

2019 FALL
VOLUME 2

The Northern Kentucky University
Journal Of Student Research

AS
S
Y
Z

TABLE OF CONTENTS

- | | | | |
|----|----------------------------------------------------------------------------------------------------------------------------------------------|----|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| 01 | About Nýsa | 25 | The Use of Qualified Medical Interpreters in Health Care: Barriers for Health Care Professionals
<i>Ana Alza-Rodriguez, Judi Wheatley Godsey</i> |
| 02 | The Effects of Reading Speed and Retrieval Practice on Reading Comprehension
<i>Carrie Weimer, Kalif E. Vaughn</i> | 32 | A Barrier to Democracy: Corruption in Former Soviet Eastern Europe
<i>Lydia Schubarth, Ryan Salzman</i> |
| 12 | Pen Chromaticity's Effect on the Recall of Details Following Lecture Annotation
<i>Jessica Raines, Kathleen Fuegen</i> | 40 | Other NKU student research publications |
| 19 | Using High School Financial Literacy Education to Predict Future Income: A Story of Selection Bias
<i>Tyler Kent, Abdullah Al-Bahrani</i> | | |

NÝSA, THE NKU JOURNAL OF STUDENT RESEARCH

Nýsa publishes research from students at NKU and across the commonwealth. It is published by NKU's Institute for Student Research and Creative Activity. All submissions are peer-reviewed by NKU faculty and students.

About The Title

Names are tricky things. Journals of student research are relatively common, and in looking for a name, it was important to find something evocative of the intellectual effort and exhilaration that accompany any research endeavor. If it could relate to our identity as The Norse, all the better. "Nýsa" worked perfectly. In the words of David Kime, Advising Coordinator for NKU's Honors College, who suggested it:

"The Viking raids were only one aspect of Norse society. The Norse were shipbuilders, farmers, philosophers, poets, artists, and merchants. The Norse were explorers who engineered new shipbuilding technology and navigation techniques. They sought new knowledge in the stars and from distant lands and cultures. In Old Norse, "nýsa" is a verb meaning to search or investigate; to peer into the unknown. The idea of "nýsa" applies to today's NKU students as much as it did to the Norse a thousand years ago as they peer into the unknown and produce new and exciting examples of research, scholarship, and creativity."

About The Cover

The cover and interior for this issue of Nýsa, The NKU Journal of Student Research was designed by Jacob Castle. Jacob is a multi-disciplined graphic designer and currently pursuing his BFA degree in Visual Communication Design at NKU. He works in a broad range of design projects, but is particularly interested in publication, brand, and packaging design. Jacob is passionate about helping elevate brands/products with the use of design. When designing the cover Jacob looked into the meaning of Nýsa, how it connects the NKU students doing research for the publication, and how it connects to innovation. "I wanted the design to connect with the process the researchers go through," Jacob says. He adds "The design not only has a very close feel to NKU's brand, but it has elements (shapes, patterns, light) that give a sophistication to it that the researchers must have to begin a project like this and an energy to it that could represent what students feel when they break ground in their research paper." Jacob wanted the cover to intrigue readers, catch their eye and encourage them to read the publication.

From The Editor

One of the pleasures of editing a journal that includes submissions from any and all fields is learning not just about the specific topics that NKU students are studying, but also discovering the different writing and article styles and standards. This issue of Nýsa includes contributions from psychology, nursing, economics, and international studies, and while formats vary, the dedication to scholarship does not. The fruits of the work by the student authors and faculty mentors included here are obvious. What remains hidden but no less important are the efforts of the NKU student and faculty reviewers, editors, editorial board members, the Institute for Student Research and Creative Activity, and administrators who support the scholarship and make this publication possible. Our students would not succeed without their efforts, and I am grateful to them all. Please enjoy this second issue.

Patrick M. Hare

Editorial Board

Patrick M. Hare, Ph.D. (Chemistry & Biochemistry)

Editor-in-Chief

E. Kobena Osam, Ph.D. (Organizational Leadership)

Julie Hart, DNP, RN, CNE (Nursing)

Kalif Vaughn, Ph.D. (Psychological Sciences)

Lexi Bensberg, student (Applied Statistics and Social Justice)

Amal Said, Ph.D. (Economics)

Abdullah Al-Bahrani, Ph.D. (Economics)

The Effects of Reading Speed and Retrieval Practice on Reading Comprehension

Carrie Weimer.

Faculty mentor: Kalif Vaughn

Psychology

Carrie Weimer

Carrie Weimer graduated Summa Cum Laude as a university honors scholar at Northern Kentucky University in May 2019. She obtained a Bachelor of Science degree in psychology and minors in French and Spanish. Carrie was grateful for the opportunity to work with her mentor, Dr. Kalif Vaughn, on this project, which she believes was an important experience that will help her in her future studies and career. She is currently pursuing dual Master's degrees in Clinical Mental Health Counseling and School Counseling at Northern Kentucky University.

KEYWORDS:

speed reading, retrieval practice, reading speed, comprehension

Abstract

Although research suggests that speed reading is ineffective, it remains a popular strategy among students. We investigated the impact of an effective strategy (engaging in self-test practice, also known as retrieval practice) on speed reading. Would retrieval practice improve the efficacy of speed reading? We had participants read two passages (one at a regular rate and one at an accelerated rate) and they were given quizzes over the material either before or after reading the passages (except for the control group, which never received a quiz). Participants then completed a final comprehension test over the material. We predicted that receiving a quiz before reading would most benefit speed reading, whereas a quiz after reading would most benefit normal reading. The results did not support our hypothesis. Results indicated that regular reading was always the superior strategy regardless of quiz placement. These results suggest that speed reading is detrimental to comprehension, even when paired with retrieval practice. Limitations and future directions are discussed.

Introduction

"Acquiring knowledge without retaining it is a fruitless venture, yet it characterizes the experience of many college students" (Hopkins, Lyle, Hieb, & Ralston, 2016, p. 854). College students attempt to absorb copious amounts of information but do not retain the information over time (Hopkins et al., 2016). Ebbinghaus (1885) originally discovered the forgetting curve, which shows that memory declines sharply within the first few days following acquisition (approximately 70-80% of what is learned is unrecalable within a matter of days; see also Murre and Dros, 2015). Clearly, unless students further process the content, they will forget a large portion of what they have studied relatively quickly.

This problem is compounded by the fact that students engage in suboptimal encoding strategies (see Dunlosky, Rawson, Marsh, Nathan, and Willingham, 2013). In the current study, we explored one particularly poor encoding strategy: speed reading. Presumably, students speed-read when they lack the time to read something at a normal speed. Students are likely to continue using this strategy despite evidence suggesting it is not effective.

Although prior research suggests that speed reading impairs comprehension (Rayner, Schotter, Masson, Potter, and Treiman, 2016), perhaps it can be made more efficacious by using a strategy known to improve learning. In particular, one promising learning strategy is known as retrieval practice (i.e., attempting to recall information from memory; see Roediger and Karpicke (2006) for a review). Before discussing the potentially interactive effects of these two methods, we first need to explain why speed reading is ineffective and why retrieval practice may help.

Why speed reading is ineffective

Peripheral vision is limited

Some studies investigated the possibility of reading a large portion of a page at a time with peripheral vision. Rayner et al., (2016) immediately discredited this idea because of the psychological and biological impossibility of performing such an action due to the limitations of visual acuity (i.e., the clarity of sight). Peripheral vision is limited; therefore, it is impossible to use it to read more than a few words at a time.

Subvocalizing is important for comprehension

Researchers have investigated whether eliminating the inner voice while reading silently could increase reading speed. For example, Slowiaczek and Clifton (1980) found that subvocalization (i.e., a mental voice) improved the durability

of memories and was essential for sentence comprehension. Overall, subvocalization is important for comprehension, as mental voices enable readers to better understand visual material after converting the material into an auditory format (Rayner et al., 2016).

Regressive eye movements are useful

Can reading speed be improved by eliminating regressive eye movements (i.e., backtracking) while reading? Acklin and Papesh (2017) found that comprehension was significantly lower when regressive eye movements were eliminated. These findings were the same regardless of the difficulty of the text. Regressive eye movements are useful, and their elimination could limit comprehension during reading.

Skimming is a Trade-off

Skimming is the act of scanning a passage to find specific information or garner the general idea (Rayner et al., 2016). Rayner et al. (2016) found numerous sources supporting skimming and strategies associated with skimming (e.g., looking at headings and the table of contents). These researchers concluded that an increased reading speed will likely result in a decrease of reading comprehension; however, the use of skimming can prove useful in managing large amounts of reading material. Overall, these researchers concluded that skimming is a trade-off (i.e., it can help manage copious amounts of reading, but does not allow for full comprehension) and the most effective way to improve reading speed and comprehension is through practice.

Does speed reading ever help?

Although speed reading is typically ineffective, there are cases where it has been shown to help. Breznitz and Share (1992) found that comprehension could be improved through the elimination of regressive eye movements, although other researchers have warned against this practice (Rayner et al., 2016). Duggan and Payne (2009) found that comprehension when skimming could be comparable to reading normally; however, they concluded that there are benefits to thoroughly reading. Although these studies show instances in which speed reading could be beneficial, these situations are exceptions to the rule that speed reading is ineffective. The purpose of the current work was to explore whether retrieval practice could improve the effectiveness of speed reading.

Retrieval Practice

Retrieval practice refers to the retrieval of information from memory, such as through testing (e.g., "What is the capital

of Australia?"). After engaging in retrieval practice, students typically remember more information and remember the information for a longer period (Karpicke, Butler, & Roediger, 2009). Below, we will review various factors that influence the benefits of retrieval practice.

Retrieval Practice Spacing

Hopkins, Lyle, Hieb and Ralston (2016) investigated spaced- (i.e., test practice distributed across numerous sessions) versus massed- (i.e., test practice crammed into one session) retrieval practice in relation to comprehension and retention. The researchers found that comprehension was higher, for both short- and long-term retention, after using spaced-retrieval practice.

Test Format

Smith and Karpicke (2014) investigated the effects of test format (i.e., short-answer, multiple-choice, and hybrid) on the benefits of retrieval practice. They also explored whether concordance between the quiz and test was necessary for the effects of retrieval practice. The researchers found that the effects of retrieval practice manifested regardless of test format and regardless of whether the quiz and test shared the same format.

Quiz Placement

Burns and Vinchur (1992) investigated the benefits and restrictions of post-lecture quizzes on comprehension. They found that post-lecture quizzes were only beneficial when the content and difficulty level were similar to those found on the final exam. Narloch, Garbin, and Turnage (2006) conducted a follow-up study to learn if the same limitations applied to pre-lecture quizzes. The study was conducted over five semesters, which allowed for a within-subjects design. Two semesters had no quizzes and functioned as a control. The other three semesters had quizzes administered to the students at the beginning of class at the start of each chapter. The researchers also investigated quiz format (i.e., matching versus fill-in-the-blank) for which they found no significant difference in student performance. The investigators did find that student performance was significantly better during semesters with a quiz than during the control semesters. The researchers did not find evidence of limiting factors for pre-lecture quizzes, such as those found by Burns and Vinchur (1992) for post-lecture quizzes (Narloch, Garbin, & Turnage, 2006). This study indicated that quizzes done prior to learning significantly improved subsequent recall of the material. A large body of research has now firmly established the benefits of pre-testing (e.g., Knight, Ball, Brewer, DeWitt, and Marsh, 2012; Kornell, Hays, and Bjork, 2009; Vaughn and Rawson, 2012).

Combining speed reading with retrieval practice

Research has been conducted to assess methods of improving reading comprehension during speed reading (e.g., Duggan and Payne, 2009; Breznitz and Share, 1992; Acklin and Papesh, 2017); however, no prior studies have investigated the combination of speed reading and retrieval practice.

The current study examines the relationship between speed reading and reading comprehension to determine the effect speed reading has on reading comprehension. More specifically, the purpose of the study was to determine if retrieval practice would improve reading comprehension following speed reading. We also investigated the effect retrieval practice has on reading comprehension in conjunction with both a normal reading rate and speed reading.

This study is important because no prior research has investigated the interaction between speed reading (generally accepted as detrimental to reading comprehension) and retrieval practice (widely accepted as beneficial to learning). If retrieval practice can improve the benefits of speed reading, then students may be able to read more in less time which could improve their academic performance. In the present study, participants engaged in one of two retrieval practice conditions (pre-quiz versus post-quiz) or a control condition (no-quiz) which did not engage in retrieval practice. In the pre-quiz condition, participants completed a quiz before they read the passage. In the post-quiz condition, participants completed a quiz after they read the passage. For both quiz types, the quiz questions were identical. The participants in each of these conditions read two passages, one at a slightly-lower-than-normal rate of speed (150 wpm) and the other at an increased rate of speed (750 wpm) (see Figure 1 for a schematic of the possible conditions).

Hypotheses

Throughout the study, we investigated numerous hypotheses, converging on the hypothesis that reading comprehension would be greater for participants when reading at a normal rate than when speed reading. We hypothesized that the use of retrieval practice would improve reading comprehension; however, the margin of improvement would be dependent upon additional factors, as outlined below.

Normal Reading Rate. We hypothesized that the participants in the regular reading group would have the best reading comprehension scores when given a post-quiz. We hypothesized this because we believed that the post-quiz would catalyze the testing effect, thus improving performance on the final test.

Speed reading Rate. We hypothesized that the participants in the speed-reading group would exhibit the worst reading

Table 1. Self-reported speed-reading usage

Speed Reading Use	Always		Frequently		Sometimes		Rarely		Never	
	n	%	n	%	n	%	n	%	n	%
	2	1.6	24	18.9	34	26.8	38	29.9	23	18.1

Table 2. Self-reported effectiveness for speed reading

Speed Reading Effectiveness ^a	More Effective		Just as Effective		Somewhat Effective		Not Effective at All	
	n	%	n	%	n	%	n	%
	4	3.1	4	3.1	50	39.4	63	49.6

a. Question asked participants to compare the effectiveness of speed reading to regular reading.

comprehension scores when given no quiz; however, we hypothesized that retrieval practice would improve reading comprehension, even for the speed-reading group. More specifically, we hypothesized that pre-quizzes would most benefit the speed-reading condition. We anticipated that the pre-quiz would prime the information and cause related information in the passage to stand out to the readers.

Methods

Participants

The participants were 128 college students (23 males, 96 females, 2 non-binaries; $M_{age} = 20.23$, $SD = 3.96$, range: 18–44 years) and were recruited via the online research participation management system SONA. Participants accessed the study through SONA in a lab setting ($n = 26$) or online ($n = 102$). According to self-reported demographics, the sample was predominantly White, Non-Hispanic (81.1%), with some individuals identifying as African American, Non-Hispanic (5.5%), Hispanic/Latino (3.9%), Asian/Pacific Islander (3.9%), and Other (0.8%). The participants consisted of freshmen (53.5%), sophomores (16.5%), juniors (10.2%), and seniors (13.4%) with a variety of majors, such as nursing (27.6%), psychology (21.3%) and pre-nursing (10.2%) ($M_{GPA} = 3.40$, $SD = 0.48$, range: 1.60–4.00). They received 2 SONA credits for their participation in the study.

Participants were excluded if they restarted or did not finish the experiment ($n = 17$) or if they wished not to be included in the final analyses ($n = 1$). Although some participants indicated that they had completed the experiment before ($n = 1$), we did not exclude them because this was our first experiment using these materials and participants could not have completed it more than once due to SONA naturally restricting based on NKU email.

Participants completed a survey to provide an estimate of the number of books they have read outside of school in the past year ($M = 8.37$, $SD = 24.72$; median = 3) and if they identify as an avid reader (30.7%), not an avid reader (64.6%), or preferred not to answer (4.7%). In a subsequent question, participants described their self-classification as an extrovert (41.7%) or introvert (53.5%), or they indicated that they preferred not to answer (4.7%). They also reported how often they engage in speed reading (see Table 1). Results indicated that, although some students avoid using the strategy completely, many students engage in speed reading (even if only sporadically). Additionally, we asked students to rate the effectiveness of speed reading compared to regular reading (see Table 2). Results indicated that almost half of the students believe that speed reading is at least somewhat effective, whereas half believe that speed reading is entirely ineffective.

Materials

The materials included two passages, a quiz, and a reading comprehension test. The passages were about sharks and the Maori people (see Appendix A and Appendix B, respectively). Each passage had 350 words and was presented on a white background with black font. The quiz consisted of five multiple-choice questions selected from the comprehension test. The comprehension test included ten multiple-choice questions (see Appendix C and D, respectively), encompassing the main ideas and information presented in the passage. After completing the study, participants completed a survey. In this survey, participants were asked to provide an estimate of the number of books they have read outside of school in the past year, if they identify as an avid reader, and their self-classification of extrovert or introvert.

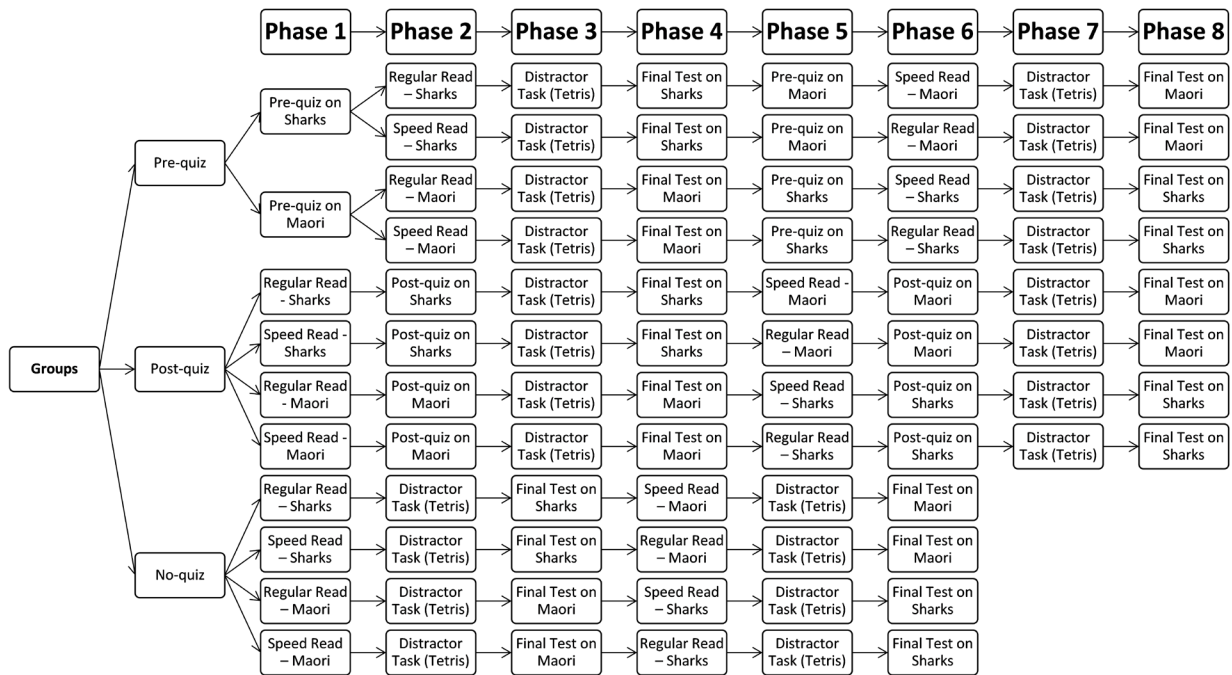


Figure 1. A schematic showing all possible groups of participants in the study.

Design

The experiment was a 3 (Quiz Placement: pre-quiz versus post-quiz versus no quiz) \times 2 (Reading Speed: regular reading versus speed reading) \times 2 (Passage: Maori versus Sharks) \times 2 (Location: Lab or Online) design, counterbalanced for the order of presentation (i.e., which passage participants read first) and reading rates (i.e., which passage they sped-read).

Procedure

All participants received two reading passages, one of which they read normally and the other they sped-read. The participants were allotted 2 minutes and 20 seconds to read at their normal rate (150 wpm) and 28 seconds to speed read (750 wpm). The participants were randomly assigned to one of three groups: pre-quiz (quiz before reading the passage), post-quiz (quiz after reading the passage), and no-quiz. All quizzes consisted of five multiple-choice questions, and participants had 30 seconds to answer each question. All participants completed a distractor task (Tetris); however, the placement of the distractor task varied based upon their group assignment. The length of the distractor task was varied across groups to ensure equal spacing between reading the passage and taking the final test. Participants in the pre-quiz group and no-quiz group played Tetris for 3 minutes after reading the passage, whereas participants in the post-quiz group played Tetris for

30 seconds after their post-quiz. This final test contained the five quiz questions presented to the pre-quiz group and post-quiz group and five new questions, and participants had an unlimited amount of time to answer each question. The same procedure was repeated for the second passage. As noted in the design section, assignment of reading speed condition (speed reading or normal reading) to passage (sharks or Maori) was approximately counterbalanced across participants. Similarly, the order of the passages (sharks first or Maori first) was approximately counterbalanced across participants.

Results

The independent variables utilized throughout the study were which passage was read at a normal rate, which passage was read quickly (speed-reading condition), which rate was utilized first, and which passage was read first. The depended variables measured during this study were the total score out of five correct answers on the quizzes and the total score out of ten correct answers on the final tests (for both normal reading and speed reading).

Quiz Performance

Quiz performance is plotted in Figure 2. We conducted a 2 (Speed First: Regular Reading or Speed reading) \times 2 (Passage First: Maori or Sharks) \times 2 (Location: Lab or Online) \times 2

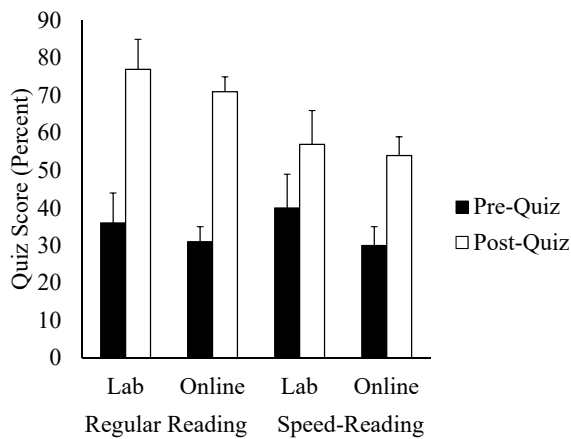


Figure 2. A bar graph of quiz score results. Participants performed significantly better in the lab. They also performed significantly better on the post-quiz than on the pre-quiz.

Group (Pre-Quiz or Post-Quiz) Repeated-Measures ANOVA with 2 within-participant levels (Regular reading quiz score and Speed-reading quiz score). There was a main effect of reading speed on quiz performance, $F(1, 76) = 4.42, p = .039$. Participants scored significantly higher when reading regularly than when speed reading. Additionally, group (pre-, post-, or no-quiz) had a significant influence on quiz performance, $F(1, 76) = 35.48, p < .001$. Participants scored significantly higher on post-quizzes than on pre-quizzes. Location had a significant effect on performance as well, $F(1, 76) = 4.72, p = .033$, with participants performing significantly better in the lab than online. Additionally, speed first * group interacted with quiz performance, $F(1, 76) = 6.93, p = .010$. Participants performed best when regular reading first and taking a post-quiz ($M = .68, SD = .24$) and worst when regular reading first and taking a pre-quiz ($M = .28, SD = .23$). Reading speed * group interacted with quiz performance, $F(1, 76) = 6.00, p = .017$. Participants scored the best on post-quizzes when reading normally ($M = .72, SD = .22$), followed by performance on post-quizzes when speed reading ($M = .55, SD = .26$). Performance on the pre-quiz was the same for both regular reading and speed reading ($M = .32, SD = .24$). Lastly, there was a significant 3-way interaction between reading speed, speed first, and passage first, $F(1, 76) = 15.62, p < .001$. No other comparisons were significant (all other $ps > .052$). The main effect of reading speed, collapsed across all other variables, is plotted in Figure 3.

Final Test Performance

Final performance is plotted in Figure 4. We conducted a 2 (Speed First: Regular Reading or Speed reading) x 2 (Passage First: Maori or Sharks) x 2 (Location: Lab or Online) x 3 Group (No Quiz, Pre-Quiz, or Post-Quiz) Repeated-Measures ANOVA

with 2 within-participant levels (regular reading final test score and speed-reading final test score). Reading speed had a significant effect on final test performance, $F(1, 105) = 27.17, p < .001$. Participants scored significantly higher on the final test after reading normally ($M = .62, SD = .21$) than after speed reading ($M = .49, SD = .21$). There was a significant 3-way interaction between reading speed * speed first * passage first, $F(1, 105) = 28.10, p < .001$, with participants performing best on the final test when regular reading first and reading the Sharks passage first ($M = .71, SD = .19$). Participants performed worst on the final test when speed reading the Maori passage second ($M = .37, SD = .19$). There was also a significant 3-way interaction between reading speed * passage first * group, $F(1, 105) = 4.06, p = .020$. Participants scored best on the final test when in the pre-quiz condition, reading normally, with the Maori passage first ($M = .64, SD = .27$), while they scored worst when in the no quiz condition, speed reading, with the Maori passage first ($M = .42, SD = .25$). No other comparisons were significant (all other $ps > .069$). Furthermore, given that there was no main effect of group, no post-hoc comparisons were conducted. The main effect of reading speed is plotted in Figure 5 as a factor of group. We collapsed across all other variables for this figure as they were not significant.

General Discussion

Reading comprehension was impaired for speed reading compared to reading normally. Taking a quiz did not help improve comprehension after speed reading. Quizzing did not have a significant effect on comprehension scores for either reading speed. The quizzes themselves showed a difference between speed reading and reading normally (with reading normally resulting in higher quiz scores). There was also an

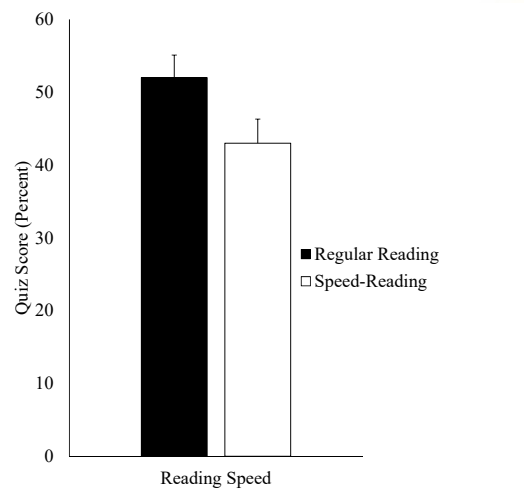


Figure 3. A bar graph showing quiz performance results. Regular reading resulted in significantly better scores than speed reading.

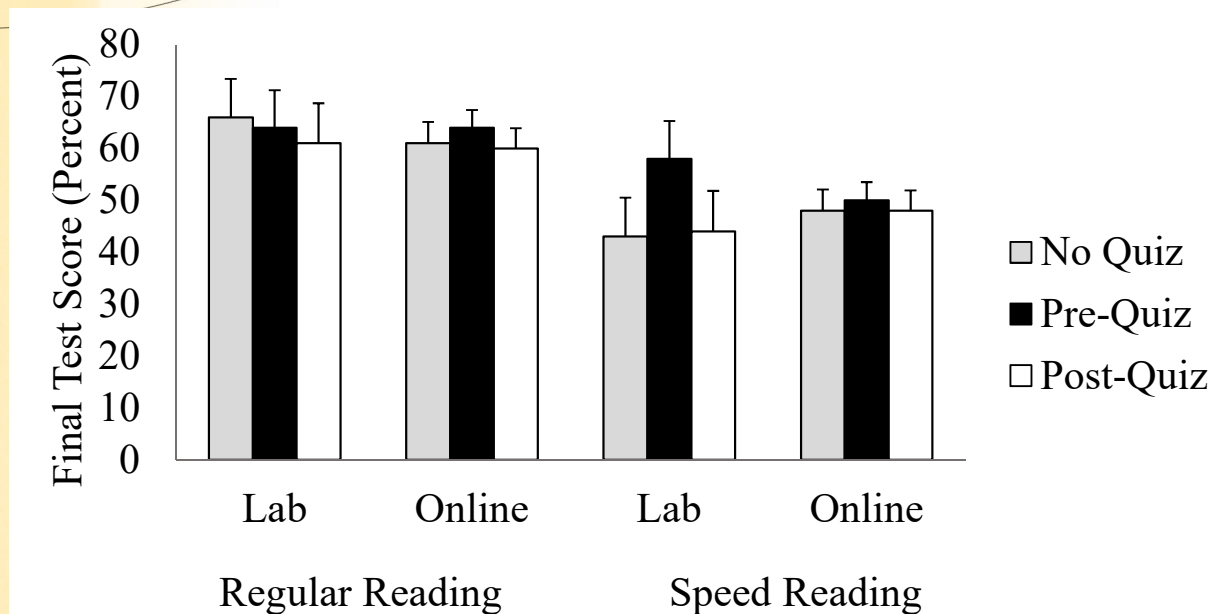


Figure 4. This is a graphic of final test score performance. There was no significant difference between final test scores for participants taking the study in the lab versus online, nor was there a significant difference between quiz type.

influence of location on the quizzes (i.e., participants scored higher on the quizzes when taken in a lab setting); however, this influence did not extend to the final tests.

As hypothesized, reading comprehension was significantly better after reading normally than after speed reading; however, not all the findings were consistent. We had anticipated that the presence of a quiz would influence performance on the final test but found no significant difference between any of the groups. This finding nullified our subsequent prediction of a pre-quiz most benefitting speed reading and a post-quiz most benefitting regular reading, as none of the groups were significantly different from one another.

These findings were partly consistent with prior research, as previous studies have supported the claim that speed reading impairs comprehension (e.g., Acklin & Papesch, 2017; Duggan & Payne, 2009; Rayner et al., 2016; Slowiaczek & Clifton, 1980). However, prior research supported the hypothesis that retrieval practice would positively influence final comprehension scores (e.g., Burns & Vinchur, 1992; Narloch, Garbin, & Turnage, 2006), which was not the case.

Most of the limitations of this study stem from it not being comparable to a typical college course. The passages were short (only 350 words), rather than the length of a characteristically assigned course reading. The quiz and final test were conducted in a single session. In most educational settings, quizzes and tests are not conducted on the same day,

especially not so close together in time. Finally, the final test was not comparable in length to an exam. The final test for this study was only 10 questions, while a standard college exam contains 50 to 100 questions with varying question types. Other limitations of the study pertain to its online nature. As it was taken on a computer, participants could have copied the text for later reference, or they could have simply chosen not to read the passage (or even speed read it despite being instructed not to in the regular reading group). Despite these limitations of an online study, it can be argued that conducting the study online made it more representative of a college setting, as an increasing number of quizzes and exams are given in an online rather than physical format. Future studies could be conducted within the laboratory using eye tracking equipment.

There are directions for subsequent investigation into speed reading and comprehension. A longitudinal study may be useful to learn some of the long-term effects of speed reading on retention. This type of study could help make it more representative of a typical college course. For example, there is commonly interference from other courses and life throughout the semester. Such a study would also allow investigation into the influence on retention of re-reading, or re-speed reading, a text. It would be useful to conduct a study with exams that are longer and more reflective of an actual college exam, including the use of multiple question types (e.g., short-answer, true / false, multiple-choice). In addition to longer exams, longer passages with more context would be more representative

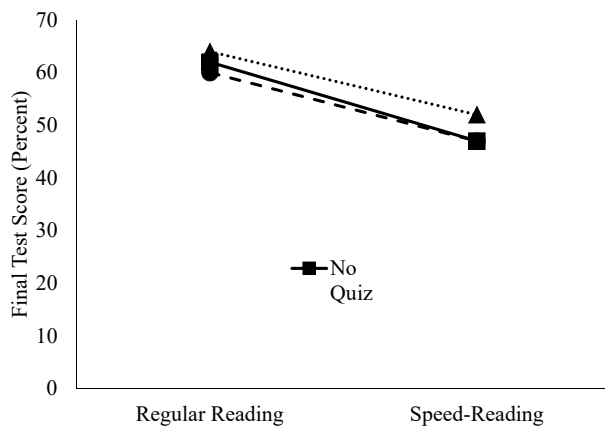


Figure 5. A line graph of final test score results. Regular reading had significantly better scores than speed reading, but there was not a significant difference between groups.

of a typical college reading. These passages would also be better suited for both fact-based and inferential questions. An additional path for examination would include the use of multiple types of reading materials (e.g., novels, textbooks, short stories) to investigate whether the findings apply to all types of reading materials, or only a select few. Another possible direction for future studies would be the impact on comprehension of reading electronic versus printed material. This is a pertinent question currently, as classes are moving toward online formats and online textbooks. This line of study also lends itself to the possibility of a cohort effect, as younger generations have grown up using and reading on computers, whereas older generations were not exposed to computers until they were adults.

This study found that speed reading, even when used in conjunction with retrieval practice, did not attain comprehension levels comparable to those when reading normally. Future studies may be able to find some way to improve comprehension when speed reading; however, as of yet, such a method is unknown and reading normally remains the most effective reading strategy.

References

- Acklin, D., & Papesh, M. H. (2017). Modern speed-reading apps do not foster reading comprehension. *American Journal of Psychology*, 130(2), 183-199.
- Breznitz, Z., & Share, D. L. (1992). Effects of accelerated reading rate on memory for text. *Journal of Educational Psychology*, 84(2), 193-199.
- Burns, D., & Vinchur, A. (1992). Effects of evaluative quizzes on test performance. *Journal of Instructional Psychology*, 19(3), 148-154.
- Duggan, G. B., & Payne, S. J. (2009). Text skimming: The process and effectiveness of foraging through text under time pressure. *Journal of Experimental Psychology*, 15(3), 228-242.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4-58.
- Ebbinghaus, H. (1885). *Über das Gedächtnis*. Leipzig: Duncker & Humblot.
- Hopkins, R., Lyle, K., Hieb, J., & Ralston, P. (2016). Spaced retrieval practice increases college students' short- and long-term retention of mathematics knowledge. *Educational Psychology Review*, 28(4), 853-873.
- Karpicke, J., Butler, A., & Roediger, H. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, 17(4), 471-479.
- Knight, J. B., Ball, B. H., Brewer, G. A., DeWitt, M. R., & Marsh, R. L. (2012). Testing unsuccessfully: A specification of the underlying mechanisms supporting its influence on retention. *Journal of Memory and Language*, 66(4), 761-746.
- Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 989-998.
- Murre, J. M., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PloS one*, 10(7), e0120644.
- Narloch, R., Garbin, C., & Turnage, K. (2006). Benefits of prelecture quizzes. *Teaching of Psychology*, 33(2), 109-112.
- Rayner, K., Schotter, E. R., Masson, M. J., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest*, 17(1), 4-34.
- Roediger, III, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181-210.
- Slowiaczek, M. L., & Clifton, Jr., C. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, 19(5), 573-582.
- Smith, M., & Karpicke, J. (2014). Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory*, 22(7), 784-802.
- Vaughn, K. E., & Rawson, K. A. (2012). When is guessing incorrectly better than studying for enhancing memory? *Psychonomic Bulletin & Review*, 19(5), 899-905.

APPENDIX A.

Reading passage on sharks

Sharks belong to a family of fish that have skeletons made of cartilage, a tissue more flexible and lighter than bone. They breathe through a series of five to seven gill slits located on either side of their bodies. All sharks have multiple rows of regenerative teeth, and while they lose teeth on a regular basis, new teeth continue to grow in and replace those they lose. The earliest known sharks date back to more than 420 million years ago. Acanthodians are often referred to as “spiny sharks”; though they are not part of Chondrichthyes proper, they are a paraphyletic assemblage leading to cartilaginous fish as a whole. Since then, sharks have diversified into over 500 species. They range in size from the small dwarf lantern shark, a deep sea species of only 17 centimeters (6.7 in) in length, to the whale shark, the largest fish in the world, which reaches approximately 12 meters (40 ft) in length. Sharks are found in all seas and are common to depths of 2,000 meters (6,600 ft). They generally do not live in freshwater although there are a few known exceptions, such as the bull shark and the river shark, both of which can survive and be found in both seawater and freshwater. Sharks have a covering of dermal denticles that protects their skin from damage and parasites in addition to improving their fluid dynamics. Shark ‘skin’ is made up of a series of scales that act as an outer skeleton for easy movement and for saving energy in the water. The upper side of a shark is generally dark to blend in with the water from above and their undersides are white or lighter colored to blend in with the lighter surface of the sea from below. This helps to camouflage them from predators and prey. Well-known species such as the great white shark, tiger shark, blue shark, Mako, thresher shark, and hammerhead shark are apex predators—organisms at the top of their underwater. They regulate the populations of species below them. Despite this, many shark populations are threatened by human activities. [Word Count: 350]

Sources:

<https://en.wikipedia.org/wiki/Shark>

<https://defenders.org/sharks/basic-facts>

APPENDIX B.

Reading passage on Māori

The Māori are the indigenous Polynesian people of New Zealand. Evidence from archaeology, linguistics, and physical anthropology indicates that the first settlers came from east Polynesia around 1280 CE, at the end of the medieval warm period, and became the Māori. Māori oral history describes the arrival of ancestors from Hawaiki (the mythical homeland in tropical Polynesia), in large ocean-going waka. Migration accounts vary among tribes (iwi), whose members may identify with several waka in their genealogies (whakapapa). In the last few decades, research has allowed an estimate to be made of the number of women in the founding population—between 50 and 100. Atholl Anderson concluded from analysis that the ancestors of Polynesian women came from Taiwan while those of Polynesian men came from New Guinea. Over several centuries in isolation, the Polynesian settlers developed a unique culture, with their own language, a rich mythology, and distinctive crafts and performing arts. Early Māori formed tribal groups based on eastern Polynesian social customs and organization. Horticulture flourished using plants they introduced; later, a prominent warrior culture emerged. The arrival of Europeans to New Zealand, starting in the 17th century, brought enormous changes to the Māori way of life. Māori people gradually adopted many aspects of Western society and culture. Initial relations between Māori and Europeans were largely amicable, and with the signing of the Treaty of Waitangi in 1840, the two cultures coexisted as part of a new British colony. Rising tensions over disputed land sales led to conflict in the 1860s. Social upheaval, decades of conflict and epidemics of introduced disease took a devastating toll on the Māori population, which fell dramatically. By the start of the 20th century, the Māori population had begun to recover, and efforts have been made to increase their standing in wider New Zealand society and achieve social justice. Traditional Māori culture has thereby enjoyed a significant revival, which was further bolstered by a Māori protest movement that emerged in the 1960s. Since the mid-19th century, Māori have gained 7 seats in Parliament and have played a role in the governing of New Zealand. [Word Count: 350]

Sources:

<https://www.britannica.com/topic/Maori>

https://en.wikipedia.org/wiki/M%C4%81ori_people

APPENDIX C.

Reading comprehension test on sharks

- Shark skeletons are made from what?
 - Bone
 - Keratin
 - Cartilage**
 - Enamel
- What is the smallest type of shark?
 - Spiny shark
 - Dwarf lantern shark**
 - Mako shark
 - River shark
- How far back do the first sharks date?
 - 420 million years ago**
 - 750 million years ago
 - 120 million years ago
 - 360 million years ago
- Which type of shark can survive in both seawater and freshwater?
 - Mako shark
 - Bull shark**
 - Thresher shark
 - Blue shark
- How many different species of shark are there?
 - 700
 - 300
 - 200
 - 500**
- Why do sharks have a light underbelly and dark backside?
 - To scare off predators
 - To attract prey
 - To blend in**
 - For no reason except genetics
- The smallest shark is how long?
 - 17 centimeters**
 - 25 centimeters
 - 8 centimeters
 - 43 centimeters
- At what depth is it common for sharks to swim?
 - 500 meters
 - 3000 meters
 - 2000 meters**
 - 1800 meters
- What do the dermal denticles protect sharks from?
 - Parasites**
 - The sun
 - Temperature
 - Infection
- The largest shark is how long?
 - 20 meters
 - 8 meters
 - 40 meters
 - 12 meters**

Note. Questions 1, 3, 5, 7, and 9 served as quiz questions.

APPENDIX D.

Reading comprehension test on the Maori

- From where did the Maori people originate?
 - Polynesia**
 - Malaysia
 - China
 - Madagascar
- The Maori and Europeans had an amicable relationship due to...
 - Nothing, the relationship was not amicable
 - The Treaty of Hawaiki
 - The Treaty of Waitangi**
 - The Treaty of Whakapaka
- The Maori protest movement occurred during which years?
 - 1960s**
 - 1980s
 - 1950s
 - 1990s
- The people that would become the Maori settled in New Zealand in what year?
 - 1280 CE**
 - 1100 CE
 - 1370 CE
 - 980 CE
- Conflict between the Maori and the Europeans was over what?
 - Civil rights
 - Trade disagreements
 - Independence
 - Land sales**
- When did the conflict between the Maori and the Europeans occur?
 - The 1320s
 - The 1930s
 - The 1540s
 - The 1860s**
- The Maori have how many seats in Parliament?
 - 0
 - 3
 - 7**
 - 15
- When did the Europeans arrive in New Zealand?
 - The 1600s
 - The 1800s
 - The 1700s**
 - The 1400s
- The people that would become the Maori came to New Zealand in what?
 - Wakas**
 - Long boats
 - Iwi
 - Ships
- The Maori had what type of culture?
 - Hunter / Gatherer
 - Warrior**
 - Agricultural
 - Nomadic

Note. Questions 1, 3, 5, 7, and 9 served as quiz questions.

Pen Chromaticity's Effect on the Recall of Details Following Lecture Annotation

Jessica Raines.

Faculty mentor: Kathleen Fuegen
Psychology

KEYWORDS:

color, red, academic performance, gender differences, psychological functioning

Jessica Raines

Jessica graduated cum laude from Northern Kentucky University with a Bachelor of Arts degree in psychology and a minor in history. She conducted this research as an Honors in Psychology Thesis and continued her exploration during an independent study guided by Dr. Kathleen Fuegen. With her help, Jessica was able to present her findings at NKU's Celebration of Student Success in 2017 and the annual meeting of the Midwestern Psychological Association in 2019. Jessica is currently pursuing her master's degree in clinical mental health counseling at Northern Kentucky University.

Abstract

When you think of failure, what color comes to mind? If you thought red, you are not alone. Previous studies have indicated that the color red, when seen on exam booklets, worsens academic performance. Nonetheless, some studies find no evidence that red worsens performance, or that its effects depend on the gender of the test taker. The purpose of this study was to investigate the effect of pen color used while taking notes during a lecture on subsequent test scores, perceptions of test difficulty, and note-taking behavior. The data were analyzed using 2 (pen color) x 2 (participant gender) ANOVAs. The results were partially consistent with the hypothesis: we found that use of a red pen resulted in worsened performance compared to use of a blue pen, but only among males. Among females, use of a red pen enhanced performance relative to a blue pen. We found this pattern of results on questions measuring recall but not recognition. We discuss the implications of pen color choice for studying and grading. If students are more aware of the effect of color on academic performance, they may be able to retain and recall more information simply by changing their pen color.

Introduction

The vast majority of student life is simply retaining information. Regardless of the class being taken, one's grade level in school, or one's grade point average, absorbing facts, ideas, and data is crucial to the learning experience. However, students may struggle to retain what they learn in class. Retention is often aided by activities like taking notes. While recently it has become common to take notes by typing on a computer, upon comparing this technique to producing hand-written notes, Mueller and Oppenheimer (2014) found that typing can lead to shallower processing, resulting in worsened performance on conceptual essay questions. If handwritten notes optimize retention, does the color of a student's pen also aid learning?

Researchers have investigated environmental effects on academic performance for decades. Michael and Jones (1955) were intrigued by the idea that color coordinating class exams may inadvertently hinder academic performance. They examined five unspecified colors of test booklets over four college courses. They reported that there were no statistically significant differences in the average scores on either multiple-choice or true-false items. The only exception was one business administration course where they found a significant difference between white and yellow test booklets on multiple-choice questions. In a more recent study examining how color affects academic performance, Sinclair, Soldat, and Mark (1998) collected data during a normally scheduled exam at a community college. They greatly improved upon Michael and Jones' study by gathering data from a larger, more diverse sample and focusing on one specific area of performance. They randomly assigned students to receive either a red or blue exam booklet with identical formatting, inside and out. They found that those using a blue test booklet performed significantly better on both easy and difficult questions, compared to those using a red test booklet (Sinclair et al., 1998).

Since these initial studies, some researchers have found that color influences academic performance, while others have not. The inconsistency in findings may depend on the nature of the test (e.g., high stakes vs. low stakes; easy vs. difficult) or the type of retrieval measured (e.g., recognition vs. recall). Most studies examine whether red influences academic performance, not which aspects of performance. The purpose of this study was to examine whether pen color (red vs. blue) influences academic performance, and, if so, whether the effect occurs for both easy and difficult test items and for both recall and recognition test items. We also examine whether pen color may influence the amount of notes students take by hand or students' perceptions of the ease or difficulty of the test. Because some studies show that females and males respond differently to color, we explore whether gender interacts with pen color to influence academic performance.

Lastly, we explore whether favorite color influences how students respond to red vs. blue pens.

Effects of color on cognitive performance

In other research conducted with students, Elliot, Maier, Moller, Friedman, and Meinhardt (2007) asked 71 undergraduates to solve 15 moderately difficult anagrams in five minutes. The only difference between the practice test and the "real" trial was that the trial included a number written in either red, green, or black at the top of the test, to identify the participant. Participants verified the number that was given to them. At the end of the study, participants were asked what they believed the study was about and to write down the color of their number as a check on the manipulation. Results indicated that those in the red condition solved significantly fewer anagrams than those in the green and black conditions (Elliot et al., 2007).

In a conceptual replication (Experiment 2), Elliot et al. (2007) changed the neutral color from black to white. They also changed the color manipulation, so that the exposure was separate from the performance period. Furthermore, they used a subscale of an IQ test rather than anagrams as the dependent measure. They found that participants in the red condition still performed worse than those in the green and white conditions. In Experiment 3 they changed the neutral color from white to gray, and they shortened the exposure to the colored number manipulation. Most importantly, they shifted from the laboratory to a high school classroom. The researchers still found a significant difference between conditions. Those in the red condition performed worse than those in the green and the gray conditions (Elliot et al., 2007).

To further examine the effect of the color red on cognition, Lichtenfeld, Elliot, Maier, and Pekrun (2009; Experiment 1) conducted a study in which they analyzed the effects of the word (i.e., "red") itself. This being a German study, participants were simply told they were either in the "Rot" (red) or "Ort" (place) anagrammatic group. Then participants completed a simple analogy IQ test. Those in the red group performed significantly worse than those who were assigned to the place group. This effect was replicated in a second experiment where the "place" condition was replaced with a "gray" condition.

To summarize, multiple experiments have found that red impairs cognitive performance relative to other colors. The effect has been found with multiple measures of performance: anagrams, an IQ test, and university exams. The exact mechanism for this effect remains unclear, though some research suggests there may be physiological changes associated with color exposure. Elliot, Payen, Brisswalter, Cury, and Thayer (2011) found that young adults exposed to the word "test" in a red font showed decreased high-frequency heart rate variability, whereas those exposed to the word "test" in a blue or gray font showed

increased high-frequency heart rate variability. Decreased high-frequency heart rate variability was associated with worse performance on an IQ test.

Culture and color effects

The results of three studies conducted in East Asia suggest that the effect of red on cognitive performance may generalize across cultures. Tanaka and Tokuno (2011) measured 63 Japanese undergraduates' willingness to choose easy or difficult analogies while the test proctor wore either a red, white, or green shirt under a lab coat during the experiment. Participants exposed to a red shirt chose to complete more easy analogies than participants exposed to a white or green shirt. The researchers concluded that red might activate avoidance motivation. In a study conducted with 58 Chinese undergraduates, Shi, Zhang, and Jiang (2015) examined whether seeing red or blue letters on a computer screen affected performance scores on a Chinese idiom test. They found that participants in the red condition performed worse than those in the blue condition. This effect occurred despite the cultural positivity surrounding the color red in China.

Additionally, Xia, Song, Wang, Tan & Mo (2016) examined the effect of word color on test performance in 125 Chinese undergraduate students. Participants identified whether pairs of the presented words were identical in spelling, capitalization, and punctuation. The researchers manipulated the difficulty of the task. The simple task had fewer letters (20-50), whereas the difficult task had more letters (50-100). Words appeared on a red, blue, or gray background. Xia and colleagues found a significant interaction between task difficulty and color indicating that performance differed by color. In the simple task condition, the accuracy rate of those in the red background condition was higher than the accuracy rate of those in the blue and gray background conditions. In the difficult task condition, the accuracy rate of those in the blue background condition was higher than that of participants in the red and gray background conditions.

Gender and color effects

The previous studies found no effect of participant gender on cognitive or physiological outcomes or treated participant gender as a covariate in the analyses. However, some studies have found that women and men react differently to red. Gnambs, Appel, and Batinic (2010; Experiment 1) quizzed 131 Austrian undergraduates on a general knowledge test. During this test, a progress bar was present at the top of each participant's computer screen. The bar was either red or green. The researchers then measured how many test questions participants answered correctly. They found that males' performance was worse in the red condition than the green condition. However, females' performance was unaffected by color. These effects remained significant even after controlling for participant age and fluid intelligence. The researchers

theorized that females were unaffected by the color of the progress bar because women prefer hues of red such as pink, fuchsia, and maroon (Cohen, 2013). Women may habituate to the color red. Thus, red may not impair their performance. Furthermore, because males prefer blue hues, they may not be habituated to the color red. Thus, their performance was impaired in the red condition (Gnambs et al., 2010).

Gnambs, Appel, and Kaspar (2015) replicated the finding of gender differences in reactions to red. One hundred and ninety Austrian teenagers were given a booklet describing medieval dining customs. The booklet title was either encased with a red rectangle or a gray rectangle. Each subsequent page had a corresponding color cue at the top. Next, they completed a knowledge test over the material in the booklet. The title on the test cover was either encased with a red rectangle or a gray rectangle, and each subsequent page of the knowledge test had a corresponding color cue at the top. The researchers found that, for boys, repeated color exposure influenced test performance more so than color presentation during just a single (i.e., learning) phase. Boys' performance was worst when red was repeatedly presented. However, girls' performance was best when red was repeatedly presented.

Context and color effects

Not all studies show that red impairs academic performance. For example, Mehta and Zhu (2009) reported that red facilitated solving anagrams of avoidance words, whereas blue facilitated solving anagrams of approach words. However, Steele (2014) conducted a direct replication with a considerably larger sample and was unable to replicate these effects. He found that neutral anagrams were solved faster than both avoidance and approach anagrams. Additionally, he found no relationship between color and speed of solving anagrams (Steele, 2014). Mehta and Zhu (2009) did not pretest the anagrams used in their study, and they were not of equal length across conditions.

Other studies conducted in college classrooms, rather than laboratories, have found no significant effect of color on academic performance. In one recent study (Smajic, Merritt, Banister, & Blinbry, 2014), college students took identical multiple-choice exams except for the booklet color which was either red or green. The color of the booklet produced no significant difference in performance, nor was color correlated with anxiety or affect. Moreover, Larsson and von Stumm (2014) found no significant differences in adults' performance on cognitive ability tests based on whether they had been previously exposed to a string of letters and numbers in red or green. Similarly, Arthur, Cho, and Muñoz (2016) indicated that, after comparing three different archival data sets of examinations, the color of the exam booklet (either red or green) did not affect test scores. The trend in these findings is that the overall test scores did not significantly differ based on test booklet color. Nonetheless, differences may be found on a

deeper level with question difficulty or in recall as opposed to the simple recognition required for multiple-choice questions (e.g., Sinclair et al., 1998; Tanaka & Tokuno, 2011; Xia et al., 2016). Reexamining their findings could illuminate potential missed analyses. Furthermore, gender differences may also be influencing their results (Cohen, 2013). Thus, a better approach to examining the effects of color on academic performance would involve measuring more than overall test performance. It is possible that color affects only some aspects of academic performance.

Present Studies

We conducted pretesting to determine the ease or difficulty of test items created by the first author. In the main study, participants are randomly assigned to use a red or blue pen while taking notes and completing a test over lecture material.

Pretesting

Methods

Participants. Twenty-one Northern Kentucky University undergraduate students (16 females, five males) recruited from Sona participated in pretesting. The students received extra course credit. The average age of this sample was 18.81 years old ($SD = 1.75$) and 76% of the participants were women. Ninety percent were Caucasian, and 10% were African American. Right-handed participants made up 91% of the sample. One participant was left-handed and one reported being ambidextrous.¹

Materials and Procedure. After providing informed consent, participants began the study. Participants wore headphones while a thirteen-minute and forty-one second video lecture about the Bay of Pigs from the Khan Academy (2011) played. The researcher asked participants to take notes during the lecture on white loose-leaf paper with a black pen, as they would during any other lecture. Following the lecture, the researcher collected the notes and administered tests on white paper. The test measured knowledge of the material provided in the video with 12 short answer and 16 multiple-choice questions. The first author created the test after reviewing key points of the Khan Academy video. Participants were instructed to answer the questions to the best of their ability with the same black pen. Demographic questions were included at the end of the test.

Results

On average, participants answered 9.00 ($SD = 2.21$) of the short answer questions correctly and 13.14 ($SD = 2.37$) of the multiple-choice questions correctly. There were no significant

differences in performance on short answer, $t(19) = 1.42$, $p = .171$, or multiple-choice, $t(19) = 1.25$, $p = .226$, questions between males and females.

Main Study

Methods

Participants. Eighty-nine Northern Kentucky University undergraduate students recruited from Sona participated in the main study.² Students received extra course credit. The average age of the sample was 19.06 years old ($SD = 3.04$ years), and 48.3% of the participants were women. Seventy-nine percent were Caucasian, 10% were African American, 6% identified as biracial, 3% were Asian, and 2% were Hispanic. Eighty-two percent reported being right-handed, while 14% reported being left-handed. The remaining three participants reported being ambidextrous.

Materials. Test questions were classified as easy or difficult if the majority of pretest participants answered them correctly (easy) or incorrectly (difficult). Twenty-four items were classified as easy, and four items were classified as difficult. The first author created a five-item perception questionnaire to examine how participants felt about their test performance. Participants rated how challenging they found the 1) test, 2) short answer questions, and 3) multiple-choice questions. Each rating was made on a scale ranging from 1 (very easy) to 10 (extremely challenging). They also indicated how many questions they struggled to answer and how many questions they found easy to answer (open-ended responses). A demographics form was used to obtain information regarding participant age, gender, year in school, race, dominant hand, color blindness, and color preference.

Procedure. Prior to the start of each session, the researcher placed white loose-leaf paper in a plain white manila envelope. Participants were randomly assigned a red or blue colored pen. One pen was placed in each envelope. The researcher shuffled the envelopes and placed one at each desk.

Participants completed the study in a laboratory in small groups of up to seven. After providing informed consent, participants viewed the same thirteen-minute and forty-one second video that participants in pretesting viewed. Next, the researcher advised the participants to take notes, as there would be a test following the video. At the conclusion of the video, the researcher collected participants' notes and administered the test. Participants completed the white paper test using their assigned pen. Following the test, the researcher administered the perception questionnaire and a demographics form. Lastly, the researcher provided participants a debriefing form.

1. One additional student participated in the study but did not use the assigned pen. This person's data are excluded from analysis.

2. Two additional students participated in the study. One participant reported being color blind, and the other did not use the assigned pen. Their data are excluded from all analyses.

Results

Test performance. We performed a 2 (pen color) \times 2 (participant gender) analysis of variance (ANOVA) on each of the following test performance variables: number of correct answers to multiple-choice questions, number of correct answers to short answer questions, number of correct answers to difficult questions, and number of correct answers to easy questions. The analysis of multiple-choice questions revealed no significant effect of pen color, $F(1, 85) = 0.02$, $p = .901$, participant gender, $F(1, 85) = 0.49$, $p = .485$, or the interaction, $F(1, 85) = 1.83$, $p = .180$.

The analysis of short answer questions showed that the effect of pen color was not significant, $F(1, 85) = 0.46$, $p = .501$. There was also no significant difference between males and females on short answer scores, $F(1, 85) = .01$, $p = .929$. However, the pen color \times gender interaction was significant, $F(1, 85) = 5.47$, $p = .022$, $\eta_p^2 = .06$. A follow up independent-samples t-test on pen color was significant for females, $t(35.78) = -2.46$, $p = .019$. Females who used a red pen ($M = 9.87$; $SD = 1.69$) performed better than females who used a blue pen ($M = 8.40$; $SD = 2.16$). The independent-samples t-test on pen color was not significant for males, $t(27.04) = 1.03$, $p = .314$. The means were in the opposite direction: males who used a red pen tended to perform worse ($M = 8.77$; $SD = 3.46$) than males who used a blue pen ($M = 9.58$; $SD = 1.38$).

We next examined performance on easy and difficult test questions. On easy test questions, neither the main effect of pen color, $F(1, 85) = 0.28$, $p = .599$, nor the main effect of participant gender, $F(1, 85) = .00$, $p = .982$, were statistically significant. However, the pen color \times gender interaction was significant, $F(1, 85) = 4.34$, $p = .040$, $\eta_p^2 = .05$. A follow up independent-samples t-test on pen color was not significant for either females, $t(41) = -1.50$, $p = .141$, or for males, $t(28.83) = 1.52$, $p = .140$. Similar to the pattern found for short answer questions, males tended to perform worse when using a red pen ($M = 19.18$; $SD = 5.62$) than a blue pen ($M = 21.17$; $SD = 2.57$). Females tended to perform better when using a red pen ($M = 20.78$; $SD = 2.37$) than a blue pen ($M = 19.60$; $SD = 2.80$).

On difficult test questions, the effect of pen color was significant, $F(1, 85) = 7.34$, $p = .008$, $\eta_p^2 = .08$. Those in the red condition ($M = 2.38$; $SD = 1.17$) performed better than those in the blue condition ($M = 1.75$; $SD = 1.16$). However, neither the effect of participant gender, $F(1, 85) = 3.26$, $p = .075$, nor the interaction, was significant, $F(1, 85) = 1.42$, $p = .237$.

Note-taking and test perceptions. Because taking notes was such a large part of this study, it was important to analyze the amount and content of what participants wrote. Because the video included many historical details, it was possible to count the number of dates, names, and countries identified in the notes, as well as the total number of words written. We

conducted a 2 (pen color) \times 2 (participant gender) ANOVA on each of these outcomes. Pen color did not have a significant effect on the number of dates, names, countries, or word total, all $ps > .30$. The interactions were not significant. However, there was a main effect of participant gender on three of the four outcomes: names written, $F(1, 85) = 7.82$, $p = .006$, $\eta_p^2 = .08$; countries written, $F(1, 85) = 10.73$, $p = .002$, $\eta_p^2 = .11$; and word total, $F(1, 84) = 13.50$, $p < .001$, $\eta_p^2 = .14$. Women ($M = 11.56$, $SD = 3.16$) wrote more names than men ($M = 9.46$, $SD = 3.77$). Women ($M = 12.91$, $SD = 5.19$) wrote more countries than men ($M = 9.43$, $SD = 4.85$). Women ($M = 169.40$, $SD = 78.76$) wrote more words than men ($M = 117.98$, $SD = 49.67$). Only the number of names written was significantly correlated with test performance: the more names written, the better the performance on short answer questions, $r = .37$, $p < .001$, multiple-choice questions, $r = .25$, $p = .018$, and easy questions, $r = .35$, $p = .001$. The correlation between number of names and performance on difficult questions was positive but not significant, $r = .20$, $p = .066$.

We also examined participants' perceptions of the test. These perception questions asked participants to report how many questions they struggled to answer, how many questions they found easy to answer, and how challenging they found the entire test, short answer, and multiple-choice questions. The latter three items were averaged (internal consistency reliability coefficient = .89). We ran a 2 (pen color) \times 2 (participant gender) ANOVA on these three variables. There were no statistically significant effects. Nonetheless, perceptions tended to mirror performance. Females in the blue condition struggled to answer more questions than females in the red condition. Conversely, males in the red condition struggled to answer more questions than males in the blue condition. Females in the red condition reported being less challenged on the test than females in the blue condition. Conversely, males in the blue condition reported being less challenged on the test than males in the red condition.

Lastly, we explored whether participants' favorite color differed by gender or affected test performance. Across all participants, 41.6% reported shades of blue, 15.7% reported shades of red, 16.9% reported purple, and the remaining 24.7% reported other colors (orange, green, white, black, etc.) as their favorite color. A Chi-Square test comparing color preferences for females and males was not significant, $\chi^2(df=1) = 2.92$, $p = .088$. Both females (60%) and males (82%) reported a shade of blue as their favorite color. Many more females (40%) than males (19%) reported a shade of red as their favorite color. We examined whether participants who were randomly assigned to use a pen that matched their favorite color tended to perform better on the test. The interactions were not significant.

Discussion

Based on previous research, we expected that participants in the red condition would perform worse on the test than participants in the blue condition. We did find some support for this expectation among male participants. Males performed slightly worse on short answer questions and easy questions in the red condition than in the blue condition. However, both males and females performed better on difficult questions when using a red pen. On short answer questions, females in the red condition performed significantly better than females in the blue condition. Also, males in the blue condition performed better than females in the blue condition.

In analyzing the handwritten notes of participants, we found that females wrote significantly more than males, regardless of pen color, in all four categories: dates, names, countries, and word total. Nonetheless, there was no significant gender difference in test performance. Furthermore, the analysis of participant perceptions of the test revealed no significant interaction between pen color and participant gender. Nonetheless, there was a trend, such that the more items participants reported struggling to answer, the worse they performed on the exam. Though not statistically significant, females in the red condition and males in the blue condition reported struggling less than females in the blue condition and males in the red condition.

By including both multiple-choice and short answer questions on the test, we were able to examine whether color affected performance on both recognition and recall questions. Although recognition (multiple-choice) was not affected by pen color or gender, we found a significant interaction between pen color and gender for recall (short answer), which requires a deeper thought process. This may suggest that color influences performance when people must write out answers. Through pretesting, the first author was able to classify test items as easy or difficult. The same interaction between pen color and gender that appeared for short answer questions appeared for easy questions. However, for difficult questions, participants in the red condition performed better than participants in the blue condition, regardless of gender.

We were able to partially replicate the finding of gender differences in the effect of color on test performance (Gnambs et al., 2010, 2015). Gnambs et al. (2010) found that male participants performed worse in the red condition than the green condition, whereas females showed no difference in performance. They speculated that, based on previous research about color preference (Cohen, 2013), females have become habituated to the color red because of gender stereotypes. Being brought up in a pink environment has forced exposure to the color red among females. Because of this, reddish hues lose their predominantly negative connotation among females. This line of reasoning fits with the findings of our study: females'

performance was not negatively affected by the color red, but males' performance was. In fact, we found that females seemed to perform better in the presence of red, though this effect may be due to the comparison group: Gnambs and colleagues compared red with green, but we compared red with blue. Expanding on Gnambs and colleagues' speculation about habituation, we posit that this works with males as well. Males are mostly exposed to blues. Although blue does not have a cultural negativity associated with it, females in this study were negatively influenced by blue. From an early age, males are exposed more often than females to the color blue. They become familiar and comfortable with the color. This may cause them to feel relaxed, free of stress, and confident in the presence of blue. On the other hand, red produces anxiety in males that is not seen in females.

We were interested in red and blue because these are the colors that have produced the most inconsistency in results. Few studies have directly compared red with blue. Shi et al. (2015) and Sinclair et al. (1998) found that those in the blue condition performed better, but they did not test for gender differences. Partially consistent with our findings, Gnambs et al. (2010) found that males performed better when exposed to blue than red, though females' performance was unaffected by color. Those studies and ours suffer from the lack of a control condition. We do not know whether red would impair performance in males compared to black. Conversely, we do not know whether red would enhance performance in females compared to black. We may draw tentative inferences from the results of pretesting. During pretesting, all participants completed the test using a black pen. The mean number of short answer questions answered correctly while using a black pen (9.00) suggests that, among males in the main study, blue enhanced performance and red impaired performance. Among females in the main study, red enhanced performance and blue impaired performance, relative to black. Nonetheless, a comparison of red, blue, and black within the same experiment is needed before any definitive conclusions can be drawn.

Limitations and future directions. Controlling the research environment is always beneficial (Elliot, 2015). However, it is not easy. It would be ideal to control what participants wear, so that others are not exposed to one color more than another. In the laboratory where the first author collected data, the desk chairs were orange. Because this was a psychology study that took place in a laboratory, this test was low-stakes. Results may vary in a high-stakes classroom environment. Further research could examine gender and color in both low and high-stakes environments (e.g., Arthur et al., 2016; Larsson & von Stumm, 2015). It is possible that in a high-stakes situation, participants may be able to suppress the connotations a color might give them in order to succeed. An additional limitation involves the way that test item difficulty was operationalized. The first author classified questions as easy or difficult if the majority of pretest

participants answered them correctly (easy) or incorrectly (difficult). There are alternative ways to classify test items as being easy or difficult. With a different classification scheme, future results may differ. Also, a larger pretesting sample would be ideal. Lastly, on the perceptions questionnaire, participants were asked to give a number rating in each section. However, many gave results such as, "most", "many", "a few", "a couple", and "some". Modifying the instructions may produce more usable data.

Conclusion. This research is important for all aspects of academia. While this does not necessarily mean that students who write more will have better grades, it is likely that they will recognize more of the key points. Because of this, male students can avoid using red as a precaution, whereas female students will likely not be affected. It is always better to take too many notes rather than not enough. With more research, we can discover why these colors are influencing males and females differently. Additionally, teachers can learn from this too. If red does encourage avoidance through its negative connotation, not using red to grade papers and exams could be beneficial to students, particularly males. It is possible that students would be more likely to go back and learn from their mistakes, read over the questions again, and analyze their wrong answers, if their pages were not completely marked up in red. A neutral color (e.g., black or gray) may have no harmful effect on either males or females.

References

- Arthur, W. J., Cho, I., & Muñoz, G. J. (2016). Red vs. green: Does the exam booklet color matter in higher education summative evaluations? Not likely. *Psychonomic Bulletin & Review*, 23(5), 1596-1601. doi:10.3758/s13423-016-1009-6
- Cohen, P. N. (2013). Children's gender and parents' color preferences. *Archives of Sexual Behavior*, 42, 393-397. doi:10.1007/s10508-012-9951-5
- Elliot, A. J. (2015). Color and psychological functioning: A review of theoretical and empirical work. *Frontiers in Psychology*, 6:368 doi: 10.3389/fpsyg.2015.00368
- Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of Experimental Psychology: General*, 136(1), 154-168. doi:10.1037/0096-3445.136.1.154
- Elliot, A. J., Payen, V., Brisswalter, J., Cury, F., & Thayer, J. F. (2011). A subtle threat cue, heart rate variability, and cognitive performance. *Psychophysiology*, 48, 1340-1345. doi:10.1111/j.1469-8986.2011.01216.x
- Gnambs, T., Appel, M., & Batinic, B. (2010). Color red in web-based knowledge testing. *Computers in Human Behavior*, 26(6), 1625-1631. doi:10.1016/j.chb.2010.06.010
- Gnambs, T., Appel, M., & Kaspar, K. (2015). The effect of the color red on encoding and retrieval of declarative knowledge. *Learning and Individual Differences* 42, 90-96. doi:10.1016/j.lindif.2015.07.017
- Kahn, S. (Producer). (2011, April 28). Bay of pigs invasion. [Audio podcast]. Retrieved from <https://www.khanacademy.org/humanities/world-history/euro-hist/cold-war/v/bay-of-pigs-invasion>.
- Larsson, E. E. C., & von Stumm, S. (2015). Seeing red? The effect of colour on intelligence test performance. *Intelligence*, 48, 133-136.
- Lichtenfeld, S., Elliot, A. J., Maier, M. A., & Pekrun, R. (2009). Fertile green: Green facilitates creative performance. *Personality and Social Psychology Bulletin*, 38(6), 784-797. doi:10.1177/0146167212436611
- Mehta, R. & Zhu, R. (2009). Blue or red? Exploring the effect of color on cognitive task performance. *Science*, 323, 1226-1229. doi:10.1126/science.1169144
- Michael, W. B., & Jones, R. A. (1955). The influence of color of paper upon scores earned on objective achievement examination. *The Journal of Applied Psychology*, 39(6), 447-450.
- Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, 25(6), 1159-1168. doi:10.1177/0956797614524581
- Shi, J., Zhang, C., & Jiang, F. (2015). Does red undermine individuals' intellectual performance? A test in China. *International Journal of Psychology*, 50(1), 81-84. doi:10.1002/ijop.12076
- Sinclair, R. C., Soldat, A. S., & Mark, M. M. (1998). Affective cues and processing strategy: Color-coded examination forms influence performance. *Teaching of Psychology*, 25(2), 130-132.
- Smajic, A., Merritt, S., Banister, C., & Blinbry, A. (2014). The red effect, anxiety, and exam performance: A multistudy examination. *Teaching of Psychology*, 41(1), 37-43.
- Steele, K. M. (2014). Failure to replicate the Mehta and Zhu (2009) color-priming effect on anagram solution times. *Psychonomic Bulletin and Review*, 21, 771-776. doi:10.3758/s13423-013-0548-3
- Tanaka, A., & Tokuno, Y. (2011). The effect of the color red on avoidance motivation. *Social Behavior and Personality*, 39(2), 287-288. doi:10.2224/sbp.2011.39.2.28
- Xia, T., Song, L., Wang, T. T., Tan, L., & Mo, L. (2016). Exploring the effect of red and blue on cognitive task performances. *Frontiers in Psychology*, 7:784. doi: 10.3389/fpsyg.2016.00784

Using High School Financial Literacy Education to Predict Future Income: A Story of Selection Bias

Tyler Kent.

Faculty Mentor: Abdullah Al-Bahrani
Economics

KEYWORDS:

NFCS, Financial literacy, Economics of Education

JEL Codes: G40, I26, D14

Tyler Kent

Tyler Kent graduated from Northern Kentucky University with a bachelors of science in Economics in December of 2019. During his time at NKU he served as a teaching and research fellow for the Center for Economic Education under the guise of Dr. Abdullah Al-Bahrani. Tyler would like to thank Dr. Al-Bahrani for his dedication to Economics and to his students. Tyler taught high school financial literacy courses in the Northern Kentucky area as well as preformed research into financial literacy retention rates of Americans, as well as unemployment benefit usage patterns of Americans. Tyler accepted a position at Fidelity Investments after graduation as a workplace investing service representative.

Abstract

I use data from the National Financial Capabilities Study for 2009 to 2015 to analyze the effect high school financial literacy education has on future income levels. I use Ordinary Least Squared regression to measure this relationship. Finding evidence of selection bias with a negative coefficient on the education participants, I control for it with a Heckman 2-step model. After adjusting for selection bias, I find participation in high school literacy education has no effect on future levels of income and is driven by confidence. Lower levels of confidence are correlated with higher participation in financial literacy education and lower levels of confidence are associated with lower incomes.

Introduction

Financial literacy is something that has overarching effects for society. If individuals had better knowledge of their finances, countless poor financial behaviors could be avoided. Research has shown that: foreclosures, bankruptcies, and high credit costs could all be drastically lower if people had a better understanding of basic personal finance topics. (Bernheim et al. 2010). Financial literacy can be described as the ability an individual has to understand and apply personal financial concepts and has been measured using Lusardi's "Big 5" measures of financial literacy (Lusardi, 2008). Research has also shown that financial literacy levels increase with income levels (Monticone, 2010). What has yet to be shown is how financial literacy education influences future income.

Existing literature has studied the returns to financial literacy education with respect to race (Al-Bahrani et al. 2018), access and use of student debt (Stoddard & Urban, 2019), use of high cost loans (Harvey, 2019) and how preexisting levels of income affect the outcomes of financial literacy education (Hamilton & Darity, 2017). These findings are important building blocks for the study of how financial literacy education affects a participant's future income level. The literature has yet to attempt to connect participation in financial literacy education on future levels of income. Connecting participation in financial education to future income levels is difficult because we cannot identify income levels before participation, and we cannot identify if low income individuals select into education at higher or lower rates. I use high school participation in financial education to estimate future earnings. I assume that the decision to participate at the high school level is not determined by income because future earning is not determined yet.

Using Ordinary Least Squares analysis (OLS) I find evidence that students actually do select into high school financial literacy education. Using a Heckman Two-Stage Least Squares to control for selection into financial literacy education, I find that there is no evidence that future incomes increase for participants. I find that those with lower confidence in their financial understanding are more likely to select into financial education.

Literature review

The effect of preexisting income on financial literacy education outcomes has several important articles in the literature. Income has been shown to have a positive effect on financial literacy (Monticone, 2010; Buckland, 2010). In situations where individuals had lower income their financial literacy was lower than those of their middle-class and upper-class counterparts (Buckland, 2010). This leads us towards the theory that there may be differences in demographic subgroups in their relation to benefits gained from financial literacy education. More

specifically, that lower income brackets would benefit more from more targeted financial literacy education than those of higher income brackets (Lyons et al. 2007). This evidence suggests that there may be a correlation between financial literacy education and the participants' income levels.

These papers show that trends are emerging in the relation between financial literacy education and income. They establish the basis for the claim that lower income earners have lower levels of financial literacy, as well as that lower income earners could benefit from financial literacy education more than higher income earners.

Additional support for the theory of financial literacy education's effects on income exists. The literature shows that schooling and financial literacy levels are positively correlated with important income related factors such as retirement planning and pension contributions. Financial literacy education has a stronger effect on the attainment of higher income levels than that of normal schooling (Behrman et al. 2010). In addition, financial literacy education has a positive correlation to other financial behaviors such as saving and participation in retirement savings instruments (Lusardi, 2008).

The contrary argument is that an individual's socioeconomic status is a determinant for outcomes of financial literacy education. The case is made that regardless of the amount of education they receive, their gains from it are limited based on their status, as they do not have the assets to fix their situation post treatment (Hamilton and Darity 2017).

What has yet to be shown in this literature is the direct effect financial literacy education has on income levels. This paper looks to add to the literature by showing correlation between participation in financial literacy education and future income level.

In Figure 1 I present the proposed relationship between high school financial literacy education and future outcomes. I assume that participation in financial education leads to an increase in financial literacy which leads to better overall decision making. Choice of major, understanding of markets, and career choices will all lead to better jobs that are possibly associated with higher incomes.

Conceptual Model

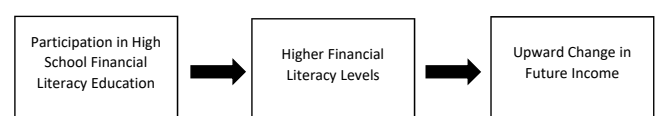


Figure 1: Description of how high school financial literacy education impacts future income

Data

I used the National Financial Capabilities Study (NFCS) from the years 2009, 2012, and 2015. The NFCS is administered every three years and samples roughly 500 individuals from every state per survey. California, Texas, New York, and Illinois are considered large states and are sampled at 1,000 individuals. I provide summary statistics for key variables in Table 1.

The income variable in this data set is broken up into buckets of the respondent's income. For example, if a respondent indicated their income was \$38,000 a year, they would fall into the 35-50k income bracket. I convert income to a continuous variable using the midpoint method, this will allow me to measure the dollar returns to high school financial literacy education.

Fully employed individuals make up 38% of the data, about 5% are currently students, and about 8% are self-employed. Using the income buckets, the sample is evenly distributed across the eight categories, with the highest number of respondents in the 50-75K range at just under 20%. The NFCS asks individuals about their participation in financial literacy education. Approximately, 14% of the respondents were offered and participated in some form of financial literacy education while in high school. The NFCS uses the "Big Five" financial literacy test to determine a respondent's financial knowledge. The average score on this exam was a 60% or 3 out of 5 correct answers, which is in line with the findings on average (Lusardi 2008, Al-Bahrani et al. 2018).

Methodology

I use an OLS regression to estimate the returns to financial literacy education. The regression I estimate is presented in equation 1.

$$I_i = \beta_0 + \beta_1 X_i + \beta_2 F + \beta_3 L + \beta_4 \text{STATEMANDATE} + \varepsilon \quad (1)$$

The dependent variable I is individual i 's income using the midpoint method. The returns to financial education participation is measured by β_2 . The variable F is a dummy variable indicating whether individual " i " participated in financial literacy education at the high school level. The variable L controls for current financial knowledge using the score on the Big Five questions. To control for state level variation in financial behaviors, I use variations in state mandates to teach financial literacy education at the high school level. This measure will capture any variations in state level efforts to influence financial behaviors, which might be correlated with other outcomes. This variable does not reflect if the individual was actually mandated to take a financial literacy course, because we do not observe when the mandate was passed and in which state the individual attended high school. I use the Stoddard and Urban (2019) database to indicate if a state

Table 1. Summary Statistics

Variable	Observations	Mean
Offered & Participated	81,219	0.139 [0.346]
Average Income	76,255	53,494 [35.833]
Financial Literacy Score	81,219	60.476 [28.959]
<15k	81,219	0.125 [0.333]
15-25K	81,219	0.116 [0.32]
25-35K	81,219	0.115 [0.319]
35-50k	81,219	0.151 [0.359]
50-75k	81,219	0.196 [0.397]
75-100k	81,219	0.125 [0.333]
100-150k	81,219	0.111 [0.315]
150k+	81,219	0.061 [0.24]
Student	81,219	0.047 [0.212]
Full time	81,219	0.379 [0.485]
Part time	81,219	0.096 [0.295]
Self employed	81,219	0.079 [0.269]

Brackets contain standard errors

Table 2. Confidence Classifications

	Confidence (n)	Percent of population
High	64,163	79
Low	17,056	21

currently mandates financial literacy education. The error term is represented by ε .

The OLS regression allows me to establish any correlation between participation in financial literacy education and future levels of income. I assume that a participant in high school financial literacy education is currently earning 0 and, therefore, financial literacy education is not driven by their current income. I do recognize that this is assumption of independence is strong, especially if individual behavior is a function of expected income. The original income variable in the NFCS was not continuous because the survey used income buckets. I transformed income into a continuous variable using the midpoint method. The results simplify the interpretation of the coefficient to a dollar measure of the return to financial education. Based on the conceptual model above, my hypothesis is that the coefficient on financial education is positive and significant. This would be evidence that participation in high school financial literacy education impact future income.

It is possible that there is selection into financial education participation. This would violate the assumptions of random selection in OLS and bias the results. Selection into financial education can be due to lack of confidence in financial understanding. If low confidence individuals are participating in financial education at higher rates, then that would bias our results because the coefficient of interest would be measuring confidence as measured through selecting to participate in financial education. I test for this possibility by using a Heckman selection model. This approach isolates the decision to participate in financial education from estimation of the returns to financial education. The two-step approach estimates equation 2.

$$P(FHS=1) = \beta_0 + \beta_1 \text{CONFIDENCE} + \beta_2 \text{STATEMANDATE} + \varepsilon \quad (2)$$

$$I = \beta_0 + \beta_1 X_1 + \beta_2 F + \beta_3 \text{ACTLITP} + \varepsilon \quad (3)$$

The Heckman works by estimating parameters (with a probit model) and then running an OLS regression with residuals of those parameters. This allows me to take into account a participant's confidence in their own financial ability. By estimating the probability of someone with low confidence in their financial ability taking a course, I can tease out the selection bias associated with confidence in a respondent's financial ability. I am able to control for financial confidence

by using the NFCS survey. The NFCS rates an individual's confidence on a scale from 1-7 with 7 being the highest. I define individuals who answer the NFCS confidence in financial ability question with a 3 or below as Low. Table 2 reports the number of people labeled as confident.

Results

Table 3 holds the initial OLS regression in which I use to find correlation between participation in high school financial literacy education and higher levels of income, the continuous income variable is the dependent variable. I find that a single person makes \$18,644.04 less annually than a married person, which is in line with research indicating that married individuals out-earn singles (Town & Antonovics, 2004). In addition, a respondent's income increases by \$269.23 for each additional percentage point increase in financial literacy score.

Table 3: Original OLS Regression

	Confidence (n)	Percent of population
Variable	Coefficient	P>T
Single	-18644.04	0
	[773.24]	
SelfEmployed	15482.89	0
	[1370.42]	
Fulltime	26198.35	0
	[919.87]	
Parttime	4456.14	0
	[1270.09]	
Unemployed	-6764.29	0
	[1536.39]	
Retired	11567.48	0
	[1146.17]	
Child1	2853.19	0
	[926.21]	
Child2	7265.31	0
	[1013.92]	
Child3	5408.15	0
	[1487.11]	
Child4plus	9216.71	0
	[1888.04]	
Offered & Participated	-3017.56	0
	[642.93]	
Financial Literacy Score	269.23	0
	[12.45]	
State Mandates	705.56	0.29
	[669.4]	

Adj R- Squared: 0.268, Prob > F: 0.000, Standard Error in brackets

Table 4: Heckman Controlled Regression

Variable	Coefficient	P>z
Single	-14996.69	0
	[707.54]	
Selfemp	12162.38	0
	[1163.69]	
Fulltime	22898.84	0
	[842.99]	
Parttime	4015.23	0
	[1178.35]	
Unemp	-5915.06	0
	[1512.53]	
Retired	9643.16	0
	[984.64]	
Child1	3680.13	0
	[802.94]	
Child2	6199.34	0
	[841.35]	
Child3	5657.39	0
	[1247.15]	
Child4plus	6864.88	0
	[1523.16]	
Offered & Participated	-1214.64	0.09
	[711.73]	
Actlitp	237.95	0
	[11.42]	
Controls for Heckman		
Offered&Part	0.09	0
	[0.03]	
Finbad	-0.16	0
	[0.04]	
Statemandate	-0.02	0.32
	[0.02]	
Robustness Checks		
Rho	0.98	
Sigma	37118.54	
Lambda	36335.29	

Chi2= 793.73

However, I find evidence of a selection bias with the high school financial literacy education participation variable. I observe a negative coefficient of \$3017.56, implying that someone who was offered and participated in financial literacy education in high school makes less than someone who didn't. This contradicts my conceptual model and would suggest that participating in financial literacy education is associated with lower incomes.

I report the results of the Heckman selection model in Table 4. This model controls for selection bias on the basis of confidence by estimating the likelihood that someone with a positive opinion of their financial ability would take a financial literacy education course. After controlling for selection, I find that participation in financial literacy in high school is insignificant and thus is not a predictor of financial outcomes. The result of the OLS were driven by confidence and indicate that less confident individuals participate in education but that results in lower labor market outcomes. Less confident individuals make less money than more confident individuals.

Interpretation of Results

I find no evidence that financial literacy education effects future incomes using a Heckman 2-Step model. My research cannot be used to assume that financial education has no impact on other financial behaviors such as saving or investing. That research is beyond the scope of this paper. These results show that a respondent who is less confident in their financial literacy knowledge is more likely to take a financial literacy course in high school than someone who is confident in their abilities. Thus, the initial OLS was measuring the impact of confidence via selection into financial literacy education. Since the OLS measure is negative, it indicates that further research should test for variations in financial confidence throughout life, and especially how it varies with participation in financial education.

Conclusions

The literature on financial literacy education supports the theory that there are positive effects of financial behaviors; from increased knowledge about credit reports, to better understanding of mortgages (Bernheim et al. 2010). This paper has shown that participation in high school financial literacy education has no effect on future income levels when controlling for selection bias based on confidence in financial ability. It also shows that an individual's confidence in their ability is an important factor in determining financial literacy knowledge and likelihood to participate in education. The implications of these results could be used to provide support for expanding research into if other levels of financial literacy education have any effect on income levels.

Acknowledgements

The author would like to thank Dr. Abdullah Al-Bahrani and the NKU Center for Economic Education for providing the support necessary to make this research possible.

References

- Al-Bahrani, A., Weathers, J., & Patel, D. (2018). "Racial Differences in the Returns to Financial Literacy Education." *Journal of Consumer Affairs*, 53 (2): 572-599.
- Antonovics, K., and Town R. (2004) "Are All the Good Men Married? Uncovering the Sources of the Marital Wage Premium." *American Economic Review*, 94 (2): 317-321.
- Bernheim, B. D., & Garrett, D. M. (2001). "The Effects of Financial Education in the Workplace: Evidence from a Survey of Households." *Journal of Public Economics*, 87: 1487-1519.
- Behrman, J. R., Mitchell, O.S., Soo, C., & Bravo, D. (2010) "Financial Literacy, Schooling, and Wealth Accumulation." National Bureau of Economic Research, Working Paper No. 16452.
- Buckland, J. (2010). "Are Low-Income Canadians Financial Literate? Placing Financial Literacy in the Context of Personal and Structural Constraints." *Adult Education Quarterly*, 60(4): 357-376.
- Hamilton, D. & Darity, W. A. (2017) "The Political Economy of Education, Financial Literacy, and the Racial Wealth Gap." *Federal Reserve Bank of St. Louis Review*, First Quarter 99(1): 59-76.
- Lusardi, A. (2008). "Household saving behavior: The Role of Financial Literacy, Information, and Financial Education Programs." National Bureau of Economic Research, Working Paper No. 13824.
- Lyons, A. C., Rachlis, M., & Scherpf, E. (2007) "What's in a Score? Differences in Consumers' Credit Knowledge Using OLF and Quantile Regressions." *Journal of Consumer Affairs*, 41(2): 223-249.
- Monticone, C. (2010). "How Much Does Wealth Matter in the Acquisition of Financial Literacy?" *Journal of Consumer Affairs*, 44(2): 403-422.

APPENDIX A.

Variable Names

SINGLE: respondent was not married

SELFEMP: respondent identified as self employed

FULLTIME: respondent identified as a full-time employee

UNEMP: respondent identified as unemployed

RETIRED: respondent identified as retired

CHILD1: respondent has 1 child

CHILD2: respondent has 2 children

CHILD3: respondent has 3 children

CHILD4PLUS: respondent has 4 or more children

F: respondent was offered and participated in financial literacy education in high school

ACTLITP: respondent's financial literacy score in percentage form

FINBAD: indicated the respondent's confidence in their financial ability as either positive or negative (dummy variable)

STATEMANDATE: participant was in a state that mandated financial literacy

The Use of Qualified Medical Interpreters in Health Care: Barriers for Health Care Professionals

Ana Alza-Rodriguez.

Faculty Mentor: Judi Wheatley Godsey,

PhD, MSN, RN

Nursing

KEYWORDS:

medical interpreters, health care

Ana Alza Rodriguez

Ana Alza Rodriguez earned a dual Bachelor in Business Management and Spanish from Northern Kentucky University and a Master of Healthcare Administration from the University of Cincinnati. She is currently a third year nursing student at NKU and has been working with her mentor, Dr. Judi Wheatley Godsey, to describe barriers among healthcare professionals in the use of qualified medical interpreters.

Ana has worked as a Certified Medical Interpreter and Translator at University Hospital Medical Center for over two decades. She plans to continue working on the creation of evidence based effective communication strategies for the Limited English Proficiency population.

Abstract

Access to language services has become a fundamental component of care for patients who are not proficient in the English language. Patients who cannot effectively communicate in their preferred language require qualified bilingual support to ensure they receive equal access to health care. Despite the understood need for quality bilingual services, many health care professionals may be reluctant to use or provide qualified medical interpreters. A review of the literature is provided regarding barriers, which prevent or negatively influence health care professionals' decision to use a qualified medical interpreter, thus disregarding policies regarding language laws. The purpose of this literature review is to increase understanding of the benefits of interpreter utilization, which acknowledges the rights of Limited English Proficient (LEP) patients, and maximizes the use of collaboration among health care professionals.

Introduction

Approximately 22% of the US population age 5 and older speak another language at home other than English (U.S. Census Bureau, 2015). It can be assumed many Limited English Proficient (LEP) persons have sought health care services in the US. Confronted with LEPs in the acute care setting, many health care providers often rely on family members and other non-qualified sources to convey messages between the healthcare staff and the patient. This mode of communication increases the risk of performing a wrong procedure, can result in medication errors, and increase readmission rates (Jacobs, Shepard, Suaya, & Stone, 2004).

The Joint Commission requires all hospital systems implement a language access plan and include health care professionals' participation in educational trainings to promote understanding regarding the benefits of using qualified medical interpreters (JCAHO, n.d.). Any individual receiving federal financial assistance from the U.S. Department of Health and Human Services (HHS) is subject to Title VI of the Civil Rights Act of 1964. Title VI prohibits discrimination based on race, color, or national origin in any program or activity receiving federal financial assistance (HHS, 2014). The U.S. Office of Civil Rights (OCR) and the Department of HHS have released a summary of compliance reviews and complaint investigation summaries describing failures to provide language access when providing healthcare services to the LEP patient population (HHS, 2014). Many hospital systems have entered into voluntary resolution agreements with the OCR to improve access to language services for LEP patients (HHS, 2014). Memorial Hospital of California is an example of a healthcare agency which has implemented an agreement to expand and improve access to health care for LEP patients to ensure compliance with all provisions of Title IV are met (HHS, 2014). The purpose of the agreement is to increase and expand language access for the LEP patient population. Conditions of the agreement include implementation of a language access plan consistent with best practices, implementation of language policy, procedures for oral interpretation and written translation for LEP patients (HHS, 2014). The agreement included creation of a position for a language assistance coordinator and a community advisory board to ensure access to qualified medical interpreters and identify and provide interpreter training (HHS, 2014).

This purpose of this literature review is to describe barriers in the use of qualified interpreters by healthcare professionals. Barriers reported in the literature include interpreter utilization, time constraints, health care professional-interpreter collaboration, language and modes of interpreting services, and cost of using interpreters. This paper also describes recommendations for best practices to improve utilization of qualified medical interpreters in health care systems.

Methods

A search was conducted to retrieve journal articles containing references related to the research question. The databases used to find information were PubMed, Google Scholar, ProQuest, Wiley and Elsevier. In addition, other resources like Agency for Healthcare Research and Quality (AHRQ), the Centers for Medicare and Medicaid (CMS), and the Mendeley Library were utilized. The Mendeley citation manager served to categorize journals and articles, and to simplify retrieval of relatable data.

Library Databases

Articles were selected from the University of Cincinnati Langsam Library and Northern Kentucky University Stealy Library, in addition to online resources. Databases used in the search for information included PubMed, Google Scholar, ProQuest, Wiley, Cochrane and Elsevier. Most of the relatable data needed to answer the research question were available in Google Scholar and Elsevier.

Journals

Mendeley citation manager was the main source used to categorize journal articles and simplify retrieval of relatable data. Among the twenty-five articles initially retrieved, six served as the primary sources of data and information from the following journals: The Journal of General Internal Medicine, American Journal of Public Health, International Journal of Evidence-Based Health Care, and Annals of Emergency Medicine.

Non-Refereed Sources

Non-refereed sources included health communications, Patient Education and Counseling, the Department of Human Health Service, the Office of Civil Rights (OCR) and Small Business Costhelper. Additionally, other resources like AHRQ, CMS and the Department of Health and Human Services (HHS) provided information on interpreter role and rules, language laws and language access plans.

Key Words

There were several key words used to search for articles and gather information to answer the research question and develop recommendations for practice. The list of key words, phrases and concepts included some or all of the following:

Factors influencing health care professionals' decision to avoid using a qualified medical interpreter,

Limited English proficiency (LEP) quality of care with and without an interpreter,

Health care professionals and interpreter collaboration,

Consequences of interpreters as patient advocate,

Bilingual health care and quality of health care delivery,

Cost of qualified medical interpreter usage and cost of not using qualified medical interpreters,

Differences in price between interpreter tools.

Literature Review

The Current Environment

Understanding how healthcare professionals view medical interpreters, along with expectations of how interpreters should behave during the patient-health care professional encounter, could provide essential clues for understanding the disregard of institutional policies regarding the use of qualified medical interpreters (Hsieh, Pitaloka and Johnson, 2013). The interpreter's potential interference with information shared may be viewed as disruptive to the patient-health care professional communication process and may offer an explanation as to why health care professionals may avoid using a qualified medical interpreter, despite the evidence of increased quality of healthcare when one is used (Hsieh et al., 2013). Allowing interpreters to modify communication methods during a medical encounter may hinder a health care professional's understanding of their responsibility over the medical encounter (Hsieh et al., 2013). Relying on another person to convey critically important information in a language not understood by the healthcare provider can create a sense of mistrust leading to conflict and lack of collaboration (Hsieh et al., 2013). The lack of understanding regarding the value of qualified medical interpreter roles and working in collaboration with the bilingual health care delivery team can negatively impact patient quality and the safety of health care delivery (Hsieh et al., 2013). Mistrust may ultimately explain why health care professionals would choose to go without using a qualified medical interpreter in future medical encounter with his or her patients (Hsieh et al., 2013).

Communication with the healthcare professional varies according to the nature of the appointment. For example, while a specialized health care professional may want to establish a long-term relationship with the LEP patient for continuing care, he or she may want to gather medical history information to decide on a prognosis that focuses on the patient's long-term care goals. An emergency care physician may want to focus on gathering information that focuses on the immediate health concern. The literature suggests health care professionals should be given the responsibility and opportunity to share their opinions as to what represents best practices when using qualified medical interpreters during bilingual health care delivery (Hsieh et al., 2013). Qualified medical interpreters must have the ability to adapt to the health care professional's needs and expectations during the encounter, without interrupting or

interfering. This adaptation could ensure quality of interpreter interaction as well as improved healthcare professional attitude toward working with interpreters (Hsieh et al., 2013).

Health Care Professional's Expectations

A study on bilingual health communication compared different expectations for medical interpreter roles from different specialties (Hsieh et al., 2013). Thirty-nine providers were surveyed, with five different specialties from a major health care facility. The surveyor ranked the health care professional's preference based on the parameters of how important the items were from extremely unimportant to extremely important. The data also measured health care professionals' view of qualified medical interpreters as a patient's ally (advocate), interpreter as a health care professional (abstain from interference) and interpreter as a health care professional proxy (ability to ensure quality of care, viewed as member of healthcare team). The results of the survey revealed support for the use of qualified medical interpreters. Interpreters were viewed as an important resource in facilitating communication with LEP patients, rather than relying solely on a health care team member (Hsieh et al., 2013). Data regarding interpreter as patient ally varies according to health care professional specialty. Nurses may prefer interpreters to be more involved in advocacy due to the potential cultural barriers, physicians may disagree. Those who oppose the use of interpreters argue the clinical significance of the patient's condition may increase the possibility of interference with effective delivery of the message intended (Hsieh et al., 2013).

Interpreter Utilization

Use of qualified interpreters remains low, despite the increased quality of care and association with patient-centered care that qualified medical interpreters support. A sample of California oncologists (n=301) reported being more likely to "sometimes" elicit assistance from family members or friends (91%) over qualified medical interpreters (40%) when providing health care to LEPs with breast cancer (Karliner, Jacobs, Chen, & Mutha, 2007). A study of 4,224 patients outlined hospital outcomes and readmission rates according to patient characteristics, including language needs (López, Rodríguez, Huerta, Soukup, & Hicks, 2015). Of the 765 patients readmitted, 96 (13%) were LEP patients who qualified for bilingual health care services. Only 32 patients (33%) of those who qualified for bilingual support actually received the services of a qualified medical interpreter present (López, et al., 2015).

Another reason healthcare professionals may avoid using qualified medical interpreters is the belief their own language skills are sufficient for effectively conveying important information, including informed consent for surgical procedures to LEP patients (Diamond, Tuot, & Karliner, 2012). One study reported that 26 out of 68 physicians used their low level of Spanish proficiency to deliver bilingual health care

services (Diamond et al., 2012). Currently, no set policies and procedures are in place in health care systems to evaluate health care professionals' language skills in health care (Diamond et al., 2012). Therefore, health care professionals with limited language proficiency may still use insufficient language skills, even when sensitive and important information is being shared with LEP patients.

The use of a qualified medical interpreter may reduce errors while delivering health care services. A study conducted in two Massachusetts pediatrics' emergency departments reported that out of 50 encounters with LEP patients requiring bilingual health care services, 1,884 errors were observed. Only 12% of those errors resulted with using qualified medical interpreters, while 22% were the result of using a non-qualified interpreter and 20% with no interpreter used (Flores, Abreu, Barone, Bachur, & Lin, 2012). It is important to note that the level of training qualified medical interpreters possess significantly improves the quality and safety of bilingual health care delivery. The findings suggest that qualified medical interpreters receive at least 100 hours of training (Flores et al., 2012). The literature further suggests more research measuring the acceptance of medical interpreters as part of the health care team is crucial in order to address conflicts resulting from tensions brought on by working with medical interpreters who tend to play the role of patient ally, also identified as advocate and/or conduit (Hsieh et al., 2013).

Another factor influencing a health care professional's decision to use a qualified medical interpreter is lack of time management and tight time constraints. According to a study conducted by the Department of Communication, at the University of Oklahoma, lack of time is one of the leading reasons affecting a health care professionals' decision to avoid using a qualified medical interpreter (Hsieh, 2015). Their choice is usually affected by health care professionals' schedule; disruptions and organizational issues are not usually addressed prior to a patient encounter. This is the result of lack of firm health care system policies and/or a language access plan with protocols regarding a patient's accurately identified language needs (Hsieh, 2015). It is very common for healthcare professionals to face making the decision to prioritize their focus as to what medical condition requires immediate attention. In many health care settings within hospital systems, health care professionals are often required to provide immediate attention to a patient who requires urgent attention. This forces health care professionals to abandon the bilingual health care service patient encounter and the scheduled interpreter is left waiting for the health care professional to return to the session. The problem arises when the scheduled interpreter assigned to be at the patient encounter appointment for a calculated unit of paid hours has to leave, thus leaving the patient without an interpreter (Hsieh, 2015).

Health care professionals may find it challenging working in collaboration with medical interpreters. This is a direct result of undefined understanding of the interpreter's role and boundaries while conveying the intended message to the patient. Differences between health care professionals and interpreter experiences presents as a major barrier to facilitate collaboration between the two parties in order to effectively communicate the intended message to the patient. This conflict is avoid by appropriate implementation of trainings defining interpreter role and how to work in collaboration with medical interpreters by having an understanding of interpreter practices (Hsieh, 2010). Health care professionals also rely on interpreter emotional support during the interpreting session. This expectation is satisfied when the interpreter is able to remain neutral and accurately conveys the intended message (Hsieh & Hong, 2010).

Qualified medical interpreters are essential resources for communicating with the LEP patient population. Their role is not to participate in decision making during bilingual health care delivery, despite their knowledge regarding the patient's culture. This literature review suggests there is a culture of overlapping roles between qualified medical interpreters and health care professionals where the interpreter acts as a patient advocate interfacing with the integrity of the intended message the health care professional is trying to deliver. Medical interpreters are not viewed as part of the health care delivery team. Their role is to remain neutral and refrain from acting as a patient's advocate (Hsieh & Kramer, 2012). The manner in which health care professionals view interpreters can significantly affect the LEP patient's understanding of the health care plan and impact the quality of care they receive (Hsieh & Kramer, 2012).

Financial Burden

Risk and volume of LEP patients seeking care from health care systems may be two factors health care systems administrators consider when choosing not to use qualified medical interpreters and/or when selecting the type of interpreting tool to use in order to remain efficient while maintaining compliance. Data reported by the illustrating the financial burden interpreting services have on hospital systems is available in the HHS.gov website. This data explains how financial burden can affect the decision for healthcare professionals to do without interpreters, taking into account that interpreting services are a non-revenue generating service that must be offered to LEP patients by health care systems receiving government subsidies (HHS, 2014). An in-person interpreter has a price tag of \$145-\$450 per hour, phone interpreters are \$75-\$180 per hour, and video interpreters are \$105-\$420 per hour (Costhelper, Inc, n.d.). According to the 2016 Bureau of Labor Statistics report, a full time interpreter can earn an average salary of \$46,120 per year, a maximum of \$83,010 per year and a minimum of \$25,370 per year (Medicaid Translation and Interpretation Services, n.d.).

Discussion

Literature regarding the rate of qualified medical interpreter use in health care is scarce. This gap in information contributes to the lack of effective solutions to address factors that may influence a health care professional's decision to avoid using a qualified medical interpreter (Meuter, Gallois, Segalowitz, Ryder, & Hocking, 2015). The current literature and research available is insufficient, does not apply to all clinical settings, and is unclear in most instances, which creates confusion about the importance of using qualified medical interpreters in health care settings.

Results of this literature review suggest there are solutions available to address language access to LEP patients; however, there is lack of evidence to address a health care professional's decision to avoid using a qualified medical interpreter. As a result, LEP patients continue to suffer adverse events in the healthcare system along with being exposed to the increased costs and liability imposed to hospital systems. Health care professionals' disregard of the importance of qualified medical interpreters and/or appropriate language tools to assess and treat LEP patients are primarily due to lack of collaboration with the qualified bilingual team and time constraints. The five elements of best practice identified solutions to increase qualified medical interpreter utilization. This model included effectiveness, reach, feasibility, sustainability, and transferability.

Interpreter Usage Protocol

Effective solutions are needed which significantly increase the use of qualified medical interpreter utilization and support improved health outcomes for the LEP patient population (Meuter, et al., 2015). Best practices suggested in the literature include addressing factors affecting a health care professional's decision to avoid using a qualified medical interpreter. The second best practice suggested by the literature is the improvement of collaboration between the health care professional and interpreter. This is to address the conflict arising from the lack of trust regarding the risk involving the integrity of the information conveyed by the interpreter. Can the interpreter effectively communicate the information the health care professional intended to deliver without changing the intended meaning? At this time, there is a lack of available measurements to evaluate the quality of the average interpreting session (Hsieh, 2010). Feasibility and considerations to increase qualified medical interpreter utilization provide barriers to implementation of solutions. Health care professionals are resistant to adopt policies that require the use of a qualified medical interpreter. In addition, the patient may decline an interpreter in many occasions do to longer wait times, privacy concerns and convenience.

Increasing Utilization of Interpreters

There are many hospitals across the U.S. who have implemented a policy requiring the use of a qualified interpreter with the LEP patient population, but they continue to see health care professionals resistant to adopt the idea of using a qualified medical interpreter (Torres, n.d.). An example of best practice recommendations for health care professional training is creating policies and procedures around language assistance. This is established by making health care professionals aware of these policies and procedures, as well as providing training on how to work in collaboration with qualified medical interpreters. Trainings are incorporated during employee orientation, staff meetings, in-services, risk management sessions, medical school seminars, grand rounds and continuing education programs (Torres, n.d.).

A barrier to solutions to increase qualified medical interpreter utilization is the lack of data available supporting effectiveness of proposed solutions. Capturing data regarding the effectiveness of these solutions, to increase qualified medical interpreter utilization, will provide case precedence to make available evidence that supports the importance of using a qualified medical interpreter.

Collaboration between Health Care Professional and Interpreter

Educating health care professionals regarding the value of working in collaboration with qualified medical interpreters is a challenge, especially if there is a diverse health care professional team varied in language themselves. Proving that effective communication is the key to healing in a way where LEP patients clearly understand is a major challenge. Centralized solutions to increase collaboration among the bilingual team and the use of qualified medical interpreters may ensure effective communication with the LEP patient population (Hsieh, 2010).

Full collaboration among the bilingual health care team offers the potential to improve their relationship. By collaborating, the bilingual health care team will improve communication between themselves, the health care professionals and the qualified medical interpreter, and they will increase cultural awareness among themselves while preventing adverse effects due to lack of communication. The benefit of establishing a relationship among the bilingual will result in improvement of patient satisfaction, it will prevent liability issues due to misinterpretation of information that can lead to misdiagnosis, and over ordering diagnostic tests that can result in an increase of health care cost (Hsieh, 2010).

Recommendations

The literature review identified time constraints as one of the factors affecting health care professional's decision to avoid using a qualified medical interpreter. One of the suggested solutions is to determine if the patient is LEP, and in what language the patient prefers to receive his or her health care service. This process will improve time management by identifying needs for bilingual health care ahead of time, and will ensure the use of a qualified medical interpreter.

To address the financial burden that may result from not complying with using qualified medical interpreters, is important to implement a language access plan. This language access plan must include solutions to increase the use of qualified medical interpreters. An example is The Health Collaborative, a Cincinnati nonprofit organization serving the tristate. This organization works alongside health care professionals and administrators to assist with creating health care solutions faced by many health care systems in the area. This Health Collaborative has implemented an initiative to control price of interpreting services by setting up price standards (Greater Cincinnati Regional Language Access Committee, 2018). This initiative forces competitive agencies to negotiate lower price. Despite this effort, the financial burden interpreting services may have over health care systems may impact administrators' decision to forgo using qualified medical interpreters.

To ensure that health care professionals use a qualified medical interpreter, a policy highlighting the benefits of using a qualified medical interpreter should be available twenty-four hours a day. In addition, this policy should stress that the interpreter has the professional obligation to maintain confidentiality during bilingual health care delivery, and that the health care professional has the responsibility to ensure care, quality and safety of effective communication via the qualified medical interpreter when it is required (Hadziabdic & Hjelm, 2013).

Conclusion

The literature was consistent regarding the necessity to implement solutions that can increase utilization of qualified medical interpreters in the healthcare environment. Effective implementation of these solutions will ensure compliance with language legislation, and with the standards of best practice set forth a customized language access plan. Health care systems that are not willing to implement solutions to increase utilization of qualified medical interpreters will be at risk of losing government subsidies (Hsieh, 2010). It is important to continue gathering data regarding utilization of qualified medical interpreters, and measurement for best solutions for health care systems based on LEP patient volume and service provided. Data which measures risks involved when qualified medical interpreters are not used will further provide

validation of their importance in the bilingual health care delivery encounter. Finally, protocols which report incidents of clinical errors resulting from lack of use of a qualified medical interpreter will be essential for quantifying the significance of this problem.

References

- Costhelper, Inc. (n.d.). Cost of interpreter: Small business. Retrieved from: <http://smallbusiness.costhelper.com/interpreters.html>
- Diamond, L. C., Tuot, D. S., & Karliner, L. S. (2012). The use of Spanish language skills by physicians and nurses: Policy implications for teaching and testing. *Journal of General Internal Medicine*, 27(1), 117–123. Retrieved from: <https://doi.org/10.1007/s11606-011-1779-5>
- Flores, G., Abreu, M., Barone, C. P., Bachur, R., & Lin, H. (2012). Errors of medical interpretation and their potential clinical consequences: A comparison of professional versus ad hoc versus no interpreters. *Annals of Emergency Medicine*, 60(5), 545–553. Retrieved from: <https://doi.org/10.1016/j.annemergmed.2012.01.025>
- Hadziabdic, E., & Hjelm, K. (2013). Working with interpreters: Practical advice for use of an interpreter in healthcare. *International Journal of Evidence-Based Healthcare*, 11(1), 69–76. Retrieved from: <https://doi.org/10.1111/1744-1609.12005>
- Hsieh, E. (2010). Provider-interpreter collaboration in bilingual health care: Competitions of control over interpreter-mediated interactions. *Patient Education and Counseling*, 78(2), 154–159. Retrieved from: <https://doi.org/10.1016/j.pec.2009.02.017>
- Hsieh, E. (2010). Provider-interpreter collaboration in bilingual health care: Competitions of control over interpreter-mediated interactions. *Patient Education and Counseling*, 78(2), 154–159. Retrieved from: <https://doi.org/10.1016/j.pec.2009.02.017>
- Hsieh, E. (2015). Not Just “Getting by”: Factors influencing providers’ choice of interpreters. *Journal of General Internal Medicine*, 30(1), 75–82. Retrieved from: <https://doi.org/10.1007/s11606-014-3066-8>
- Hsieh, E., & Hong, S. J. (2010). Not all are desired: providers’ views on interpreters’ emotional support for patients. *Patient Education and Counseling*, 81(2), 192–197. Retrieved from: <https://doi.org/10.1016/j.pec.2010.04.004>
- Hsieh, E., & Kramer, E. M. (2012). Medical interpreters as tools: Dangers and challenges in the utilitarian approach to interpreters’ roles and functions. *Patient Education and Counseling*, 89(1), 158–162. Retrieved from: <https://doi.org/10.1016/j.pec.2012.07.001>

- Hsieh, E., Pitaloka, D., & Johnson, A. J. (2013). Bilingual health communication: Distinctive needs of providers from five specialties. *Health Communication*, 28(6), 557–567. Retrieved from: <https://doi.org/10.1080/10410236.2012.702644>
- Jacobs, E. A., Shepard, D. S., Suaya, J. A., & Stone, E.-L. (2004). Overcoming language barriers in health care: costs and benefits of interpreter services. *American Journal of Public Health*, 94(5), 866–869. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15117713>
- Joint Commission on Accreditation of Healthcare Organizations (JCAHO, n.d). Overcoming the challenges to providing care to LEP populations. Retrieved from <https://www.jcrinc.com/assets/1/7/PS5.pdf>
- Karliner, L. S., Jacobs, E. A., Chen, A. H., & Mutha, S. (2007). Do professional interpreters improve clinical care for patients with limited English proficiency? A systematic review of the literature. *Health Services Research*, 42(2), 727–754. Retrieved from: <https://doi.org/10.1111/j.1475-6773.2006.00629.x>
- Greater Cincinnati Regional Language Access Committee (2018). Language access plan. Retrieved from <http://healthcollab.org/wp-content/uploads/2018/02/2018-Regional-Language-Access-Plan-template-for-organizations.pdf>
- López, L., Rodriguez, F., Huerta, D., Soukup, J., & Hicks, L. (2015). Use of interpreters by physicians for hospitalized limited English proficient patients and its impact on patient outcomes. *Journal of General Internal Medicine*, 30(6), 783–789. Retrieved from: <https://doi.org/10.1007/s11606-015-3213-x>
- Meuter, R. F. I., Gallois, C., Segalowitz, N. S., Ryder, A. G., & Hocking, J. (2015). Overcoming language barriers in healthcare: A protocol for investigating safe and effective communication when patients or clinicians use a second language. *BMC Health Services Research*, 15, 371. Retrieved from: <https://doi.org/10.1186/s12913-015-1024-8>
- Torres, B. (n.d.). Commonwealth of Massachusetts executive office of health and human services Massachusetts department of public health best practice recommendations for hospital-based interpreter services office of minority health. Retrieved from <http://www.mass.gov/eohhs/docs/dph/health-equity/best-practices.pdf>
- Translation and Interpretation Services | Medicaid.gov. (n.d.). Retrieved April 3, 2018, from <https://www.medicaid.gov/medicaid/finance/admin-claiming/translation/index.html>
- US Census Bureau (2015). Detailed language spoken at home and ability to speak English. Retrieved from: <https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html>
- US Department of Health and Human Services Office for Civil Rights (2014). Mee Memorial Hospital agreement. Retrieved from: <https://www.hhs.gov/civil-rights/for-providers/compliance-enforcement/mee-memorial-hospital/index.html>

A Barrier to Democracy: Corruption in Former Soviet Eastern Europe

Lydia Schubarth.

Faculty mentor: Ryan Salzman

International Studies

Lydia Schubarth

Lydia Schubarth is currently a senior at Northern Kentucky University and will graduate in May 2020. She is working towards a Bachelor of Arts degree in international studies with minors in political science, history, and European studies. After graduation, she hopes to obtain a Master's degree in security or conflict resolution. Lydia is enthusiastic about this opportunity and thanks her faculty mentor, Dr. Ryan Salzman, for his help and guidance on this project over the past year.

KEYWORDS:

democracy, corruption, Eastern Europe

Abstract

One of the difficulties that a country faces when working toward democratization is the corruption that may run rampant in their government. This paper will analyze corruption in government and how it affects democracy, particularly in four of the former Soviet Union countries that are located in Eastern Europe: Moldova, Ukraine, Belarus, and Georgia. Qualitative and quantitative analysis is provided regarding each of the four former Soviet Union countries listed. By using the Corruption Perceptions Index (CPI) scores given by Transparency International, the countries are compared and contrasted in order to examine how their corruption has changed over time and how their situations and governments have affected their level of corruption. Scores given to each of the countries by Freedom House are also used to compare democracy. CPI and Freedom House scores are used in conjunction to analyze the relationship between corruption and democracy and to locate a correlation between the two. After analysis of the democracy and levels of corruption in each country in the study and seeing the correlation, it is clear that democracy and corruption affect each other.

Introduction

When a country is striving toward democracy, one of the biggest issues is the possibility of corrupt officials holding positions of power. According to the website Transparency International, corruption is defined as “the abuse of entrusted power for private gain. It can be classified as grand, petty, and political, depending on the amounts of money lost and the sector where it occurs” (Corruptions Perceptions Index, 2017).

Due to the large impacts that corruption can have on a newly-established or even a well-maintained democracy, it is important to understand those impacts, the relationship between corruption and democracy, and how countries can move away from corrupt practices. The problems are especially evident in post-Soviet Union Eastern Europe.

After the fall of the Soviet Union in the late 1980s and early 1990s, the former Soviet countries were left on their own. As a result, corruption that became commonplace during the time of the Soviet Union not only continued, but flourished. This paper will analyze the relationship between corruption and democracy in general and specifically regarding four post-Soviet Union Eastern European countries: the Republic of Moldova, Ukraine, the Republic of Belarus, and Georgia. Qualitative and quantitative evidence will illuminate the relationship between corruption and democracy in the case studies those countries.

Literature Review

There is no government that has not experienced some degree of corruption, but the scope of corruption varies between different countries.

Charron and Bagenholm (2016) claim that corruption not only results in mistrust of politicians and disbelief in the claims or promises that they put forth, but also other undesirable consequences. They state that some of these undesirable consequences are less economic development within the country, more inequality—both based on demographics and income inequality—poor health outcomes and environmental conditions, and less trust in the government, which in turn affects the morale of the people and creates a less happy population. They go on to explain that when citizens feel as if their government is not listening to them, they are less likely to be happy with the governmental structure and the system as a whole.

Democracy can be corroded away—a basis of democracy is the concept of politicians representing citizens, and when they are not doing that, it takes away the essence of democracy. According to Diamond (2008), there are few things that will

erode away trust in government as quickly and effectively as corruption will, and citizens should expect their representatives in government to abide by the same laws and rules that the citizens must.

There are many essential parts of a democracy, some of which concern corruption. According to Charron and Bagenholm (2016), one of these most central parts is electoral accountability, which operates by punishing politicians that are clearly corrupt and behaving inappropriately in their position of power, and by rewarding those politicians that don't misuse the power that they have obtained. Charron and Bagenholm explain that by voting out the politicians that are not in line with the interests of the citizens or that demonstrate corrupt motivations, people are able to clearly show their government officials that they do not support their behavior and refuse to accept failure, mismanagement, and criminal activities engaged in by their representatives.

Thoughts by Diamond (2008) demonstrate the expectation that the citizens in a democracy place on their government officials and representatives to follow the laws that all of the citizens have to follow and to represent the citizens' interests in government to the best of their ability. When it is proven that officials are corrupt and are not doing that, it prevents citizens from trusting in those people—or perhaps the system of government—in the future.

According to Drury, Kriekhaus, and Lusztig (2006), democratic institutions are important in order to prevent corrupt authoritarian leaders from taking over. But if there are already corrupt authoritarian leaders in the government, it is nearly impossible to establish and maintain these democratic institutions in order for them to do their job; they need to be implemented before those corrupt officials have gained power. They argue that democracy doesn't only decrease the level of corruption that is within a government, but democracy also affects the nature of the corruption. They support their argument by saying that because citizens in a democracy hold power, they are able to vote politicians out, and therefore the level of corruption decreases. Politicians may not act corruptly not because they are concerned for the well-being of their country and government, but because they are worried about holding their jobs and keeping the power—however minimal it may be—that they possess. However, the authors make the point that in cases of corruption that may have very minimal consequences, such as nepotism or small bribes, corruption in the democracy may be unabated due to the benefits outweighing the minor political costs that may come as a result.

The position taken by Drury, Kriekhaus, and Lusztig (2006)

seem to be supported by Rose-Ackerman (1996). According to Rose-Ackerman, democracy is the least corrupt form of government due to the fact that the need for reelection outweighs the short-term benefits that a politician could earn from corruption. She states that “the protection of civil liberties and free speech, which generally accompany democratic electoral processes, make open and transparent government possible” (Ackerman, 1996, p. 83). Because of this, corruption thrives more in an authoritarian or totalitarian regime than in democracy. The author also makes it clear that corruption in any form or case by a politician in a democratic state reduces the legitimacy of the democracy and the state.

Warren (2004) claims that there are many different pathologies of politics, and the one that thrives the most in democracies is political corruption. This is different from the other literature surrounding this topic, as other literature points to political corruption not being able to thrive due to the institutions set up in a democracy. Warren even makes the statement that corruption in a democracy may have the potential to be beneficial, by “lowering transaction costs, reducing the inefficiencies of cumbersome rules, and generally making things happen” (Warren, 2004, p. 328). Despite this, he does accept that corruption undermines the political culture that surrounds democracy, and can cause citizens to become cynical toward their government and its officials, whether or not each official has been proven to be corrupt or not.

According to Azfar, Lee, and Swamy (2001), citizens respond to the services that they are provided by public officials as what is referred to be a tip, or even a gift. In Western countries this may seem corrupt, but in their cultures it is normal and customary. The authors point out the necessity of defining corruption, due to the fact that the lines can be extremely blurry in what is or is not considered corruption. Among these blurry qualifications of corruption are at what point a tip or gift becomes a bribe and how campaign donations and financing can potentially be considered corruption, citing that in America, campaign donations serve the same roles and purposes that corruption may serve in other countries.

Similar to Warren (2004), Azfar, Lee, and Swamy (2001) are able to recognize the potential benefits of low-levels of corruption in some states. They state that corruption has the possibility to be socially advantageous. They give the example of a government employee that doesn’t make much money and due to that, is unmotivated to effectively perform the work that they have to do. A small bribe—or tip, or gift—could potentially expedite the process and puts a bit of money in that employee’s pocket, and in that case makes it a beneficial situation. While that is a specific instance that by no means applies to the majority of corruption, it is important to be aware of the variety of purposes that corruption may play in the state. Even in this

case, it does undermine democracy and unfairly favor those who have the money to pay a bribe.

In total, the literature seems to emphasize the point that corruption ideally has no place in a democratic system. While there are small benefits that can occur on an individual basis from corruption, overall it tends to hurt the system. Corruption favors the rich and the poor are treated unfairly, therefore widening the economic income inequality gap and undermining the entire concept of democracy where everyone is equal and citizens’ best interests are served in government. Thus, our expectation is that we will find a negative relationship between corruption and democracy in our analysis of post-Soviet Eastern European states.

Methods

A helpful tool in analyzing the level of corruption that is present in countries across the world is the Corruption Perceptions Index (CPI). The CPI is maintained and was created by Transparency International. The methodology is relatively simple; all countries are scored and ranked according to the same scale, focusing on corruption such as nepotism, bribery, and public office usage in order to gain privately. Country experts and business people score each country, based on set criteria, every four years (Corruption Perceptions Index, 2017). The scoring process by Transparency International is as follows (2017):

“Standardise data sources to a scale of 0-100 where a 0 equals the highest level of perceived corruption and 100 equals the lowest level of perceived corruption. This standardisation is done by subtracting the mean of each source in the baseline year from each country score and then dividing by the standard deviation of that source in the baseline year. This subtraction and division using the baseline year parameters ensures that the CPI scores are comparable year on year since 2012. After this procedure, the standardised scores are transformed to the CPI scale by multiplying with the value of the CPI standard deviation in 2012 (20) and adding the mean of CPI in 2012 (45), so that the data set fits the CPI’s 0-100 scale”.

After three sources assess each country, the average is calculated and a score is given, with lower values denoting more corruption while higher scores imply less corruption. According to the score that is given, the countries are ranked relative to one another. There are 180 countries that are ranked, therefore each country is given a ranking between 1 and 180.

Corruption and Democracy in Eastern Europe

Corruption is a problem that post-communist Eastern Europe has been dealing with for decades. Ever since the start of the fall of the Soviet Union in the late twentieth century, many of

Table 1. Perceived Corruption in Eastern Europe

Country	2018	2017	2016	2015	2014	2013
Moldova	33	31	30	33	35	35
Ukraine	32	30	29	27	26	25
Belarus	44	44	40	32	31	29
Georgia	58	56	57	52	52	49

Source: Corruption Perceptions Index

Table 2 Freedom House Scores of Eastern Europe

Country	2019	2018	2017	2016	2015	2014
Moldova	58	61	62	60	63	64
Ukraine	60	62	61	61	62	55
Belarus	19	21	20	17	14	14
Georgia	63	64	64	64	64	63

Source: Freedom House

those countries have had to work on their own democratic governmental systems. In the cases of Moldova, Ukraine, Belarus, and Georgia, none of those countries are considered to be full democracies. Working toward a democratic system has proven difficult, in part because of the corruption that is present. Table 1 shows perceptions of corruption that were sourced from the Corruption Perception Index.

According to the Corruption Perceptions Index, Eastern Europe is among the worst regions in the world for corruption with an average CPI score of 34 (Corruption Perceptions Index, 2018). Diamond (2008) also brings up the point that in most post-Soviet countries, democracy is not the norm. Older generations were born during communism and authoritarianism; the Soviet Union was the only government they knew for a long time. Because of that, it's difficult to establish strong democracy, especially because of the political culture of the people. He explains that due to the corruption of the Soviet Union, the citizens in post-Soviet countries have a general mistrust for the government. Democracy is difficult to maintain, and impossible to not have corruption undermine it in some way. Some degree of corruption can be seen in all post-communist Eastern European countries, particularly Moldova, Ukraine, Belarus, and Georgia.

An indicator of the level of democracy of a country is the score that it is given from Freedom House. Higher values indicate a higher level of democracy, while lower values indicate a lower level of democracy. Table 2 shows the values given by Freedom House measuring democracy in Moldova, Ukraine, Belarus, and Georgia from the years of 2014 to 2019.

The emphasis of this paper is on corruption in government and its relationship with democracy, so it is useful to compare the aggregate scores provided by Freedom House over time with each country and in relation to their CPI scores. The data from the CPI consists of the years 2013-2018, while the Freedom House scores are from 2014-2019. This is due to the fact that the CPI has not yet released their 2019 scores, and Freedom House does not have aggregate scores for countries before 2014. The years were chosen in order for them to overlap as much as possible and while still providing six years of data.

Moldova

The Republic of Moldova is a post-communist country that was at one time a part of the Soviet Union. Moldova declared its independence in 1991 after the dissolution of the Soviet Union had begun, and continued to create its own constitution

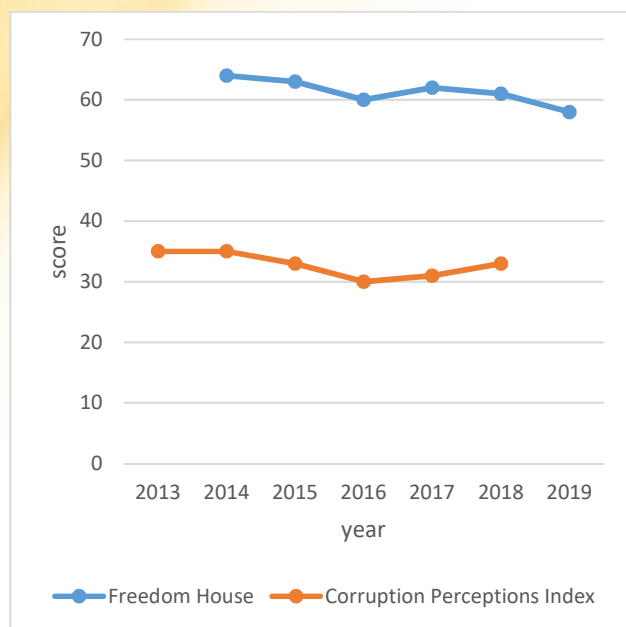


Figure 1: Freedom House and CPI's for Moldova.

in 1994. Moldova was one of the only former Soviet Union countries that had some aspects of democracy right after the fall of the Soviet Union (Diamond, 2008).

Despite being independent for a few decades, corruption remains normal in Moldova (figure 1). According to Freedom House, "Corruption remains a widespread problem at all levels of government, and existing anticorruption laws are inadequately enforced. In a report published in 2017, the National Anticorruption Center recorded a 23 percent increase in corruption cases in 2016 compared to 2015" (Freedom House, 2018).

Regarding the transparency of the government of Moldova, Freedom House also states that "the government does not operate with transparency. Most political activity takes place behind the scenes" (Freedom House, 2018). When such political activity is not permitted to be in the public eye, it is worrisome. Citizens are unable to be involved in the amount of politics in which they should be involved, and that is a problem when the country is working toward being a real democracy.

Moldova is riddled with economic crime and corruption. According to Lilia Carasciuc, the cause and baseline of this corruption is the "lack of control of state employees' activity and a low enforcement rate, as a well as delays in the payment of wages for state workers" (Carasciuc, 1999, p. 128). When state employees are not monitored and are not held accountable for their actions, that is when corruption is able to run rampant and it is evident that this has occurred in Moldova.

Moldova is ranked low on the Corruption Perceptions Index with a score of 31, ranking 122nd out of 180 countries that were scored (Corruption Perceptions Index, 2017). Unlike the other countries that will be discussed, Moldova's score has been decreasing over time, holding a 36 in 2012 and a 31 in 2017. Corruption is rampant there, and it could be concluded that because of that, their democracy has weakened. That low score is only second lowest to Ukraine, and is a few points lower than the average CPI score for Eastern Europe. Moldova's Freedom House score has steadily decreased at a rate similar to its CPI, with a Freedom House score of 64 in 2014 and 58 in 2019. Both scores have decreased around the same time, which indicates a relationship between corruption and democracy. Upon a correlation test from the years 2014-2018 (as those are the years where there are both CPI and Freedom House values available), it was discovered that the R value of Moldova's Corruption Perceptions Index and Freedom House scores is 0.811, indicating a strong positive correlation between levels of corruption and democracy.

Ukraine

Similar to Moldova, Ukraine was part of the Soviet Union. While Moldova has an extremely low score on corruption, according to the Corruption Perceptions Index, Ukraine is even lower (figure 2). With a score of 30 and a ranking of 130, it is clear that the country has some problems with corruption. Despite having a low score as of 2017, it is important to note that in comparison to past years for Ukraine, the score has been steadily increasing over the years.

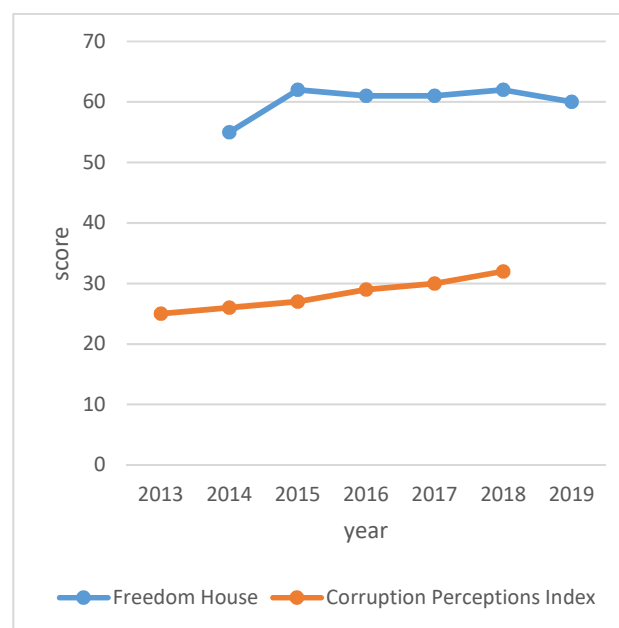


Figure 2: Freedom House and CPI's for Ukraine.

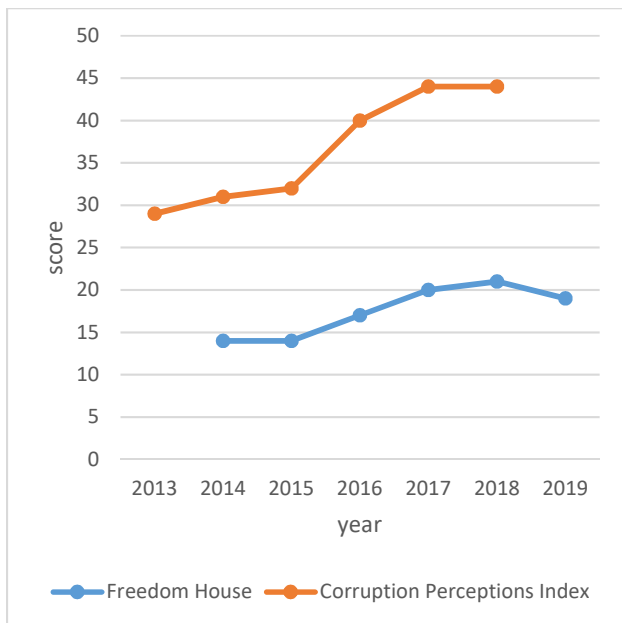


Figure 3: Freedom House and CPI's for Belarus.

Corruption is one of Ukraine's biggest problems and threats to democratic growth, so much so that the government—from pressure from external factors such as a desire for foreign investment—turned toward legislation to try to decrease it. In 2010, former Ukrainian President Yanukovich—who later was removed from power in 2014—proposed a law in order to decrease corruption (Hitch & Kuchma, 2011, p. 844). Despite efforts such as this, corruption has continued.

Ukraine's government actually puts surprisingly little effort into dealing with corruption. For example, in 2017, the National Anti-Corruption Bureau of Ukraine spearheaded a case against Ukraine's National Agency on Corruption Prevention due to allegations of extortion schemes (Freedom House, 2018). Dismissals and other shady business within Ukraine have even caused European Union officials to express concern about Ukraine's corruption problem. Ukraine is improving according to their Corruption Perceptions Index score in recent years, but out of the four countries discussed in this paper, they have the lowest score and are well below the average CPI score for the region. According to the analysis and research, it would be surprising for Ukraine to make any substantial strides in the near future regarding corruption and democracy.

Unlike Moldova, their Freedom House score has been consistently increasing over the years, indicating improvement in their democracy. This steady increase in the Freedom House scores also correlates with a steady increase in CPI scores. Just like in the case with Moldova, the scores indicate a relationship

between democracy and corruption. After a correlation test with the CPI and Freedom House scores for Ukraine from 2014-2018, there is an R value of 0.646, indicating a positive correlation. This correlation is not as strong as the one seen with Moldova, but still strong enough to take note of. As Ukraine's CPI score increases, indicating that corruption has gotten lower, the Freedom House score increases as well.

Belarus

Belarus, formally the Byelorussian Soviet Socialist Republic, has a higher CPI score than Moldova and Ukraine, but still has a low score compared to most of Europe (figure 3). Even though it is a low score compared to Europe as a whole, it is a relatively high score for Eastern Europe. As of 2017, Belarus has a score of 44. It is a bad score, but it isn't as low as one might think for that country. It's above the average of Eastern Europe, and has substantially increased, gaining 15 points in four years (Corruption Perceptions Index, 2017).

Belarus is generally considered a dictatorship, as its president has been in power for 24 years and is currently in his fifth term in office (Besemeres, 2016). Belarus undeniably lacks a main feature of democracy: free, fair, and frequent elections. None were shocked when President Lukashenka emerged victorious in every election. President Lukashenka has had an interest in the European Union, and in order to achieve some credit with them and potentially get some financial backing, he worked to disguise the presidential election and campaigns to appear legitimate. This fell apart when, during a political demonstration against the regime, the members of his security forces beat up and arrested hundreds of citizens, truly showing the corrupt nature of the government system and how far it really was from being a democracy (Besemeres, 2016).

Unsurprisingly, with a lack of democratic accountability and institutions in place comes large amounts of corruption. Corruption is only further fed by the fact that the vast majority of the economy and the media is controlled by the government, that there is no accountability or transparency within the government, and there aren't bodies in place to investigate and bring forth corruption cases (Freedom House, 2018). It's clear that Belarus is not making the same kind of strides toward minimizing corruption and becoming a true democracy that many Eastern European countries are.

Belarus was given extremely low scores from Freedom House, but they have been gradually improving, similar to their CPI scores. As seen in the cases of Moldova and Ukraine, the Freedom House and CPI scores for Belarus indicate a relationship between democracy and corruption, as they increase together and decrease together. As for Belarus's correlation test for the years of 2014-2018 between the CPI and Freedom House scores, there is an R value of 0.974,

suggesting an extremely strong positive correlation and relationship between corruption and democracy in Belarus.

Georgia

While Georgia is by no means a healthy democracy, it is by far the least corrupt country out of the sample chosen for this analysis (figure 4). With a 2017 Corruption Perceptions Index score of 56—a large leap from its score of 49 in 2013, and over 20 points above the Eastern European regional average score—it's clear that the government is consciously making strides toward removing corruption. There has been quite a bit of progress against petty corruption, but there are not the high levels of enforcement of anticorruption measures that the country needs (Freedom House, 2018).

Georgia has free, fair, and frequent elections for the most part, and is working toward making changes to their constitution that will help increase their electoral systems. It could be predicted, based on the information that is given, that Georgia will continue improving. According to the road that Georgia is on, it will most likely become a stronger democracy as time goes on, and corruption will continue to diminish. The corruption in place can definitely hinder the democracy, but Georgia is making conscious anti-corruption efforts in order to prevent this as much as possible.

For Georgia, while their CPI score has increased from 2012 to 2017, their Freedom House score has stayed very constant at a score of 63 or 64 from 2014 to 2019. This doesn't indicate a relationship or correlation between corruption and democracy like the other countries discussed, but does not indicate a lack

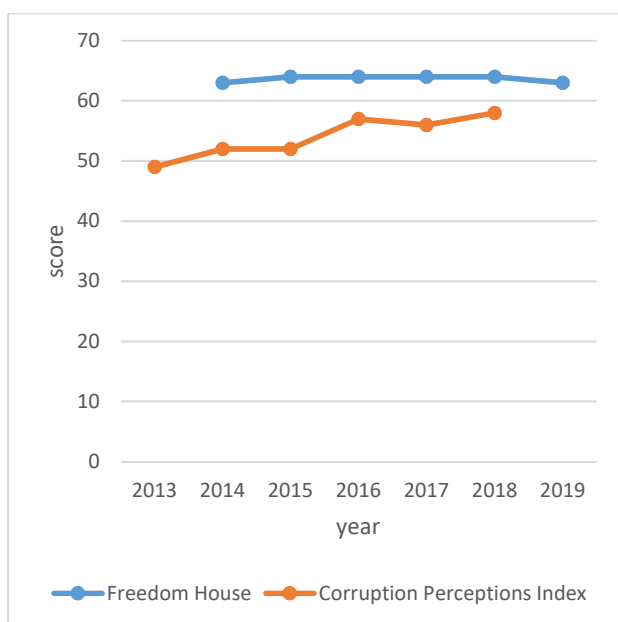


Figure 4: Freedom House and CPI's for Georgia.

of a relationship either. In regards to Georgia's correlation test between CPI and Freedom House scores over 2014-2018, it has the weakest correlation out of the countries included in this study with an R value of 0.593, which still indicates a positive correlation.

Conclusion

After analysis and research on corruption and the role that it plays in democracy and in former Soviet Eastern European countries, it is clear that corruption has no place in a democracy. Based on the research that has been done and comparing the CPI scores and Freedom House scores, it is obvious that there is a positive correlation between democracy and corruption. This correlation is clear when looking at the scores for Moldova, Ukraine, and Belarus. Corruption and democracy do not act independently and they do affect each other, which is indicated by the positive correlation in CPI and Freedom House scores. All four countries had an R value on the correlation test that indicated there was a notable correlation between corruption and democracy.

Due to the fact that the entire claim of democracy is that it is a government for the people by the people, it is not fair nor morally correct for politicians and government officials to take advantage of the system for their own personal gain. When corrupt government officials only look out for themselves, there is a lot that the country—and its citizenry—has to lose. Every country has a different way to deal with corruption and has to find a method of keeping their officials accountable that works for them. Corruption doesn't only affect the people directly tied to it; it hurts the entire political system in some way, and can hurt many people indirectly. Corruption undermines efforts that countries make to democratize.

References

- Azfar, O., Lee, Y., & Swamy, A. (2001). The causes and consequences of corruption. *The Annals of the American Academy of Political and Social Science*, 573, 42-56. Retrieved from <http://www.jstor.org/stable/1049014>
- Besemeres, J. (2016). In A Difficult Neighbourhood: Essays on Russia and East-Central Europe since World War II. Australia: ANU Press. Retrieved from <http://www.jstor.org/stable/j.ctt1rqc96p.17>
- Carasciuc, L. (1999). The Republic of Moldova: corruption as a compelling force in the shadow economy. *SEER: Journal for Labour and Social Affairs in Eastern Europe*, 2(4), 123-136. Retrieved from <http://www.jstor.org/stable/43291683>
- Charron, N., Bagenholm, A. (2016). Ideology, party systems and corruption voting in European democracies. *Electoral Studies*, 41, 35-49.

- Corruption Perceptions Index. (2018). In Transparency International online. Retrieved from https://www.transparency.org/news/feature/corruption_perceptions_index_2017.
- Diamond, L. (2008). *The Spirit of Democracy: The Struggle to Build Free Societies Throughout the World*. New York: Holt.
- Drury, A., Kriekhaus, J., & Luszti, M. (2006). Corruption, democracy, and economic growth. *International Political Science Review / Revue Internationale De Science Politique*, 27(2), 121-136. Retrieved from <http://www.jstor.org/stable/20445044>
- Freedom House. (2019). In Freedom House online. Retrieved from: <https://freedomhouse.org/>
- Hitch, J., & Kuchma, Y. (2011). Ukraine's new anti-corruption law: Will it Really Stop Corruption in Ukraine? *The International Lawyer*, 45(3), 839-855. Retrieved from <http://www.jstor.org/stable/23824394>
- Rose-Ackerman, S. (1996). Corruption and democracy. *Proceedings of the Annual Meeting (American Society of International Law)*, 90, 83-90. Retrieved from <http://www.jstor.org/stable/25659012>
- Warren, M. (2004). What does corruption mean in a democracy? *American Journal of Political Science*, 48(2), 328-343. doi:10.2307/1519886

Other Publications For NKU Student Research And Creative Activity

Pentangle

Research and critical essays on all areas of literary studies. Published by the Pi Omega chapter of Sigma Tau, the International English Honor Society.

[*https://inside.nku.edu/artsci/departments/english/get-involved/pentangle.html*](https://inside.nku.edu/artsci/departments/english/get-involved/pentangle.html)

Loch Norse

The NKU undergraduate literary magazine. Published by Creative Writing students.

[*https://inside.nku.edu/artsci/departments/english/get-involved/loch-norse-magazine.html*](https://inside.nku.edu/artsci/departments/english/get-involved/loch-norse-magazine.html)

Perspectives in History

A peer-reviewed scholarly journal published by NKU members of Alpha Beta Pi, a chapter of the Phi Alpha Theta Interlational History Honor Society.

[*https://inside.nku.edu/artsci/departments/hisgeo/history/student-groups-2/phi-alpha-theta-perspectives.html*](https://inside.nku.edu/artsci/departments/hisgeo/history/student-groups-2/phi-alpha-theta-perspectives.html)

Anthropolgical Perspectives

The journal of the NKU Gamma Chapter of the Lambda Alpha National Anthropolgical Honor Society.

[*https://inside.nku.edu/artsci/departments/sapdept/anthropology/student-organizations/lambda-alpha.html*](https://inside.nku.edu/artsci/departments/sapdept/anthropology/student-organizations/lambda-alpha.html)

The Flame

The journal of the NKU Honors College.

[*https://nku.edu/honors-college/contact/theflame.html*](https://nku.edu/honors-college/contact/theflame.html)



NÝSA 2019 FALL
VOLUME 2

The Northern Kentucky University
Journal Of Student Research