, which is the role of the superior medial frontal region (Stuss et al., 2001).

A limitation to this study was the relatively low sample size of 172 participants. In order to accurately assess the validity and reliability of any new measure, as compared to a more traditional approach, a much larger sample size would be needed. Additionally, the majority of the participants were young, female, right-handed, and well educated. A much more diverse sample with greater variability in age and education level is needed to accurately compare measures. An increase in older adults may also make the difference in errors and self-corrections as compared to younger adults on the Stroop incongruent condition much more pronounced. Lastly, although best efforts were made to exclude any individuals from a clinical population with any psychiatric history, some could have made their way into the sample. Perhaps, a more thorough questionnaire, or evaluation, should be given prior to the test administration.

As stated earlier, a much larger sample size would be needed in order to assess the validity and reliability of the computer-assisted versions of the Stroop Test and Trail Making Test as compared with the traditional forms. After this, data collection from a clinical population would be reasonable.

For example, errors and self-corrections could be examined for people with Attention Deficit Hyperactivity Disorder (ADHD). Because it is hypothesized that self-corrections require mental flexibility and attention, we would expect individuals from this population to make fewer self-corrections and more errors on the Stroop Test. Additionally, it would be interesting to further examine the location of these errors or self-corrections. As the test progresses and becomes more mentally taxing, I expect these individuals to make more and more errors, due to limited attention resources. It would also be interesting to include patients with circumscribed frontal damage in future research. Previous studies have found these individuals to exhibit an increase in self-monitoring errors. In this way, we can validate that the increase in self-monitoring errors is truly a function of the superior medial frontal region.

REFERENCES

Army Individual Test Battery. (1944). *Manual of Directions and Scoring*. Washington, DC: War Department, Adjutant General's Office.

Comalli, P.E., Jr., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test in childhood, adulthood and aging. *Journal of Genetic Psychology*. 100, 47-53.

Golden, C.J. (1975). The measurement of creativity by the Stroop color and word test. *Journal of Personality Assessment*, 39, 502-506.

Groth-Marnat, G (Ed.). (2000). *Neuropsychological Assessment in Clinical Practice*. Canada: John Wiley & Sons, Inc.

Jaensch, E.R. (1929). Grundformen menschlichen Seins. Berlin: Otto Elsner.

Lee, D.Y., Lee, K.U., Lee, J.H., et al. (2004). A normative study of the CERAD neuropsychological assessment battery in the Korean elderly. *Journal of International Neuropsychological Society*, 10, 72-81.

Nagandu, T., von Strauss, E., Helkala, E.L. (2007). Education and dementia: what lies behind the association? *Neurology*, 69, 1442-1450.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behavior. *Neuropsychologia 12*: 323-330

Poreh, A. (Ed.). (2006). *The Quantified Process Approach to Neuropsychological Assessment*. New York, NY: Taylor & Francis.

Seo, E.H., Lee, D.Y., Choo, I.H., Kim, S.G., Kim, K.W., youn, J.C., Jhoo, J.H. & Woo, J.I. (2008). Normative study of the Stroop color and word test in an educationally diverse elderly population. *International Journal of Geriatric Psychiatry*, 23, 1020-1027)

Stern Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of International Neuropsychology Society*, 8, 440-460.

Stroop, J. R. (1932). Is the judgment of the group better than that of the average member of the group? *Journal of Experimental Psychology, 15,* 550-562.

Stuss, D.T., Floden, D., Alexander, M.P., Levine, B. & Katz, D. (2001). Stroop performance in focal lesion patients: dissociation of process and frontal lobe lesion location. *Neuropsychologia, 39*, 771-786.