FORMAT READABILITY ENHANCING IN BASIC MATHEMATICAL OPERATIONS

Caterina B Azzarello School of Psychological Sciences University of Northern Colorado Greeley CO, USA caterina.azzarello@unco.edu
Dave B Miller Department of Mechanical Engineering Tufts University Medford MA, USA dave.miller@tufts.edu 0000-0001-9706-4630
Ben D Sawyer Department of Industrial Engineering and Management Systems University of Central Florida Orlando, USA sawyer@inhumanfactors.com
Joanna E Lewis School of Psychological Sciences University of Northern Colorado Greeley CO, USA joanna.lewis@unco.edu 0000-0003-0974-2230

Individuated font selection, which can increase text reading speed, may be able to increase mathematical expression reading speed and influence reasoning accuracy. To investigate whether the same font that increases a participant's reading speed enhances mathematical reading, we compared their speed in evaluating mathematical expressions as true or false presented in their fastest reading font as determined by empirical test and with a control font, Times Roman. Participants were faster in completing mathematical problems when using typography selected for individual readability, but no difference occurred in task accuracy, matching patterns previously seen in interlude reading. Future research should assess the impact of elements of time pressure, math complexity, numeric versus text-based questions, and associated math anxiety.

INTRODUCTION

Readability is a measure of how written text influences the ability to comprehend materials, reading speed, and engagement (Dale & Chall, 1949). For both physically printed texts and screen-presented texts, content readability can be increased by selecting shorter sentences in the passages, shorter words in the sentences, and familiar words (Perera, 1980). A growing body of work suggests that beyond content, typographical aspects of a document may strongly influence reading speed and comprehension, through format readability adjustment (Beier et. al., 2022; Wallace et. al., 2022; Wallace et. Al, 2022; Cai et. al; 2022; Wallace et. al., 2021). Individuated changes to typography, such as font selection, are possible on screen-based text in ways that are not practical in printed matter. One practical application of the flexibility afforded by screens is to provide better readability in screen-presented text by selecting typographic variables to match the reader's needs.

Improving the readability of materials through customized typography has previously focused on text-based content. For example, ease of reading (Chaparro et. al, 2010), assessing comprehension of passages (Cai et. al, 2022; Wallace et. al, 2021), reading speed (Blanc-Goldhammer & MacKenzie, 2018), or effectiveness of reading short text in complex environments (Sawyer et. al, 2020; Sawyer et. al. 2020). Specific elements of the typography can influence reading performance, such as case (Kember & Varley, 1987), spacing or spatial frequencies (Patching & Jordan, 2005; Paterson & Jordan, 2010), or the use of a sans-serif or serif font (Moret-Tatay & Perea, 2011; Arditi & Cho, 2005; Josephson, 2008). Manipulation of font size and font selection have been shown to increase reading speed (Cai et. al, 2022; Soleimani & Mohammad, 2012). Individuation is necessary as one font does not optimize speed for all readers—one reader's fastest font may be another's slowest (Wallace et. al, 2022). Specifically, typographic individuation can help a reader to find a format which promotes success in their reading task, and digital technologies can allow this format to appear across all a reader's screens, such as files, applications, and devices. To determine one's individuated font and spacing parameters, empirical testing is needed, but fortunately a short reading test can determine these. While there is a speed-comprehension tradeoff, this appears to be relatively small and thus optimizing for speed likely does not incur serious comprehension penalties, but more research is required (Wallace et. al., 2021).

Previous assessment of the benefits of individuated typography has primarily focused on text-based or narrative content. However, we aim to seek if the same optimizations that work for increasing reading speed and processing also personally benefit speed-optimized numeric typography formatting, as in mathematics. Mathematics is complex for readers who may be much more fluent in terms of reading other written content, and this can lead to errors and an elevated level of anxiety associated with interacting with this type of material (Ashcraft & Krause, 2007). With mathematics already having been stigmatized as 'difficult' (Aschraft, 2002), an opportunity to minimize barriers to learning seems prudent to explore.

In our current work, we aim to evaluate whether individuated, format readability-oriented typography settings intended for increasing reading speed in narrative content increases speed of evaluating mathematical expressions. Investigating this allows us to assess if the benefits of speed-optimal typography are equally beneficial as task difficulty increases, as well as if speed-optimal typography adjustments may reduce state math anxiety. In this research, we study the performance of students evaluating simple mathematical expressions, as a first step towards understanding ways to use typography individuation in math more broadly.

METHODS

Participants

We recruited a sample of 30 university students (*M* age = 19.56; 9 Males, 21 Females, 0 Other) from The University of Northern Colorado, a mid-sized public university. An a-priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007) to determine the minimum sample size required to test the study hypothesis. Results indicated the required sample size to achieve 80% power for detecting a medium effect (Cohen, 1988), at a significance criterion of α = .05, was N = 23 for a two-tailed t-test. Thus, the obtained sample size of 28 is adequate to evaluate the study hypothesis, that individuated font and spacing will increase the speed and/or accuracy when evaluating whether mathematical expressions are true or false.

Design and Procedure

Participants completed an online assessment (https://readabilitylab.xyz) to determine both their fastest font (Times, Roboto, EB Garamond, Montserrat, or Noto Sans) and spacing (-0.05, 0.0, +0.05, +0.10, +0.30 em) for a passage-length (approximately 150 words) reading task. Table 1 shows the number of participants with each font as their fastest, and the number with each spacing as their fastest.

Table 1

Number of participants with each font or spacing as their fastest.

Fastest reading Font	*Times*	*Roboto*	*EB Garamond*	*Montserrat*	*Noto Sans*
	9	10	4	5	2
Fastest reading Spacing	*-0.05em*	*0em*	*+0.05em*	*+0.10em*	*+0.30em*
	2	9	6	10	1

After participants determined their font and spacing settings, Participants evaluated two blocks of 80 basic arithmetic problems indicating whether each expression was true or false, one block utilizing a 'standard' typography setting (Font: Times; Spacing: 0 em) and one block utilizing their fastest typography settings based on the individualized initial assessment (Times, Roboto, EB Garamond, Noto Sans, or Montserrat). The arithmetic problems consisted of two pairs of two-digit numbers added together to a single solution that was either correct (e.g., 45 + 23 = 68) or incorrect (e.g., 10 + 15 = 26), 50% of the solutions were incorrect. The error in the computations was in 50% of the incorrect statements due to an error in the tens place, 50% due to an error in the ones place, based on Hunt and Sandhu's previous work (2017). The blocks of problems were presented in random order, to mitigate order effects of typography settings [Times 0em then optimal; optimal then Times 0em]. Participants viewed the entirety of the problem set on a single scrolling page, and indicated whether each expression was true or false by pressing 0 to indicate incorrect, and 1 to indicate correct.

RESULTS

Two participants' datasets were removed due to their failure to complete the experiment, leaving 28. Participants with Times 0em as their fastest settings completed the study with the same settings twice. The most common fastest typography settings were predominantly Roboto and Times and +0.10em spacing. Participants completed the arithmetic in their fastest and 'standard' typography settings with similar accuracy (~94%), $t(27) = 0.40$, $p = .348$. For the most part participants were faster to complete the true/false arithmetic statement evaluation presented in their speed-optimal settings (*M* speed-optimal = 474.91s, *M* standard = 519.37s), $t(27) = -1.82$, $p = .040$.Removing the data for the two participants whose fastest font was Times 0em (and thus who completed the study with the same font and spacing settings in both conditions), and the effect was still significant, $t(25) = -2.08$, $p = .024$.

Figure 1. Average time required to complete the evaluation of a block of 80 arithmetic expressions. Participants completed the task faster when material was presented in their speed-optimal font.

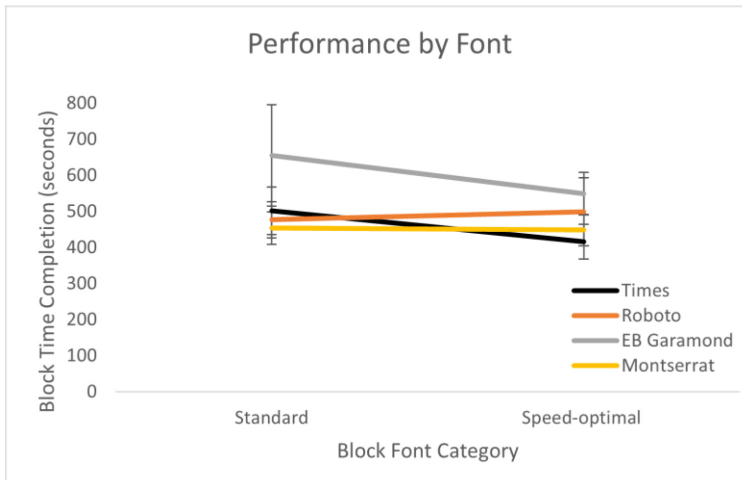
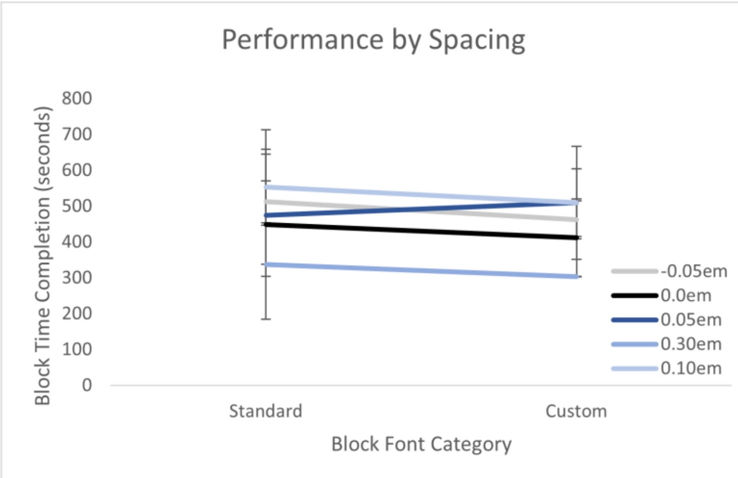


Figure 2. Average time required to complete the evaluation of a block of 80 arithmetic expressions based on spacing differences. Participants completed the task similarly when material was presented in their speed-optimal spacing versus the standard font spacing.



DISCUSSION

We found that the speed-optimal individuated typography setting for paragraph-length text as determined by the online readability assessment increased student performance in a basic arithmetic task, significantly reducing their total task completion time, confirming our prediction. While we found no difference in task accuracy, this would be expected given the ease of the correct/incorrect mathematical expression identification task. This is in line with previous work which found that reading performance may be improved in either speed or comprehension, but not both (Chaparro et. al, 2010; Wallace et. al, 2022).

Format readability optimization is a new and rapidly growing area of inquiry, and the present work recommends several refinements to the methods used both for this specific topic, and more generally. Future research should assess performance at a micro-level (i.e., per-question), rather than measuring completion time for a complete set of questions. Assessing overall time informs overall performance but does not account for individualized pacing or breaks. In addition, testing fonts more commonly used in mathematics (i.e., Latin Modern, Lucida, Cambria) could improve the ecological validity of future work. The current work employs fonts previously tested by The Readability Consortium (https://thereadabilityconsortium.org/) and available in the free research tools that group makes available, but future work should assess both more conventional mathematical font choices, and new technologies in mathematical fonts, such as variable typefaces which offer other dimensions of adjustment within the same font.

Future work would be well advised to look more granularly at time, fatigue, and mathematics strategy. It could be hypothesized that reduced readability increases perceived difficulty. As perceived difficulty of math problems increases because of reduced readability, participants may spend more time during a break or while solving a single problem, which can be disambiguated by timing individual responses. Furthermore, Increasing the complexity of the mathematical equations will likely lead to increased time spent on reading and processing. Complexity could be increased through differing basic arithmetic tasks (e.g., multiplication, division, subtraction) or by investigating basic arithmetic word problems. In addition, the current paradigm could be made more difficult by invoking exogenous time pressure, which would be typically seen in most real-world contexts (e.g., testing, or social pressure to calculate a solution quickly) (Wallace et. al, 2022).

Finally, future research should look beyond mathematics to other non-traditional reading domains where those fluent with conventional written words may yet struggle. If mathematics can benefit from improvements in format readability, such as speed-optimal settings, findings such as these will allow us to continue to explore readability optimization's role in the digitalization of the educational system.

This demonstration of the application of format readability, in the form of speed-optimal individuated typography into successfully transferring to a novel domain indicates that designers, researchers, instructors, and developers should consider investigating and implementing speed-optimal fonts and spacing in a variety of circumstances. As mathematics and other disciplines rely on technology-based education resources, we must investigate the nuances that will lead to the best success for students.

REFERENCES

Arditi, A., & Cho, J. (2005). Serifs and font legibility. *Vision research*, *45*(23), 2926-2933.

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current directions in psychological science*, *11*(5), 181-185.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic bulletin & review*, *14*, 243-248.

Ball, R. V., Miller, D. B., Wallace, S., Macias, K. C., Ibrahim, M., Gonzaga, E. R., ... & Sawyer, B. D. (2021, June). Optimizing Electronic Health Records Through Readability. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care. 10*(1), 65-70.

Beier, S., Berlow, S., Boucaud, E., Bylinskii, Z., Cai, T., Cohn, J., ... & Wolfe, B. (2022). Readability research: An interdisciplinary approach. *Foundations and Trends® in Human–Computer Interaction*, *16*(4), 214-324.

Cai, T., Wallace, S., Rezvanian, T., Dobres, J., Kerr, B., Berlow, S., ... & Bylinskii, Z. (2022). Personalized font recommendations: combining ML and typographic guidelines to optimize readability. *Designing Interactive Systems Conference*. 1-25.

Chaparro, B. S., Shaikh, A. D., Chaparro, A., & Merkle, E. C. (2010). Comparing the legibility of six ClearType typefaces to Verdana and Times New Roman. *Information Design Journal*, *18*(1), 36-49.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, *26*(1), 19-26.

Blanc-Goldhammer, D. R., & MacKenzie, K. J. (2018). The effects of natural scene statistics on text readability in additive displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 62*(1), 1281-1285).

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Hunt, T. E., & Sandhu, K. K. (2017). Endogenous and exogenous time pressure: Interactions with mathematics anxiety in explaining arithmetic performance. *International Journal of Educational Research*, *82*, 91-98.

Josephson, S. (2008). Keeping your readers' eyes on the screen: An eye-tracking study comparing sans serif and serif typefaces. *Visual communication*.

Kadner, F., Keller, Y., & Rothkopf, C. (2021). Adaptifont: Increasing individuals’ reading speed with a generative font model and bayesian optimization. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1-11.

Kember, P., & Varley, D. (1987). The legibility and readability of a visual display unit at threshold. *Ergonomics*, *30*(6), 925-931.

Mehler, B., & Reimer, B. (2020). Glanceable, legible typography over complex backgrounds. *Ergonomics*, *63*(7), 864-883.

Moret-Tatay, C., & Perea, M. (2011). Do serifs provide an advantage in the recognition of written words?. *Journal of Cognitive Psychology*, *23*(5), 619-624.

Patching, G. R., & Jordan, T. R. (2005). Assessing the role of different spatial frequencies in word perception by good and poor readers. *Memory & Cognition*, *33*(6), 961-971.

Paterson, K. B., & Jordan, T. R. (2010). Effects of increased letter spacing on word identification and eye guidance during reading. *Memory & Cognition*, *38*(4), 502-512.

Perera, K. (1980). The assessment of linguistic difficulty in reading material. *Educational review*, *32*(2), 151-161.

Sawyer, B. D., Dobres, J., Chahine, N., & Reimer, B. (2020). The great typography bake-off: comparing legibility at-a glance. *Ergonomics*, *63*(4), 391-398.

Sawyer, B. D., Wolfe, B., Dobres, J., Chahine, N., Mehler, B., & Reimer, B. (2020). Glanceable, Legible Typography Over Complex Backgrounds. *Ergonomics*.

Soleimani, H., & Mohammadi, E. (2012). The Effect of Text Typographical Features on Legibility, Comprehension, and Retrieval of EFL Learners. *English Language Teaching*, 5(8), 207-216.

Wallace, S., Bylinskii, Z., Dobres, J., Kerr, B., Berlow, S., Treitman, R., ... & Sawyer, B. D. (2022). Towards individuated reading experiences: different fonts increase reading speed for different individuals. *ACM Transactions on Computer-Human Interaction*, *29*(4), 1-56.

Wallace, S., Dobres, J., & Sawyer, B. D. (2021). Considering the speed and comprehension trade-off in reading mediated by typography. *Journal of Vision*, *21*(9), 2249-2249.

Wallace, S., Dobres, J., Bylinskii, Z., & Sawyer, B. (2022). Space for readability: effects on reading speed from individuated character and word spacing. *Journal of Vision*, 22(14), 3349-3349.