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, & Blinebry, 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) x 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 x 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 x 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) x 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) x 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, $X^2(df=1) = 2.92$, $p = .088$. 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., & Blinebry, 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