

Working Memory Capacity and Familiarity with Testing Environment

Yu Yi Lu¹, Jonathan Jones², and Rachel Booth³

¹Leland High School, San Jose, CA, USA

²University of Cambridge, Department of Clinical Neurosciences, Cambridge, UK

³Ulm University, Germany

ABSTRACT

Previous studies have shown converging evidence that negative perceptions of the surrounding environment lead to lower standardized test performance among stigmatized individuals. However, there has been minimal research done about the underlying cognitive mechanism that may account for these effects. I hypothesized that unfamiliarity with the surrounding environment interferes with test performance because it limits individuals' working memory capacity. This within-subjects experiment, with a total of 35 Leland High School students, tested that hypothesis. A self-produced version of the working memory span task was given to all participants in both of their prospective classrooms, familiar and unfamiliar. Through a matched-paired t-test analysis, the results demonstrated that unfamiliarity with the surrounding environment significantly limited one's working memory capacity. Implications for future studies are discussed.

Introduction

According to the College Board, approximately 10 million high school students took the SAT (Scholar Aptitude Test) or ACT (American College Testing) related assessments during the 2017-2018 school year (College Board, 2019). Recent research has been dedicated in discussing the fair and objective standards, as well as the unjust and discriminatory flaws, of the standardized testing system. The current system distributes students across a variety of testing centers. Unfamiliarity with the surrounding environment, then, becomes a potential factor in students' test scores. In my research, I will examine working memory capacity, one's ability for short-term maintenance and manipulation of information, as an underlying cognitive mechanism that may explain why familiarity with the surrounding environment could affect performance on standardized tests.

Literature Review

Perception of the Surrounding Environment and Standardized Testing Performance

Throughout the years, educators have discovered that environment, emotion, and cognition are all factors that highly contribute to student learning (Brown & Harris, 2009; Van Gennip, Segers, & Tillema, 2009). Students' attitudes and perceptions of their learning environment dramatically change how they engage in the learning process. When students possess positive attitudes and perceive their classroom environment as safe and valuable, their achievements are expected to improve (Chu, Babenko, Cui & Leighton, 2014).

For the past few decades, mandatory standardized tests have been administered to measure students' achievement "with significant consequences for students" (Chu et al., 2014). From secondary exit examinations, like SAT, to

Medical College Admission Test (MCAT), students are categorized and graded based on how well they perform on such tests. At times, students are prevented from securing placements at prestigious programs and/or universities due to their results. Increasingly, students' standardized test scores are also used to compare student achievements and school systems both nationally and internationally. This poses significant consequences on "educational policy, post-secondary acceptance rates, and school closures" (Ravitch, 2010).

Determining the aspects that enable or hinder students' academic achievement on these standardized tests is vital (Aldridge & Templeton, 1999). In a 2014 study, Professors Chu et al. examined the correlation between students' standardized test performance and their perceptions of learning environments. Researchers adapted a 4-point Likert scale survey to evaluate how students perceive their learning environments. The Programme for International Student Assessment (PISA) is then used to measure students' academic achievement. The survey and the PISA were given to 5,233 students from 165 schools in the United States and 23,000 students from 978 schools in Canada. Through sequential hierarchical linear modeling of the data, Chu et al. clearly concluded that how students perceive their surroundings explained a significant, yet modest, 2.86% of the total difference in students' standardized test scores. This provides a rationale for further investigation on the environmental variables that contribute to students' standardized test performance. However, instead of finding the correlational evidence for the relationship, my experiment studies the cause and effect of the scenario. It also highlights the importance of researching potential mechanisms, such as working memory, that lead to such effects.

Working Memory Definition

In 1690, the idea of short-lived memory was first coined by philosopher John Locke. He proposed that only a limited amount of knowledge can be stored at once (Cowan, Rouder, Blime & Scott, 2012). This view has since then been developed, reflecting the modern conceptualization of working memory most renownedly presented by Baddeley and Hitch in 1974. These two researchers built on the theory of short-term memory, the part of the memory that retains information for immediate use. They presented a new multi-component system that stores and processes together. Together, Baddeley and Hitch (1974) termed "working memory," developing a model with three distinctive, yet interactive, cognitive functions. These are the phonological loop, the visuospatial sketchpad, and the central executive processor. Baddeley and Hitch (1974) defined the phonological loop as a short-term store of spoken and written material, the visuospatial sketchpad as the internal eye that stores visual and spatial information, and central executive as a system that controls attentional processes. In the early 2000s, Engle's definition of working memory centered entirely on the central executive process. He explained it as one that "focuses attention on preserving temporarily activated information of interest and inhibiting distractions that are irrelevant to the tasks at hand" (Engle, 2002). Ever since, modern theories of working memory have drifted to include attentional capability (central executive) in addition to the temporary storage of information (short-term memory).

For my research, I have adopted Nelson Cowan's attention-control working memory definition. Cowan defined it as "the use of attention to preserve information about goals and sub-goals for ongoing processing and to inhibit distractions from those goals" (Cowan, 2016). Individuals with higher working memory capacity can suppress task-irrelevant details and focus on task-essential information for ongoing processing (Simon, Tusch, Holcomb & Daffner, 2016).

Working Memory Capacity and Cognitive Tasks

Another study by Engle, Tuholski, Laughlin, and Conway (1999) explored relationships between working memory capacity and real-world cognitive tasks, such as complex learning, language comprehension, writing, reasoning, and other executive functions. Complex cognitive span tasks measure individual differences in working memory. These tasks were designed "to engage multiple aspects of working memory and impose interleaved processing demands that limit the use of mnemonic strategies" (Mathy, Chekaf & Cowan, 2018). Working memory capacity was distinctly

positively correlated (correlation coefficient of $r = 0.59$) with fluid intelligence—“one’s ability to think logically and solve new problems in novel situations without reference to previously acquired knowledge” (Thorsen, Gustafsson & Cliffordson, 2014). A high correlation was also found between students’ working memory capacity and their verbal ($r = 0.40$) and qualitative ($r = 0.43$) SAT scores. In this paper, working memory is shown to be moderately involved in complex cognitive tasks and standardized tests.

Underlying Cognitive Mechanism

Prior research has shown converging evidence that students’ perceptions of their surrounding environment negatively affects their standardized test performance. However, little is known about its underlying cognitive mechanisms. Therefore, the focus of my research centers around assessing working memory capacity as a potential cognitive mechanism that contributes to differences in students’ standardized test results due to unfamiliarity with the testing environment. As there is a wide range of elements that can alter students’ perceptions of their surrounding environment, I have narrowed my topic specifically to examining students’ familiarity with the surrounding environment as the factor. Since most standardized tests are provided during school hours in a classroom setting, and most of my subjects will be participating in my experiment at my high school, it would be most logical for me to define surrounding environments as classrooms on a school campus. The ultimate goal of my research is to examine whether students’ familiarity with their surrounding environment affects their working memory capacity.

Working Memory Span Tasks

The Daneman and Carpenter’s version of the reading span task is used to measure working memory capacity. In a methodological review on working memory span tasks, professors evaluate the positives and the negatives of the task, from the reliability and validity of the assessment to optimal administration and scoring procedures. Conway et al. (2005) provides strong evidence that working memory span tasks are accurate predictions of one’s cognition. In statistical analysis, Cronbach’s alpha coefficients are calculated as an estimate of the reliability of a psychometric test. Three verbal span tasks—counting, operation, and reading—are shown to be constructively valid and reliable with high Cronbach’s alpha coefficients of 0.78, 0.80, and 0.77, respectively (Conway et al., 2005). Proven to be extremely useful in cognitive psychology research, I will be applying one of these specific working memory span tasks (reading) in my experiment. More detailed descriptions of the specific procedural of the reading span task will be discussed in the methodology section.

Hypothesis

Working Memory Capacity and Testing Environment

Previous research has not directly examined how environmental manipulation can temporarily limit one’s working memory capacity. However, some evidence suggests that high levels of anxiety might correlate with reduced levels of working memory capacity (Eysenck & Calvo, 1992), and that variations with the surrounding environment can increase or decrease one’s stress level (Zatz & Chassin, 1985). Individuals, who report more real-life stress and score higher on anxiety trait, perform worse on working memory capacity measures than their less-stressed counterparts (Derakshan & Eysenck, 1998; Klein & Boals, 2001). In a 2001 paper, Klein and Boals claimed that individuals under high stress levels utilize a great amount of their “mental resources” to suppress those negative and unwanted thoughts, instead of focusing on the tasks at hand. Stress-related thoughts occupy valuable cognitive resources, preventing stressed individuals from operating at their optimal working memory capacity. In another paper, Taylor (2004) examined the roles of the classroom environment on students’ increasing or decreasing anxiety levels. Qualitative data

analyses confirmed that familiarity with classroom environments can lower students' levels of anxiety and increase positive attitudes during learning and testing (Taylor, 2004). With these previously correlational relationships established, it is expected that the environmental unfamiliarity can temporarily lower students' working memory capacity. It will then be the goal of my experiment to test such a hypothesis. As previous studies are mostly correlational research, my study will help establish the cause and effect relationship between the two variables, if any.

Methodology

Participants

The participants in this study were recruited from a convenience sample of students at Leland High School from ages 14 to 18, as explained by **Table 1** below.

Table 1. Demographics of Participants. Note that percentages do not add to 100.00% due to rounding.

	Numbers (out of 35)	Percentage (out of 100 %)
Gender		
Male	13	37.1%
Female	22	62.9%
Age		
14	1	2.9%
15	5	14.3%
16	10	28.6%
17	18	51.4%
18	1	2.9%
Race		
Asian	34	97.1%
White	1	2.9%

As most participants were underaged (<18) students, consent forms were provided and signed electronically by both the parents and the students. The forms request information regarding participants' age, gender, race, test anxiety scores, and working memory task results. Participants also consented to taking a working memory task at two prospective environments at the school, with a 30 minutes run-time each. To ensure subjects' confidentiality, participants were labelled by numbers at the start of the experiment without any other personally identifiable information.

Procedure

Participants were instructed to indicate a classroom on Leland High School's campus where they had taken at least 5 academic tests and where they felt the most "comfortable/relaxed" in. The classroom participants reported was then used as the "familiar environment" in the study. Participants then identified a classroom on campus that they have never been in before, and this was the "unfamiliar environment."

Prior to the test date, participants were asked to take a Likert survey, describing their test anxiety level. With that information, I have ensured that all participants have a moderate level of test anxiety—one that is not too extreme comparably to the rest of the subjects and one that will not strongly fluctuate their reading span scores. Along with the gathered information, I grouped the students into blocks based on all three factors: age, gender, and anxiety level.

Within each block, participants were then separated and assigned into two groups through a random generator. Ideally, the testing should have been counterbalanced such that, half of the participants took the working memory span tasks in their familiar locations first and the unfamiliar second. The other half, vice versa. Since participants would already have more practice and experience with the tasks taking it a second time, this was proposed to ensure the reliability of the test, such that the test-retest effect—also known as the practice effect—would not be the factor affecting students' working memory capacity scores. However, due to scheduling issues and time constraints, 21 participants ultimately took the test at the familiar environment first and unfamiliar second, while 14 others did vice versa. Therefore, the average scores in the unfamiliar environment may be slightly biased upwards, as more participants took the test in the familiar environment first after practice. Note that this confounding effect would act in the opposite direction to the hypothesis.

After the data was obtained, a paired samples t-test analysis was conducted to test for a significant difference in students' working memory capacity scores in the familiar and unfamiliar testing environments. Because test-retest effect was a potential bias in my study, another t-test analysis was performed to examine any significant difference in students' scores between the first-time and the second time testing.

Other factors that were taken into consideration include sleeping time and noise level. To prevent sleep deprivation as a factor affecting students' test scores, I have asked all participants to sleep at least 7 hours prior to their test dates. As for noise level, I have made sure that the room is quiet as if it was during a real testing in school.

Materials

Test Anxiety and Stress Survey

A simple test anxiety questionnaire developed by Nist and Diehl (1990) was given to participants to fill out prior to the test dates. The survey aims to determine if some students experience cases of test anxiety. Participants were to carefully read through each statement and reflect upon their past (pleasant and unpleasant) testing experiences. Then, participants indicated the frequency of such stressful events happening in their life by choosing a number from one to five, from 1 being "never" to 5 being "always."

The survey includes a total of 10 statements, as displayed in *Appendix A*. Participants' scores were tallied and ranged between 10 and 50. A low score between 10 to 19 signifies that the participants do not suffer from test anxiety. A median score between 20 and 35 indicates that the subjects exhibit some characteristics of test anxiety, but the stress level is healthy. Participants who scored over 35 experience an unhealthy level of anxiety. Since emotional instability plays a large role in students' working memory span tasks scores (Krause-Utz, 2012), it is important to consider it alongside any potential effects of environmental familiarity.

Reading Span Task

To investigate the effects of environmental familiarity on students' working memory capacity, a version of Daneman and Carpenter's (1980) dual processing reading span task was implemented in my study. During my pilot study, I conducted the experiment with both the reading span and operation span tasks. It was found that participants' scores

in the reading span task had a normal distribution, whereas the results in the operation span displayed a ceiling effect—meaning that the task was too easy. Ultimately, the reading span task was considered as the best fit for my research.

For the purpose of my experiment, I have coded my own version of the reading span task, which involves two separate, yet concurrently performed, tasks: a processing task and a memory span task. Participants evaluate the logic of written sentences while memorizing single digits for later recalls.

In a trial, subjects are first presented with a number from 0 to 9, followed by an English sentence. This sentence may or may not make logical sense. The following are examples of a correct sentence and an incorrect sentence, respectively: “I like to swim in the pool” and “I like to swim in the sun.” After reading the sentence, participants are to respond to the question “does the sentence make logical sense?” Subjects reply “y” for “yes, it makes sense,” or “n” for “no, it doesn’t make sense.” Following the response, another number and sentence combination is then displayed. At the end of a series of sentences and numbers, also known as a trial, participants are asked to recall all the digits previously displayed in the correct order. The sentences are introduced merely as a way to engage individuals in some sort of cognitive processing. The working memory capacity is then measured by the numbers of digits that the participants can correctly recall at the end of each trial.

Each participant is given a total of 34 trials—that is, 2 repetitions of each set size of 2 and 3, followed by 6 repetitions of each set size from 4 to 8. A set size is essentially the amount of numbers that the participants are asked to recall on a trial. First 4 trials at the beginning of the test with set sizes of 2 and 3 are merely for participants to practice and get used to the test. The remaining 30 trials are utilized to calculate participants’ working memory span scores. The order of set sizes is sequentially determined. The numbers to-be-recalled are randomly generated without replacement—meaning that the same number can appear twice in a row. The sentences presented were taken from previously designed tests and selected to be of moderate difficulty. The sentences were no longer than 20 words, and errors were hidden within the sentences.

Scoring

Reading span task scores were calculated based on the total numbers that participants recalled correctly after all 30 trials. For each set size from 4 to 8, participants received 6 trials—that is, 6 repetitions of trying to recall the “set size” amount of numbers after a series of number/sentence sequences. A score will be calculated for each set size by adding up the amount of numbers correctly recalled for each of the 6 trials. The total possible correct for each set size would be (set size) times (6 trials). With the scores for each set size, a total correct score is then calculated for each environment as the sum of all scores for all set sizes. The maximum total correct for the test was 180.

Data Collection

The means of the participants’ scores at the two environments are displayed in **Table 2**, and the total mean difference between the scores are further emphasized by **Figure 1**.

Table 2. This table displays the mean scores of all 35 students for each set size and at each prospective environment. The row entitled “total correct” represents the participants’ mean scores after all the raw points of each set size were added and taken the average of.

Students’ Average Scores At The Two Prospective Environments

Set Size	Familiar	Unfamiliar
4	23.200	22.543
5	28.200	27.914
6	33.286	32.714
7	35.057	35.657
8	38.914	37.743
Total Correct	160.657	156.571

Average Total Scores at The Two Environments

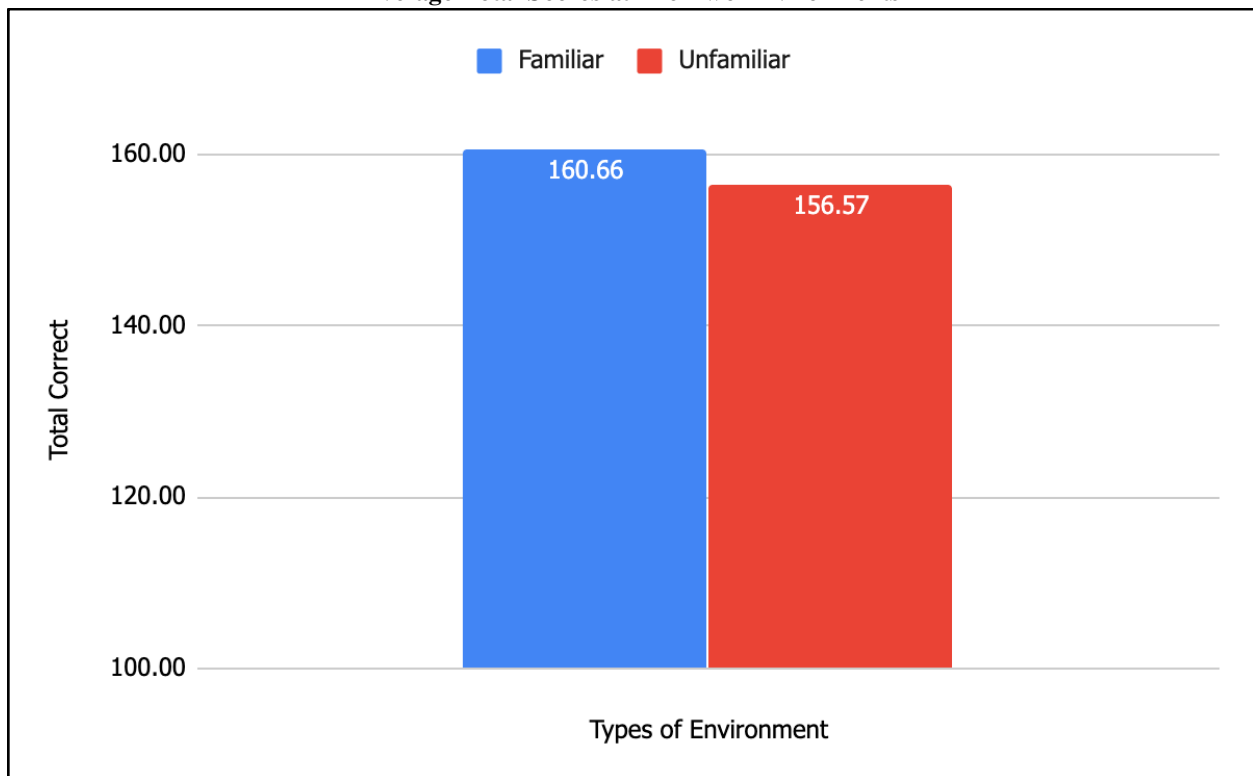


Figure 1. The total correct scores displayed represent the averages of all 35 participants’ counts for each prospective location. Note that the vertical axis starts at 100, not 0.

Analysis

Familiarity with Environments

To calculate for a significance difference between the two mean scores between the familiar environment versus the unfamiliar environment, I performed a paired samples t-test analysis.

Hypothesis

The goal in my research is to test for a significant difference between the students' working memory capacity mean scores at their familiar environment (μ_f) versus their scores at their unfamiliar environment (μ_u) at $\alpha = 0.05$ (95%) level of significance. The difference between students' scores at the two prospective locations is then denoted by μ_D , or $\mu_f - \mu_u$.

My null hypothesis (H_0) is that there is no significant difference between students' working memory capacity scores at their familiar environment compared to their scores at their unfamiliar environment. That is, $\mu_f = \mu_u$, or $\mu_D = 0$. My alternative hypothesis (H_A) is students' working memory capacity scores at their familiar environment are higher than to their scores at their unfamiliar environment. That is, $\mu_f > \mu_u$, or $\mu_D > 0$.

Conditions

To qualify for the proper usage of t-test, the data must be random, independent, and normal. For analysis purposes, I assume my conveniently sampled data is representative and random. My data is considered independent because each student's working memory capacity scores is not influenced by any other individuals. Also, the school population is at least 10 times the sample size (>350). Lastly, my data is said to follow a normal distribution due to the central limit theorem, which states that "the distribution of sample means approximates a normal distribution as the sample size gets larger" (Investopedia, 2019). The central limit theorem typically requires a sample size equals to or greater than 30, which fits for my sample of size 35.

Calculation

The mean difference was calculated as the average of each student's score difference between the two environments. The standard deviation values were evaluated to indicate the deviation of each participant's score difference away from the mean difference. The t-statistic scores were calculated utilizing mathematical formulas for a matched-paired two sample t-test on Excel. The p-values were found in correspondence of the t-statistics. The results are explicitly exhibited in **Table 3**.

Table 3. Positive mean difference signifies that students' average score at the familiar environment is greater than their average score at the unfamiliar environment.

Set Size	Mean Difference	Standard Deviation	T-Statistics	P-Value
4	0.6571	2.5661	1.515	0.0695
5	0.2857	2.2172	0.762	0.2255
6	0.5714	3.4580	0.978	0.1675
7	1.4000	4.1245	2.008	0.0263
8	1.1714	6.4965	1.067	0.1467
Total Correct	4.0857	10.7877	2.240	0.0158

Also note that values of the “Total Correct” are not simply the sums of each column. Rather, they are values conducted with each column’s cumulative mean scores.

Conclusion

At $\alpha = 0.05$ level of significance, the p-value (0.0158) is less than the level of significance (0.05) for the total score, and therefore, the null hypothesis is rejected. This suggests that students scored higher on a working memory span task when taken in their familiar environment than when they took it in an unfamiliar environment. It can then be concluded that unfamiliarity with the surrounding environment does significantly occupy students’ working memory.

Looking at each set size, set size of 7 is where the biggest and the only significant difference in students’ scores at the two locations lies. In comparison with the rest of the set sizes, set size 4 also displays quite a large difference, but it is not significant. These differences—though not greater than the difference observed with the cumulative mean—may hint at some possible theories to be analyzed in the Discussion section.

Order of Testing

To test whether the order of testing contributes significantly to the displayed results, I conducted another paired sample t-test. This analysis compares students’ mean scores of their first-time testing versus their second-time testing—no matter the familiarity of the environment that they were in the first or second time. Students’ average scores for the first-time testing and second time retesting are displayed in **Table 4** below. The mean difference, standard deviation, t-statistics and p-value calculated are shown in **Table 5** below.

Table 4. Average Total Corrects During First-Time Testing Versus Second-Time Retesting

Total Correct	First Time Testing	Second Time Retesting
	157.571	159.657

Table 5. Mean Difference, Standard Deviation, T-Score and P-Value Examining for Significant Difference In Order of Testing

Order of Testing	Mean Difference	Standard Deviation	T-Statistics	P-Value
	-2.0857	11.3614	-1.086	0.1425

Note that the negative in mean difference and t-statistics simply indicate that the mean scores during the second time testing is greater than the scores during the first time.

Based on the result of this t-test analysis, students’ scores did not significantly differ between the two time points (p-value 0.1425 > level of confidence $\alpha = 0.05$). Therefore, it is reasonable to suggest that the order of testing did not greatly influence on the significant result concluded previously.

Discussion

The goal of my research was to examine the correlation between students' working memory capacity and familiarity with their surrounding environment. It has been previously established that unfamiliarity with environment increases anxiety levels (Zatz & Chassin, 1985), and that chronic anxiety level correlates with lower working memory capacity (Eysenck & Calvo, 1992). From this, I hypothesized that environmental unfamiliarity will limit students' working memory capacity temporarily. The result supports my hypothesis. Students' working memory scores at the familiar environment were significantly higher than their scores at the unfamiliar location.

Based on prior literature, individuals are more likely to be stressed, or at least more aware of their environment if placed in an unfamiliar environment (Klein & Boals, 2001). When individuals place more conscious thought about their surrounding environment, they are more likely to be distracted and focus less on the task—which in and of itself already requires a huge amount of attention. With stress and anxiety occupying students' mental resources, it is very possible that these individuals are not operating at their optimal working memory capacity (Klein & Boals, 2001). In addition, previous research on stereotype threats depict similar decreases in students' working memory capacity, explaining such findings with heightened levels of anxiety and arousal (Schmader & Johns, 2003). It further associates increased stress levels with decreases in mental resources and lowered levels of working memory capacity.

Differences at Each Set Size

The biggest and most significant difference in students' mean scores at the two environments is observed at span 7. This trial is where students' working memory was most challenged and thus most prominently influenced by the difference in the environment. The lack of a significant difference at set size of 8 may conclude that such level of difficulty was beyond the working memory capacity of high school individuals—also known as a floor effect. Variance in scores is limited by the extremely hard difficulty on these trials, and therefore, scores are less likely to be influenced by other factors.

It is also interesting to note that scores at span 4 almost significantly differed. This can be observed as where students' working memory is not greatly challenged. This was also the start of the assessment. Students may be more nervous or anxious, and therefore, more aware of their surrounding environment. It is possible that environmental factors have a much greater effect on students at this time of the task, and if follows that performance on standardized academic tests may be most influenced by the environment at the beginning of the test.

Comparing the t-statistic values at each span size, it suggests that the working memory assessment was most sensitive at span 7 and 4, and less sensitive at span 5, 6, and 8.

Limitations

Uneven Distribution in the Order of Testing

Although the additional t-test demonstrated that the order of testing does not significantly contribute to the results concluded, this uneven distribution might still introduce some bias. 21 students took the task in a familiar environment first and unfamiliar second. 14 students did vice versa. In theory, these 21 students should have performed better during the second-time at their unfamiliar environment, while the other 14 students should have done better during the second-time at their familiar environment. This suggests that more students should have performed better in the unfamiliar environments if the participants' performance is much affected by the order of testing. However, students' mean scores at the familiar environments still ended up being significantly higher than the scores at the unfamiliar environments. This indicates that students' familiarity with their surrounding environment essentially caused more of a factor than the order of testing. This analysis also conveys the possibility of having an even more significant result if the participants were truly to take the task at each environment in the order assigned.

Uncontrollable Personal Biases

Participants were asked to have at least 7 hours of sleep prior to each test date and to be as stress free as possible. However, each individual may react differently to the test due to the test dates, their mood, their sleep schedule, and/or any other personal issues. These may all create additional bias in the results.

Convenience Sampling

As subjects were pulled from a convenience sample of my high school population, they may not accurately represent high school students in general. In addition, conditions of t-test require the sample to be random. But my convenience sample may not be random as I assumed in the t-test. This may affect the reliability of the t-test procedure in my analysis.

Hawthorne Effect

Although I did not directly inform any of my participants about the design and the purpose of my experiment, some of the students were able to discover the motive behind my study. As the Hawthorne effect and demand characteristics can be a large determinant in experimental studies, it brings bias to students' scores on the task.

Other Environmental Factors

Even though I had controlled the noise level in each classroom, there are still various environmental factors that could affect students' testing. These include the amount of light, the heating, the aesthetics of the room. These potential factors could all influence students' performance on the task.

Significance and Call for Future Research

This experiment advances researchers' understanding on how familiarity of the environment affects high school students' cognitive function. Manipulations of the surrounding environment can produce situational reductions in working memory capacity, which is an important cognitive mechanism utilized in standardized testing. This underscores the need for educators to re-examine the reliability of the current standardized testing system. With students spread across school campuses and classrooms, the environment plays a huge role in their test scores. The finding demonstrates a potential bias factor in standardized testing that is uncontrollable by the students.

Conclusion

Previous research on environmental familiarity predominantly focused on influential variables that cause fluctuations in students' standardized test scores. Little to no work has been done to examine the consequences of environmental unfamiliarity on the cognitive processing of a high school student. The primary focus of this research is to advance current knowledge on how environmental manipulation decreases students' scores on complex cognitive tests because it lowers their working memory capacity. This study found converging evidence that performing in unfamiliar settings can deplete individuals' working memory capacity and impair their results on challenging academic tasks. Although working memory capacity may not be the only cognitive function involved in standardized testing, these findings still demonstrate how environments can undermine the "talents and abilities of individuals who might otherwise meet their full potential in high-stakes performance situations, such as college entrance exams" (Schmader & Johns, 2013). I hope that this research meets its maximum potential and contributes to the development of a fair and just standardized testing system, where environmental unfamiliarity should not and cannot play a role in influencing students' performance.

Acknowledgements

This research project would not have been possible without the guidance of my AP Research teacher Anu Sarkar, from Leland High School. Her passion and knowledge greatly motivated me and were great forces that guided me throughout the whole process. I also greatly appreciate Dr. Jonathan Jones, a professor at University of Cambridge - MRC Cognition and Brain Sciences Unit, and Rachel Booth, a research assistant at Ulm University, for their dedication and support. Their feedback and suggestions only allowed me to strive for better throughout the various revisions of my paper. I thank them for all their hard work as well as my very own parents' support throughout the journey.

References

- Aldridge, Jill, and Rosalyn Anstine Templeton. "School Climate: Measuring, Improving and Sustaining Healthy Learning Environments. H. Jerome Freiberg (Ed.)." *Learning Environments Research*, vol. 2, Oct. 1999, pp. 331–334., doi: 10.1023/A:1009926031342.
- Alloway, Tracy Packiam, et al. "A Structural Analysis of Working Memory and Related Cognitive Skills in Young Children." *Journal of Experimental Child Psychology*, vol. 87, no. 2, 27 June 2003, pp. 85–106., doi:10.1016/j.jecp.2003.10.002.
- Baddeley, Alan David, and Graham Hitch. "Working Memory." *Psychology of Learning and Motivation*, vol. 8, 1974, pp. 47–89., doi:10.1016/S0079-7421(08)60452-1.
- Chu, Man-Wai, et al. "Using HLM to Explore the Effects of Perceptions of Learning Environments and Assessments on Students' Test Performance." *International Journal of Testing*, vol. 14, no. 2, 28 Apr. 2014, pp. 95–121., doi:10.1080/15305058.2013.841702.
- Conway, Andrew R. A., et al. "Working Memory Span Tasks: A Methodological Review and User's Guide." *Psychonomic Bulletin & Review*, vol. 12, no. 5, 2005, pp. 769–786., doi:10.3758/bf03196772.
- Cowan, Nelson, et al. "Models of Verbal Working Memory Capacity: What Does It Take to Make Them Work?" *Psychological Review*, vol. 119, no. 3, July 2012, pp. 480–499., doi:10.1037/a0027791.
- Cowan, Nelson. "The Many Faces of Working Memory and Short-Term Storage." *Psychonomic Bulletin & Review*, vol. 24, no. 4, 28 Nov. 2016, pp. 1158–1170., doi:10.3758/s13423-016-1191-6.
- Daneman, Meredyth, and Patricia A Carpenter. "Individual Differences in Working Memory and Reading." *Journal of Verbal Learning & Verbal Behavior*, vol. 19, no. 4, 1980, pp. 450–466., doi:10.1016/S0022-5371(80)90312-6.
- Draheim, Christopher, et al. "What Item Response Theory Can Tell Us about the Complex Span Tasks." *Psychological Assessment*, vol. 30, no. 1, 9 Mar. 2017, pp. 116–129., doi:10.1037/pas0000444.
- Engle, Randall W. "Working Memory Capacity as Executive Attention." *Current Directions in Psychological Science*, vol. 11, no. 1, Jan. 2002, pp. 19–23., doi:10.1111/1467-8721.00160.
- Engle, Randall W., et al. "Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach." *Journal of Experimental Psychology: General*, vol. 128, no. 3, 1999, pp. 309–331., doi:10.1037/0096-3445.128.3.309.
- Eysenck, Michael W., and Manuel G. Calvo. "Anxiety and Performance: The Processing Efficiency Theory." *Cognition & Emotion*, vol. 6, no. 6, 1992, pp. 409–434., doi:10.1080/02699939208409696.
- Eysenck, Nazanin Derakshan Michael W. "Working Memory Capacity in High Trait-Anxious and Repressor Groups." *Cognition & Emotion*, vol. 12, no. 5, 1998, pp. 697–713., doi:10.1080/026999398379501.

- Ganti, Akhilesh. "What Is the Central Limit Theorem (CLT)?" *Investopedia*, Investopedia, 13 Sept. 2019, www.investopedia.com/terms/c/central_limit_theorem.asp.
- Gennip, Nanine A.e. Van, et al. "Peer Assessment for Learning from a Social Perspective: The Influence of Interpersonal Variables and Structural Features." *Educational Research Review*, vol. 4, no. 1, Dec. 2009, pp. 41–54., doi:10.1016/j.edurev.2008.11.002
- Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy and Practice*, 16(3), 365-381. doi:10.1080/09695940903319745
- Hockley, William E., et al. "Associative and Familiarity-Based Effects of Environmental Context on Memory." *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, vol. 66, no. 2, 2012, pp. 81–89., doi:10.1037/a0027136.
- Kail, Robert, and Lynda K. Hall. "Distinguishing Short-Term Memory from Working Memory." *Memory & Cognition*, vol. 29, no. 1, 2001, pp. 1–9., doi:10.3758/bf03195735
- Kenton, Will. "T-Test Definition." *Investopedia*, Investopedia, 22 Mar. 2020, www.investopedia.com/terms/t/t-test.asp.
- Klein, Kitty, and Adriel Boals. "Expressive Writing Can Increase Working Memory Capacity." *Journal of Experimental Psychology: General*, vol. 130, no. 3, 2001, pp. 520–533., doi:10.1037/0096-3445.130.3.520
- Krause-Utz, A., et al. "Influence of Emotional Distraction on Working Memory Performance in Borderline Personality Disorder." *Psychological Medicine*, vol. 42, no. 10, Aug. 2012, pp. 2181–2192., doi:10.1017/s0033291712000153
- Mathy, Fabien, et al. "Simple and Complex Working Memory Tasks Allow Similar Benefits of Information Compression." *Journal of Cognition*, vol. 1, no. 1, 25 May 2018, doi:10.5334/joc.31
- "More Than 2 Million Students in the Class of 2018 Took the SAT, Highest Ever." *The College Board*, The College Board, 18 Mar. 2019, www.collegeboard.org/releases/2018/more-than-2-million-students-in-class-of-2018-took-sat-highest-ever.
- Quinn, Diane M., and Steven J. Spencer. "The Interference of Stereotype Threat With Womens Generation of Mathematical Problem-Solving Strategies." *Journal of Social Issues*, vol. 57, no. 1, 2001, pp. 55–71., doi:10.1111/0022-4537.00201
- Ravitch, Diane, and Lois A. Stoehr. "The Death and Life of the Great American School System." *Early Learning*, 2 Mar. 2010, pp. 125–128., doi:10.4324/9781315429977-13
- Schmader, Toni, and Michael Johns. "Converging Evidence That Stereotype Threat Reduces Working Memory Capacity." *Journal of Personality and Social Psychology*, vol. 85, no. 3, 2003, pp. 440–452., doi:10.1037/0022-3514.85.3.440
- Simon, Sharon S., et al. "Increasing Working Memory Load Reduces Processing of Cross-Modal Task-Irrelevant Stimuli Even after Controlling for Task Difficulty and Executive Capacity." *Frontiers in Human Neuroscience*, vol. 10, no. 380, 3 Aug. 2016, doi:10.3389/fnhum.2016.00380
- Swanson, H. Lee. "Short-Term Memory and Working Memory." *Journal of Learning Disabilities*, vol. 27, no. 1, 1 Jan. 1994, pp. 27–34., doi:10.1177/002221949402700107
- "T Test (Student's T-Test): Definition and Examples." *Statistics How To*, 2020, www.statisticshowto.datasciencecentral.com/probability-and-statistics/t-test/.

Taylor, Bret Allen. "The Influence of Classroom Environment on High School Students' Mathematics Anxiety and Attitudes." Aug. 2004, <http://hdl.handle.net/20.500.11937/73>.

Taylor, Lee, et al. "The Impact of Different Environmental Conditions on Cognitive Function: A Focused Review." *Frontiers in Physiology*, vol. 6, 6 Jan. 2016, doi:10.3389/fphys.2015.00372

Thorsen, Cecilia, et al. "The Influence of Fluid and Crystallized Intelligence on the Development of Knowledge and Skills." *British Journal of Educational Psychology*, vol. 84, no. 4, Sept. 2014, pp. 556–570., doi:10.1111/bjep.12041

Wilhelm, Oliver, et al. "What Is Working Memory Capacity, and How Can We Measure It?" *Frontiers in Psychology*, vol. 4, 24 July 2013, doi:10.3389/fpsyg.2013.00433.

Williams, Tony. *To Hell and Back Three Times: the Cancer Experience*. MLW Inc., 2003.

Zatz, Sheri, and Laurie Chassin. "Cognitions of Test-Anxious Children under Naturalistic Test-Taking Conditions." *Journal of Consulting and Clinical Psychology*, vol. 53, no. 3, 1985, pp. 393–401., doi:10.1037/0022-006x.53.3.393