

Test of Times New Roman: Effects of Font Type on Mathematics Performance

Jenny Yun-Chen Chan*¹

Lilly-Beth D. Linnell

Cindy Trac

Kathryn C. Drzewiecki

Erin Ottmar

Social Sciences and Policy Studies, Worcester Polytechnic Institute

Worcester, Massachusetts, United States of America

* Corresponding author: chanjyc@eduhk.hk; jchan2@wpi.edu; <https://orcid.org/0000-0003-1479-5935>

Acknowledgments

We would like to thank the students for their participation, Anthony Botelho for his guidance on programming, and the members of the MAPLE Lab and the ASSISTments team for their support.

Funding: None

Conflicts of Interest: None

Availability of Data and Materials: Please contact the authors for any additional information or requests for summary statistics

Authors' Contributions

Jenny Yun-Chen Chan: Conceptualization, Formal analysis, Data curation, Writing, Project administration. **Lilly-Beth D. Linnell:** Methodology, Writing, Visualization. **Cindy Trac:** Software, Writing, Visualization. **Kathryn C. Drzewiecki:** Methodology, Writing, Visualization. **Erin Ottmar:** Conceptualization, Writing

¹ Jenny Yun-Chen Chan, Department of Early Childhood Education, The Education University of Hong Kong; No 10, Lo Ping Rd, Ting Kok, Hong Kong

Test of Times New Roman: Effects of Font Type on Mathematical Performance

Abstract

Mathematics is presented in a variety of font types across materials (e.g., textbooks, online problems); however, little is known about the effects of font type on students' mathematical performance. Undergraduate students ($N = 121$) completed three mathematical tasks in a one-hour online session in one of three font conditions: Times New Roman ($n = 45$), Kalam ($n = 41$), or handwriting ($n = 35$). We examined whether font type impacted students' performance, as measured by accuracy and response time, on the Perceptual Math Equivalence Task, error identification task, and equation solving task. Compared to students in the Kalam and handwriting conditions, students in the Times New Roman condition were less accurate on the Perceptual Math Equivalence Task in which they judged whether two expressions were equivalent or not equivalent. We did not find differences between conditions in performance on error identification and equation solving tasks. The findings have implications for research and practice. Specifically, researchers and educators may choose font types in which they present mathematics information with consideration, as font types may impact students' mathematical processing and performance.

Keywords: Font Type, Times New Roman, Handwriting, Perceptual Features, Mathematical Performance

Test of Times New Roman: Effects of Font Type on Mathematical Performance

Students learn and practice mathematics using a variety of materials, including textbooks, handwritten notes, and educational technology tools; font type, such as Times New Roman and Arial, varies across these materials. In addition to typed fonts, students and teachers often write and solve equations by hand. Perceptual features of symbols, such as proximity and color, do not bear any mathematical implications, yet influence encoding, interpretation, and performance on mathematics (e.g., Alibali et al., 2018; Landy & Goldstone, 2007). We posit that font type, including handwriting style, may also influence the ways in which students perceive, process, and perform mathematics. Although the influences of font type on reading and recall performance have been investigated (Gasser et al., 2005), little is known about its effects in mathematical contexts. Understanding the effects of font type on mathematical thinking and performance has implications for research and practice. Because there are seemingly endless font type options for presenting mathematical information in textbooks and online platforms, it is important to examine how font type variations may influence mathematical thinking and learning. In the current study, we take the first step in this endeavor by exploring whether and how font type influences undergraduate students' performance on three online mathematical tasks. We operationally define mathematical performance as accuracy (i.e., correctness) and efficiency (i.e., response time) on the three mathematical tasks varying in task complexity to examine the font type effects across task contexts. As teaching and learning are increasingly integrated with educational technologies, we focus on students' performance in an online learning platform so the findings may have practical implications for current and future practices of presenting on-screen materials in ways that support mathematical teaching and learning.

1.1 Perceptual Features Influence Mathematical Performance

Substantial empirical work suggests that mathematical reasoning is grounded in perceptual and embodied processes (Alibali & Nathan, 2012; Goldstone et al., 2010; Kirshner & Awtry, 2004; Marghetis et al., 2016). For instance, students use *proximity* as a perceptual cue to group symbols aligning with the order of operations, and write numbers around the multiplication sign closer than around the addition sign (e.g., $3 + 4 \times 5$; Landy & Goldstone, 2007). When the symbols are spaced in a manner that is incongruent (e.g., $3+4 \times 5$) vs. congruent (e.g., $3 + 4 \times 5$) with the order of operations, students are slower at solving and tend to solve expressions incorrectly (i.e., add before multiply; Braithwaite et al., 2016; Harrison et al., 2020; Jiang et al., 2014). Similarly, using *color* to highlight mathematical structures (e.g., the equal sign in equations; Alibali et al., 2018) or to connect mathematical ideas (e.g., corresponding elements between diagrams and formulas; Chan et al., 2019) can improve problem-solving performance. These findings demonstrate that perceptual features, such as proximity and color, do not bear mathematical implications, but they do influence encoding, interpretation, and performance on mathematical tasks.

Similar to other perceptual features, font types used to present mathematical expressions may play a role in how people reason about mathematics. However, little is known about how font type, a perceptual feature that often varies across learning contexts, influences the ways in which students perceive, review, and solve mathematical problems. Prior work has examined the effects of font type on reading comprehension and recall performance. The current study builds on the literature of reading and perceptual influences to examine whether font type, a feature that is seemingly irrelevant to mathematics, impacts students' mathematical performance.

1.2 Effects of Font Type on Reading Performance

While research on the effects of font type in mathematical contexts is limited, several studies have revealed the effects of font type in reading contexts. Gasser and colleagues (2005) gave 149 undergraduate students a reading passage printed on paper in either a serif font (i.e., fonts with decorative strokes at the end of a letter stem; e.g., Times New Roman) or a sans-serif font (i.e., fonts without the decorative strokes; e.g., Arial). They found that students in the serif font condition recalled more details of the reading passage compared to the students in the sans-serif font condition. However, Hojjait and Muniandy (2014) found the reverse effect when students read passages on computer screens. They gave 30 postgraduate students four passages, two presented in serif font and two presented in sans-serif font. They found that students were faster at reading and more accurate on recall when the passage was presented in sans-serif font instead of serif font on a computer screen.

Together, these findings suggest that font type influences reading and recall performance, and this effect may vary by the medium in which information is delivered. Whereas printed materials are often presented with higher resolution in serif fonts, it is recommended that on-screen materials should be presented in sans-serif fonts because these fonts are easier and faster to read, especially when the resolution is relatively poor (Peck, 2003; Wilson, 2001). While it is possible that presenting on-screen materials in sans-serif fonts increases their readability and subsequently improves readers' performance on comprehension and recall, it is unclear whether such effects were solely caused by the fonts themselves or by the ample experiences of reading on-screen materials in sans-serif fonts. With previous findings demonstrating the effects of font type on reading and recall performances, further investigations are warranted to understand its effects in mathematical contexts.

1.3 Handwriting and Mathematical Learning

While reading materials are often presented in typed fonts, mathematics materials are often presented in handwriting. Previous research on learning and note-taking has found that undergraduate students who took notes by hand performed significantly better on conceptual recall tasks than students who typed their notes (Mueller & Oppenheimer, 2014). Further, in that study, students who studied their handwritten notes performed significantly better on assessments than students who studied their typed notes. Although there is evidence supporting the notion that writing and studying handwritten notes may be better for learning compared to typing notes, this benefit may depend on other factors, such as the learning contexts (Fiorella & Mayer, 2017; Luo et al., 2018) and the affordance of the modality to adequately capture the key information (e.g., speed, representations; Morehead et al., 2019).

Mathematics teachers often write out equations and derivations during instruction, and students are used to solving mathematical problems by hand (Anthony et al., 2005). Handwriting offers unique advantages to teaching and learning mathematics as opposed to typing, such as the flexibility of using special symbols or representations (Petrescu, 2014) that are not present in standard keyboards, (i.e., \sum , $\sqrt{\quad}$, \int , etc.). A study reported that students were faster and made fewer errors when writing equations by hand than typing them using keyboards, and more than 50% of the students in the sample preferred handwriting over typing mathematical expressions (Anthony et al., 2008). Although some researchers have compared typing vs. using digital ink tools for writing mathematics (Imtiaz et al., 2017), the effects of *viewing* mathematics presented in typed vs. handwritten fonts is unclear. In textbooks and online problems, mathematics is often presented in typed fonts; however, handwriting is still common in classroom and homework contexts.

Given that the effects of font type may vary depending on the readers' experience with the presentation medium and that handwriting is preferred and often used in mathematical contexts, it is important to advance the knowledge of how font type impacts students' mathematical performance. Specifically, in an online learning context, if students are faster or more accurate on mathematical tasks when the problems are presented in handwriting on a screen, the findings suggest that the experience with viewing and doing mathematics by hand may have facilitative effects on students' mathematical performance in online contexts. However, if students perform better when on-screen mathematical problems are presented in a sans-serif font compared to serif font, the findings will extend prior work on reading performance and suggest that sans-serif font may have facilitative effects on students' mathematical performance in online contexts. Here, we explore whether and how typed and handwritten fonts influence students' performance on three mathematical tasks: Perceptual Math Equivalence, error identification, and equation solving.

1.4 Mathematical Task Contexts

In their everyday life, students experience a variety of task contexts when they learn and do mathematics. As such, we examine the effects of font type on students' mathematical performance in three different task contexts to investigate its effects on various aspects of mathematical performance. These three tasks are designed to measure the processes of *perceiving*, *reviewing*, and *solving* mathematical problems, respectively. These processes are involved at different stages of mathematical learning and reasoning (Booth & Davenport, 2013; Booth et al., 2013; Carroll, 1994; CCSS, 2010; Kellman et al., 2010), and the three tasks together capture the common processes that support students' mathematics problem solving.

When students are presented with an equation, they need to first encode—or in other words, accurately represent—the problem features in their mind (Booth & Davenport, 2013). This perceptual processing from the external problems to the internal representation has implications for downstream performance. As an example, children's and adults' encoding of the equal sign is related to their strategies of solving equations (Crooks & Alibali, 2013). Further, training eighth- and ninth-grade students' ability to identify different forms of equivalent equations (e.g., $5x + 17 = 32$ vs. $5x = 15$) is associated with improved performance on solving algebraic equations (Kellman et al., 2010). Given the importance of noticing abstract structures (Kaput, 1998; Venkat et al., 2019) and understanding mathematical equivalence (Kieran, 2007; Knuth et al., 2006; Stephens et al., 2013), we developed the *Perceptual Math Equivalence Task* based on prior research (Bufford et al., 2018; Kellman et al., 2010; Marghetis et al., 2016) to measure the impacts of font type on students' perceptual processing of mathematics.

While the Perceptual Math Equivalence Task is novel, its conceptualization and design are informed by the prior research on perceptual learning of mathematical features. Specifically, this task built on Marghetis et al.' (2016) algebraic equivalence task in which adults judged whether two algebraic expressions were equivalent or not equivalent (i.e., $a \times b + c \times d$ vs. $b \times a + d \times c$). In that study, Marghetis et al. found that the adults who were more accurate at judging the equivalence of the two expressions were also more likely to perceive the mathematical structures of expressions following the order of operations (i.e., grouping symbols around a higher order operator [e.g., multiplication sign] rather than a lower order operator [e.g., addition sign]). The task was also inspired by Bufford et al.'s (2018) Algebraic Transformation Task in which adults selected one of four equations that mapped onto a target expression (e.g., $12/-s = 8$). In that study, Bufford et al. found that practicing problems within the Algebraic

Transformation Task improved adults' performance on encoding equations, as measured by a psychophysical task. Similarly, asking eighth- and ninth-grade students to practice recognizing equivalent equations in the Algebraic Transformation Task improved their performance on algebraic equation solving (Kellman et al., 2010). Supported by prior studies validating the tasks that asked participants to make quick judgements about the perceptual structure of equivalence, we developed the Perceptual Math Equivalence Task to measure the effects of font type on students' perceptual processing of mathematical expressions.

When learning mathematics, students often study worked examples of problem-solving derivations in textbooks, and teachers often demonstrate the problem-solving process by writing the derivations in front of the class (Carroll, 1994). In addition to studying correct examples, an important aspect of mathematical problem-solving is to learn from errors (Booth et al., 2013; Durkin & Rittle-Johnson, 2012; Ohlsson, 1996), and doing so requires students to review worked examples and identify errors. Therefore, in the current study, we examine whether font type impacts students' performance on identifying errors in worked examples of algebraic equation-solving.

In addition to viewing mathematical expressions and studying worked examples, students need to learn how to solve equations. In fact, prior research examining the influence of perceptual features on mathematical performance often focuses on students' accuracy in equation solving (Braithwaite et al., 2016; Harrison et al., 2020; Jiang et al., 2014; Landy & Goldstone, 2010). Because students' equation-solving performance is a goal of mathematics education (CCSS, 2010) and it is often an outcome of interest in cognitive research, we examine the influences of font type on students' equation-solving performance. By testing the effects of font

type on three tasks that tap into different aspects of mathematical processing, we aim to investigate the ways in which font type impacts mathematical thinking and performance.

1.5 The Current Study

To examine the influences of font type on mathematical performance, we conducted a between-subject experiment presenting undergraduate students with mathematical tasks in one of three font types: Times New Roman, Kalam, and handwriting (see Figure 1). An author wrote the materials for the handwriting condition; this font type served as a proxy for the handwriting students might encounter when learning mathematics. Times New Roman is a serif font type commonly used in paper and online materials, and in mathematics textbooks when presenting equations (e.g., Lay et al., 2016); it served as a comparison to examine whether students' experience with handwriting in mathematical contexts significantly impacted their performance on mathematical tasks. Kalam is a sans-serif font type derived from handwriting and optimized for on-screen use (Indian Type Foundry, 2020); it served as an intermediate font type between Times New Roman and handwriting.

We examined whether font type influenced student performance on (a) perceptual processing of expressions using the Perceptual Math Equivalence Task, (b) reviewing mathematical derivations using the error identification task, and (c) solving problems using the equation-solving task. We designed the three tasks varying in task demands to afford investigation of how font type influenced different aspects of mathematical performance. Because there is limited work on the effects of font type in mathematical contexts, the current study is exploratory and we do not have a priori, directional hypotheses.

2 Methods

2.1 Participants

A total of 167 undergraduate students were recruited from a technical university in the Northeastern United States. The recruitment occurred online through the Social Science Research Participation System. Students completed the study online in an hour and they received psychology course credits for their participation. Of the 167 students enrolled in the study, 46 were excluded due to incompleteness on the experimental tasks ($n = 33$), experimenter error ($n = 10$), or noncompliance with the task instruction ($n = 3$). The final sample comprised 121 undergraduate students (63 females; 54 males; 3 non-binary; 1 other). The students were 18 to 22 years old ($M = 19.70$, $SD = 1.16$), and were White ($n = 71$), Asian ($n = 25$), Hispanic ($n = 10$), Black ($n = 5$), multi-racial ($n = 9$), or prefer not to say ($n = 1$). Students were randomly assigned to one of the following three font conditions: Times New Roman, Kalam, or handwriting. The demographic information by condition is reported in Table 1 (see Table 1). The analytic sample of 121 affords 80% power to detect the small to medium effects of font type on mathematical performance at $f \geq 0.16$. The conventional cutoffs are 0.10, 0.25, and 0.40 for small, medium, and large effect sizes respectively.

2.2 Procedure

Informed consent was obtained online prior to the study. Students were informed that they would complete some mathematical tasks and answer some survey questions on their web browser, but they were naive about the study's focus on font type. After obtaining informed consent, students completed all three mathematical tasks in one of the three font types and reported their mathematics experiences and demographic information. All mathematical problems in the three tasks were identical between conditions except for the font type in which the problems were presented (see Figure 1). The task instructions and survey questions were presented in another font (i.e., Verdana; a sans-serif font), to avoid students receiving more

exposure to one of the three font types tested in the current study. We used Verdana font for the instructions and survey because it was the default font type used in the study platform and a sans-serif font commonly used in online materials. At the end of the study, students received a debriefing form explaining the purpose of the study and were requested to not share the study information with their classmates.

2.3 Measures and Task Procedures

In the three mathematical tasks, students were instructed to respond as quickly and accurately as possible, mimicking timed tests, so we could examine how font type influenced the initial processing of mathematics without much deliberation. In each task, the problems varied in complexity and were presented one at a time. The order of the problems in each task were randomized to avoid presenting similar problems (e.g., $-(x+5) = 20$ and $-3(y+4) = 18$) consecutively or presenting problems in the order of difficulty. After the initial randomization, all students viewed the problems in the same order. We recorded accuracy and response time on each problem and aggregated these values across problems for each task to conduct student-level analyses.

2.3.1 Perceptual Math Equivalence Task

In this task, students determined whether two algebraic expressions (e.g., $2 \cdot (3y-6)$ and $-2 \cdot 6 + 2 \cdot 3y$) were equivalent or not equivalent by selecting their answer on screen. The instruction was, "In the next set of problems, you will see two expressions. Your job is to decide whether these two expressions are equivalent or not equivalent. Do it as quickly and as accurately as you can!" The task comprised 16 problems that covered a wide range of mathematical topics at the middle-school level, such as arithmetic operations (8 items), inverse operations (4 items), and factoring and distribution (4 items). Both the accuracy of their judgment and the response time

was recorded. The reliability of this task as measured by the Kuder-Richardson-20 coefficient (KR-20; a reliability measure for items with binary scoring) was 0.37. The poor reliability might be due to the limited number of items and the diverse topics covered in the task.

2.3.2 Error Identification Task

In this task, students identified the line in worked examples that contained an error. The instruction was, "In the next set of problems, you will find errors in worked examples. Please look at each worked example carefully. Assuming the first line of the equation is correct, identify where the first error is made in the process of solving for the variable. Please answer each question as quickly and as accurately as possible!" The task comprised 24 incorrect worked examples. Each worked example contained one incorrect line that reflected one of three following types of errors commonly made by middle-school students (Booth et al., 2014): dropping the negative sign (see Figure 2a; 8 items), incorrectly applying the distributive property (see Figure 2b; 8 items) and performing an incorrect operation across the equal sign (see Figure 2c; 8 items). The first line of each worked example was adapted from materials in a middle-school algebra project (Rittle-Johnson & Star, 2007). The worked examples were constructed such that the correct and the incorrect solutions were integers. Although the worked examples ranged from four to seven lines long, students were always given four multiple-choice options. For the worked examples that consisted of only four lines (9 items), the multiple-choice options were: (a) Line 2, (b) Line 3, (c) Line 4, (d) No lines contain errors. For the worked examples that consisted of more than four lines (15 items), the last option was (d) Line 5. None of the errors occurred on Line 6 or Line 7. The reliability of this task was $KR-20 = 0.90$, indicating excellent reliability.

2.3.3 Equation-Solving Task

In this task, students solved 12 one-variable equations (e.g., $5(x+2) = 3(x+2)+10$) selected and adapted from materials developed by Rittle-Johnson and Star (2007) for seventh-grade students. A sample instruction was, "Please solve for x by hand, then enter the value of the variable." The problems varied in complexity (e.g., easy problem: $4(x+2) = 12$; complex problem: $-3(x+5+3x)-5(x+5+3x) = 24$); there were four easy problems that involved one set of parentheses, and eight complex problems that involved two sets of parentheses. The reliability of this task was $KR-20 = 0.72$, indicating acceptable reliability.

2.4 Analytic Approach

First, we conducted descriptive analysis on the accuracy and response time in the three tasks to examine the data distribution and to inform subsequent analyses. To examine whether font type impacts performance on each mathematical task, we conducted two one-way ANOVAs to test the effect of font type condition on task accuracy and response time, respectively. Parallel models were conducted for each of the three mathematical tasks. We used Bonferroni-corrected pairwise comparisons for all post-hoc tests.

3 Results

3.1 Preliminary Analyses

Overall, undergraduate students were accurate on the Perceptual Math Equivalence ($M = 84.9\%$, $SD = 10.5\%$), error identification ($M = 89.8\%$, $SD = 16.2\%$), and problem-solving tasks ($M = 86.2\%$, $SD = 16.7\%$), confirming that most students had the algebraic knowledge to solve these problems. If students were responding to these problems by guessing, the percent accuracy should be 50% for the Perceptual Math Equivalence Task and 25% for the error identification task. Because of the relatively high accuracy on the three tasks among our sample, we reviewed the skewness and kurtosis of students' accuracy to check the normality of the data distribution.

We found that the skewness and kurtosis were beyond ± 2 for students' accuracy on the error identification and equation-solving tasks, suggesting non-normal distributions (George & Mallery, 2010). In response to the potential violation of the normality assumption, we reported the statistical values from the non-parametric Kruskal-Wallis Test for the ANOVAs on error identification accuracy and equation-solving accuracy. Students took an average of 8.96 seconds ($SD = 3.40$) on each Perceptual Math Equivalence problem, 21.98 seconds ($SD = 10.06$) on each error identification problem, and 43.33 seconds ($SD = 19.60$) on each equation-solving problem. As expected, the response times suggested that the three tasks tapped into different levels of mathematical processes —perceiving, reviewing, and solving, respectively (see Table 2).

3.2 Primary Analysis

3.2.1 Perceptual Math Equivalence Task

A one-way ANOVA revealed a significant main effect of condition on response accuracy, $F(2,118) = 4.67$, $p = .011$, $\eta_p^2 = .073$. Post-hoc tests revealed that students in the Times New Roman condition ($M = 81.3\%$, $SD = 10.8\%$) were less accurate on the Perceptual Math Equivalence Task compared to the students in the Kalam condition ($M = 86.9\%$, $SD = 9.9\%$), $p = .034$, and the students in the handwriting condition ($M = 87.3\%$, $SD = 9.8\%$), $p = .028$. The response accuracy on this task was not significantly different between students in the Kalam condition and students in the handwriting condition, $p = 1.00$ (see Figure 3). A one-way ANOVA on response time revealed that the effect of font type was not significant, $F(2,118) = 1.44$, $p = .241$.

3.2.2 Error Identification Task

A one-way ANOVA with the Kruskal-Wallis Test revealed that the effect of font type condition on error identification accuracy was not significant, $F(2,118) = 0.040$, $p = .908$. A one-

way ANOVA on response time also revealed that the differences between conditions were not significant, $F(2,118) = 0.88, p = .417$.

3.2.3 Equation-Solving Task

A one-way ANOVA with the Kruskal-Wallis Test revealed that the effect of font type condition on equation-solving accuracy was not significant, $F(2,118) = 1.383, p = .501$. A one-way ANOVA on response time also revealed that the differences between conditions were not significant, $F(2,118) = 0.81, p = .447$.

4 Discussion

In summary, we found that undergraduate students were less accurate at judging whether two expressions were equivalent or not equivalent when the expressions were presented in Times New Roman compared to Kalam or handwriting. We did not find differences between conditions on students' response time in the Perceptual Math Equivalence Task. For the error identification task and the equation-solving task, we also did not find condition effects on accuracy and response time. Below, we discuss the findings and their implications for research and practice.

4.1 Times New Roman vs. Kalam vs. Handwriting

Extending previous studies that revealed the effects of font type on reading and recall performance (Gasser et al., 2005; Hojjait & Muniandy, 2014), we found that font type also impacted students' mathematical performance, as measured by their accuracy of judging equivalent expressions. Specifically, students in the Kalam condition (a sans-serif font) and the handwriting condition were more accurate on the Perceptual Math Equivalence Task compared to students in the Times New Roman condition (a serif font). This study contributes novel findings to perceptual learning theories and suggests that, just like proximity and color of

symbols (Alibali et al., 2018; Landy & Goldstone, 2007), font type may influence students' performance on mathematical tasks.

The significant difference in accuracy on the Perceptual Math Equivalence Task between handwriting and Times New Roman conditions suggests the influences of perceptual features on mathematical performance. Specifically, the experience of writing mathematical derivations and seeing mathematics presented in handwriting may contribute to how surface-level perceptual features, like font type, influence students' mathematical performance. Prior research has found that students prefer writing equations by hand and make fewer errors doing so compared to typing (Anthony et al., 2008); however, little research has been conducted on the effects of *viewing* font type in mathematical contexts. The handwriting used in the current study was created by an author and was not representative of any participants' experience in mathematical learning, yet we found a facilitative effect of viewing mathematics in handwriting compared to Times New Roman on screen. This finding suggests that the handwriting we used in the current study may resemble the writing and materials students often encounter in mathematics, and that handwriting may provide unique affordances for connecting with students' experience of viewing and writing mathematics. This connection to the experience of writing and solving mathematics by hand may support mathematical performance, just as studying handwritten notes improves learning compared to studying typed notes (Mueller & Oppenheimer, 2014). Further, the positive effect of handwriting may not be specific to the handwriting that students personally experience but may be more general to any handwritten mathematics.

With respect to students' accuracy on the Perceptual Math Equivalence Task in the Kalam condition, we found that students in the Kalam condition outperformed students in the Times New Roman condition, but their accuracy was comparable to that of the handwriting

condition. One possible explanation is that both Kalam and handwriting fonts resemble the writing and materials students often encounter in mathematics, and the mechanism through which Kalam font supports students' performance may be similar to that of handwriting. Alternatively, the reason underlying the advantage of Kalam font may be aligned with prior research on reading. Specifically, the significant difference between Kalam and Times New Roman conditions may be related to the advantage of presenting materials in sans-serif font as opposed to serif font on screen (Hojjait & Muniandy, 2014). This interpretation supports the recommendation that on-screen materials should be presented in sans-serif font (Peck, 2003; Wilson, 2001) and further extends this recommendation from reading to mathematics.

More research is needed to explore these two hypotheses and delineate the underlying mechanisms through which font type influences mathematical thinking and learning. However, our study provides an important foundation for examining the effects of font type on students' mathematical performance, and a starting point for informing practices that facilitate students' processing of mathematics. In particular, if future research reveals that presenting mathematical information in handwriting and handwriting-like fonts supports mathematical performance regardless of the medium, it may be worthwhile to present both printed and on-screen materials in these fonts. If future research reveals differential benefits in mathematical performance based on font type and medium (i.e., an advantage of serif fonts in printed materials and sans-serif font in on-screen materials), it may be worthwhile to vary the font type based on the medium. These subtle changes can be low cost yet have notable impacts on supporting students' mathematical thinking and learning.

4.2 Influences of Task Context

Although we found the effect of font type on the accuracy of the Perceptual Math Equivalence Task, we did not find effects of font type on students' performance in the error identification task or the equation-solving task. The findings suggest that handwriting and Kalam font may be beneficial for perceptual processing of mathematical information, such as quickly judging the equivalence of two expressions; however, the benefit may not extend to more cognitively involved tasks, such as error identification and equation solving, at least among undergraduate students who have acquired the mathematical content tested in the current study. Although the effect of font type may be limited to perceptual processing of mathematics in the current study, prior work has demonstrated the importance of perceptual processing on student learning and problem solving (Crooks & Alibali, 2013; Kellman et al., 2010; McNeil & Alibali, 2004). Therefore, the effect of font type on perceptual processing of mathematics may still be important and relevant to mathematical reasoning.

We noted that in all three tasks, we instructed participants to respond as quickly and accurately as possible. This instruction is often used in experimental studies and mimics timed tests in classrooms; however, it remains unclear whether and how the effects of font type may emerge differently depending on the task instruction and contexts, potentially limiting the generalizability of the current findings to real-world problem solving. Even so, the current findings provide a foundation for future research investigating the influences of font type on mathematical performance and demonstrate support for the perceptual learning theory positing that mathematical reasoning is grounded in perceptual and embodied processes (Alibali & Nathan, 2012; Goldstone et al., 2010; Kirshner & Awtry, 2004; Marghetis et al., 2016).

4.3 Limitations and Future Directions

Given that our study is exploratory in nature, there are several limitations worth noting.

First, the sample is limited to undergraduate students, and they presumably have acquired the mathematical content tested in the current study. We chose to focus on the middle-school content because undergraduate students might have varying levels of mathematical knowledge due to differences in mathematics requirements across majors, and our content focus would ensure that all students had the knowledge to complete the tasks. Even among undergraduate students who were likely to be proficient in solving middle-school algebra problems, we found the effects of font type emerging in an aspect of mathematical performance. Future work should replicate this study with middle-school students to examine developmental differences and the effects of font type on mathematical performance among students who are still learning the content.

Second, while the current study only tested three font types varying in the modality through which the materials are created—from typing to writing, we tested their effects on students' performance in three mathematical tasks and carefully controlled for the font type students encountered throughout the study by using Verdana for non-experimental content. Further, the Perceptual Math Equivalence Task was a novel measure with poor reliability. However, the task was developed based on extensive prior research and included relatively few items that covered a wide range of algebraic topics. Building on the current study, future research would benefit from examining how measures like the Perceptual Math Equivalence Task can be used in online settings to assess students' perceptual processes of mathematics, validating the novel measure, and replicating and extending the findings with other font types. Doing so will contribute to delineating the mechanisms underlying the font type effects we observed in the current study. Specifically, if students perform better on mathematical tasks when the information is presented in handwriting or handwriting-like fonts (e.g., Kalam) as

opposed to other serif (e.g., Times New Roman) or sans-serif fonts (e.g., Arial), the findings may provide support for the influences of experience with writing and viewing mathematical writing on performance. If students perform better when the information is presented in sans-serif (e.g., Arial) as opposed to serif (e.g., Times New Roman) fonts, the findings may support the recommendation that sans-serif fonts are more suitable for presenting learning materials on screen compared to serif fonts.

Finally, due to the COVID-19 pandemic, the current study consists of only online computerized tasks. While the study is limited to computer-based contexts and the findings may not generalize to other task contexts, the computerized tasks provide affordances for measuring response time precisely. Further, given the increasing prevalence of online learning due to the pandemic, the current work provides an important foundation for research that informs practices for online instruction. When in-person data collection is possible again, future work can investigate whether the effects of font type on mathematical performance vary depending on the medium, such as textbook and lecture notes. The findings will extend prior research on reading and recall performance and have implications for designing research and learning materials across learning subjects and contexts.

4.4 Conclusion

In sum, this study investigates how font type influences aspects of mathematical thinking and performance, and suggests that even subtle design decisions, such as the font type chosen to present mathematical symbols, may contribute to how people reason about mathematics. Although the effects of font type only emerged in the Perceptual Math Equivalence Task and did not persist in other tasks that require reviewing or solving equations, the findings nonetheless provide evidence for the relevance of seemingly irrelevant features. If font type influences

mathematical performance, researchers and educators should be aware of its effects in experiments, classrooms, online platforms, and instructional materials. For instance, as teachers and content designers present instructional materials in online contexts, they may deliberate and make intentional decisions on the font type in which to present the materials.

References

- Alibali, M. W., Crooks, N. M., & McNeil, N. M. (2018). Perceptual support promotes strategy generation: Evidence from equation solving. *British Journal of Developmental Psychology*, 36(2), 153–168. <https://doi.org/10.1111/bjdp.12203>
- Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, 21(2), 247–286. <https://doi.org/10.1080/10508406.2011.611446>
- Anthony, L., Yang, J., & Koedinger, K. R. (2005). Evaluation of multimodal input for entering mathematical equations on the computer. *Conference on Human Factors in Computing Systems - Proceedings*, 1184–1187. <https://doi.org/10.1145/1056808.1056872>
- Anthony, L., Yang, J., & Koedinger, K. R. (2008). Toward next-generation, intelligent tutors: Adding natural handwriting input. *IEEE Multimedia*, 15(3), 64–68. <https://doi.org/10.1109/MMUL.2008.73>
- Booth, J. L., Barbieri, C., Eyer, F., & Paré-Blagoev, E. J. (2014). Persistent and pernicious errors in algebraic problem solving. *Journal of Problem Solving*, 7(1), 10–23. <https://doi.org/10.7771/1932-6246.1161>
- Booth, J. L., & Davenport, J. L. (2013). The role of problem representation and feature knowledge in algebraic equation-solving. *Journal of Mathematical Behavior*, 32(3), 415–423. <https://doi.org/10.1016/j.jmathb.2013.04.003>
- Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning and Instruction*, 25, 24–34. <https://doi.org/10.1016/j.learninstruc.2012.11.002>

- Braithwaite, D. W., Goldstone, R. L., van der Maas, H. L. J., & Landy, D. (2016). Non-formal mechanisms in mathematical cognitive development: The case of arithmetic. *Cognition*, 149, 40–55. <https://doi.org/10.1016/j.cognition.2016.01.004>
- Bufford, C., Mettler, E., Geller, E., Kellman, P., Mettler, E., & Bufford, C. (2018). The Psychophysics of algebra: Mathematics perceptual learning interventions produce lasting changes in the perceptual encoding of mathematical objects. *Journal of Vision*, 18(10), 1073. <https://doi.org/10.1167/18.10.1073>
- Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, 86(3), 360–367. <https://doi.org/10.1037/0022-0663.86.3.360>
- Chan, J. Y. C., Sidney, P. G., & Alibali, M. W. (2019). Corresponding color coding facilitates learning of area measurement. In S. Otten, A. G. Candela, Z. de Araujo, C. Haines, & C. Munter (Eds.), *Proceedings of the forty-first annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (p. 437). St Louis, MO: University of Missouri. Retrieved from <https://www.pmena.org/pmenaproceedings/PMENA 41 2019 Proceedings.pdf>
- Common Core State Standards: National Governors Association and Council of Chief School Officers (2010). Retrieved June 2020, from: <http://www.corestandards.org/>.
- Crooks, N. M., & Alibali, M. W. (2013). Noticing relevant problem features: activating prior knowledge affects problem solving by guiding encoding. *Frontiers in Psychology*, 4(November), 884. <https://doi.org/10.3389/fpsyg.2013.00884>

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, 22(3), 206–214.

<https://doi.org/10.1016/j.learninstruc.2011.11.001>

Fiorella, L., & Mayer, R. E. (2017). Spontaneous spatial strategy use in learning from scientific text. *Contemporary Educational Psychology*, 49, 66–79.

<https://doi.org/10.1016/j.cedpsych.2017.01.002>

Gasser, M., Boeke, J., Haffernan, M., & Tan, R. (2005). The influence of font type on information recall. *North American Journal of Psychology*, 7(2), 181–188.

George, D., & Mallery, P. (2010). SPSS for Windows Step by Step: A Simple Guide and Reference 18.0 Update.

Goldstone, R. L., Landy, D., & Son, J. Y. (2010). The education of perception. *Topics in Cognitive Science*, 2(2), 265–284. <https://doi.org/10.1111/j.1756-8765.2009.01055.x>

Harrison, A., Smith, H., Hulse, T., & Ottmar, E. R. (2020). Spacing out! Manipulating spatial features in mathematical expressions affects performance. *Journal of Numerical Cognition*, 6(2), 186–203. <https://doi.org/10.5964/jnc.v6i2.243>

Hojjait, N., & Muniandy, B. (2014). The effects of font type and spacing of text for online readability and performance. *Contemporary Educational Psychology*, 5(2), 161–174.

<https://doi.org/10.1038/nphys2346>

Indian Type Foundry, (2020). Google font: Kalam. Retrieved from

<https://fonts.google.com/specimen/Kalam#about>

Imtiaz, M. A., Blagojevic, R., Luxton-Reilly, A., & Plimmer, B. (2017). A survey of intelligent digital ink tools use in STEM education. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3014812.3014825>

- Jiang, M. J., Cooper, J. L., & Alibali, M. W. (2014). Spatial factors influence arithmetic performance: The case of the minus sign. *Quarterly Journal of Experimental Psychology*, 67(8), 1626–1642. <https://doi.org/10.1080/17470218.2014.898669>
- Kaput, J. J. (1998). Representations, inscriptions, descriptions and learning: A kaleidoscope of windows. *The Journal of Mathematical Behavior*, 17(2), 265–281. [https://doi.org/10.1016/s0364-0213\(99\)80062-7](https://doi.org/10.1016/s0364-0213(99)80062-7)
- Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, 2(2), 285–305. <https://doi.org/10.1111/j.1756-8765.2009.01053.x>
- Kieran, C. (2007). Learning and teaching of algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 2, pp. 707–762). Charlotte, NC: Information Age.
- Kirshner, D., & Awtry, T. (2004). Visual salience of algebraic transformations. *Journal for Research in Mathematics Education*, 35(4), 224–257. <https://doi.org/10.2307/30034809>
- Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, 37(4), 297–312. <http://doi.org/10.2307/30034852>
- Landy, D., & Goldstone, R. L. (2007). Formal notations are diagrams: Evidence from a production task. *Memory & Cognition*, 35(8), 2033–2040. <https://doi.org/10.3758/BF03192935>

- Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly Journal of Experimental Psychology*, 63(10), 1953–1968.
<https://doi.org/10.1080/17470211003787619>
- Lay, D. C., Lay, S. R., & McDonald, J. (2016). *Linear algebra and its applications*. Boston: Pearson.
- Luo, L., Kiewra, K. A., Flanigan, A. E., & Peteranetz, M. S. (2018). Laptop versus longhand note taking: Effects on lecture notes and achievement. *Instructional Science*, 46(6), 947–971. <https://doi.org/10.1007/s11251-018-9458-0>
- Marghetis, T., Landy, D., & Goldstone, R. L. (2016). Mastering algebra retrains the visual system to perceive hierarchical structure in equations. *Cognitive Research: Principles and Implications*, 2(25). <https://doi.org/10.1186/s41235-016-0020-9>
- McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science*, 28(3), 451–466.
https://doi.org/10.1207/s15516709cog2803_7
- Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). How much mightier Is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). *Educational Psychology Review*, 31(3), 753–780. <https://doi.org/10.1007/s10648-019-09468-2>
- Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard. *Psychological Science*, 25(6), 1159–1168. <https://doi.org/10.1177/0956797614524581>
- Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, 103(2), 241.
<https://doi.org/10.1037/0033-295X.103.2.241>
- Peck, W. (2003). *Great web typography*. New York: Wiley.

Petrescu, A. (2014). Typing or writing? A dilemma of the digital era. In *The 10th International Scientific Conference eLearning and software for Education Bucharest* (pp. 393–397).

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, 99(3), 561–574. <https://doi.org/10.1037/0022-0663.99.3.561>

Stephens, A. C., Knuth, E. J., Blanton, M. L., Isler, I., Gardiner, A. M., & Marum, T. (2013). Equation structure and the meaning of the equal sign: The impact of task selection in eliciting elementary students' understandings. *The Journal of Mathematical Behavior*, 32(2), 173–182. <http://doi.org/10.1016/j.jmathb.2013.02.001>

Venkat, H., Askew, M., Watson, A., & Mason, J. (2019). Architecture of mathematical structure. *For the Learning of Mathematics*, 39(1), 13–17. Retrieved from <https://flm-journal.org/Articles/15CDA7D47D380050F5587BF8C2BCB8.pdf>

Wilson, R. (2001). *Text font readability study*. Retrieved on 14 July 2021 from <http://www.wilsonweb.com/wmt6/html-email-fonts.htm>.

Figure 1

The same equation is presented in (a) Times New Roman (b) Kalam, or (c) handwriting

Figure 2

Three incorrect worked examples that vary in complexity, error type, and the line containing an error. Panel (a) shows the error of dropping the negative sign on Line 2. Panel (b) shows the error of incorrectly applying the distributive property on Line 3. Panel (c) shows the error of performing an incorrect operation across the equal sign (i.e., subtracting instead of dividing by 10 on both sides of the equation) on Line 5.

Figure 3

The main effect of font type on students' percent accuracy of the Perceptual Math Equivalence Task. Error bars represent standard deviations.

a) Times New Roman	b) Kalam	c) Handwriting
$4(x+2) = 12$	$4(x+2) = 12$	$4(x+2) = 12$

a)		b)		c)
$-2(x+5) = 20$	Line 1	$1(x+9)+2(x+9) = -12$	Line 1	
$2x+10 = 20$	Line 2	$3(x+9) = -12$	Line 2	
$2x = 10$	Line 3	$3x+9 = -12$	Line 3	
$x = 5$	Line 4	$3x = -21$	Line 4	
		$x = -7$	Line 5	
				$17(x+6) = 7(x+6)+100$
				$10(x+6) = 100$
				$10x+60 = 100$
				$10x = 40$
				$x = 30$

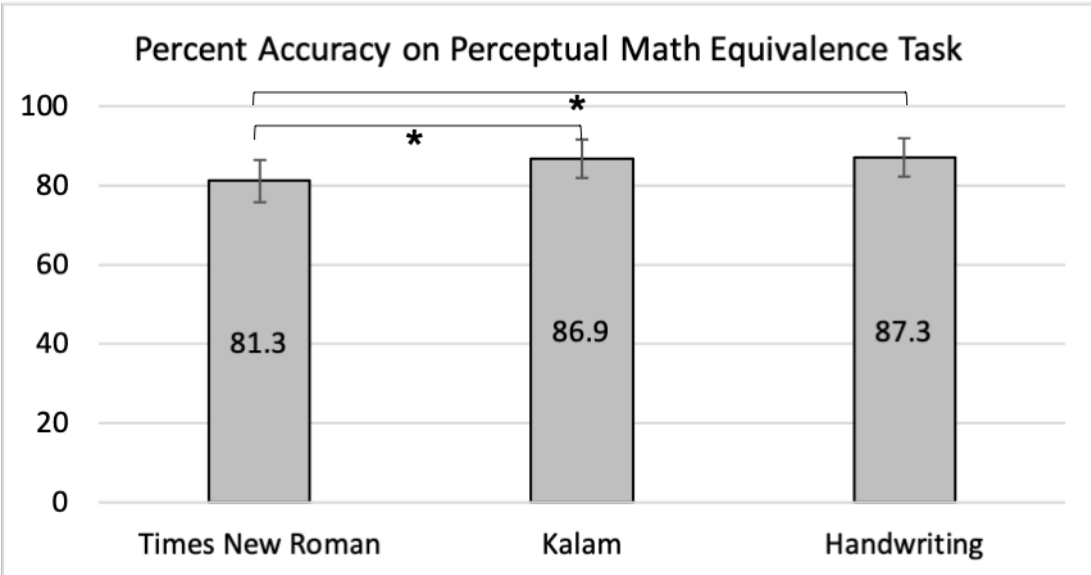


Table 1

<i>Demographic Information by Condition</i>			
	Times New Roman (<i>n</i> = 45)	Kalam (<i>n</i> = 41)	Handwriting (<i>n</i> = 35)
Age (year)	19.80 (1.08)	19.60 (1.22)	19.70 (1.21)
Gender			
Female	23	23	17
Male	22	17	15
Non-binary	0	1	2
Other	0	0	1
Race			
White	21	29	21
Asian	9	9	7
Hispanic	5	2	3
Black	4	1	0
Multi-racial	5	0	4
Prefer not to say	1	0	0

Table 2

Means and Standard Deviations of Percent Accuracy and Response Time (in Seconds) on the Three Mathematical Tasks for the Three Font Type Conditions

Task	Times New Roman (<i>n</i> = 45) <i>M</i> (<i>SD</i>)	Kalam (<i>n</i> = 41) <i>M</i> (<i>SD</i>)	Handwriting (<i>n</i> = 35) <i>M</i> (<i>SD</i>)
Perceptual Math Equivalence			
Accuracy	81.3% (10.7%)	86.9% (9.9%)	87.3% (9.76%)
Response Time	8.91 (2.90)	9.60 (4.02)	8.28 (3.13)
Error Identification			
Accuracy	88.8% (15.9%)	91.0% (14.4%)	89.6% (18.7%)
Response Time	23.03 (12.37)	22.41 (8.81)	20.12 (7.90)
Equation Solving			
Accuracy	87.2% (17.8%)	86.4% (16.7%)	84.8% (15.5%)
Response Time	40.90 (21.61)	46.31 (15.58)	42.87 (21.14)