The Effects of Font Disfluency on Reading Retention

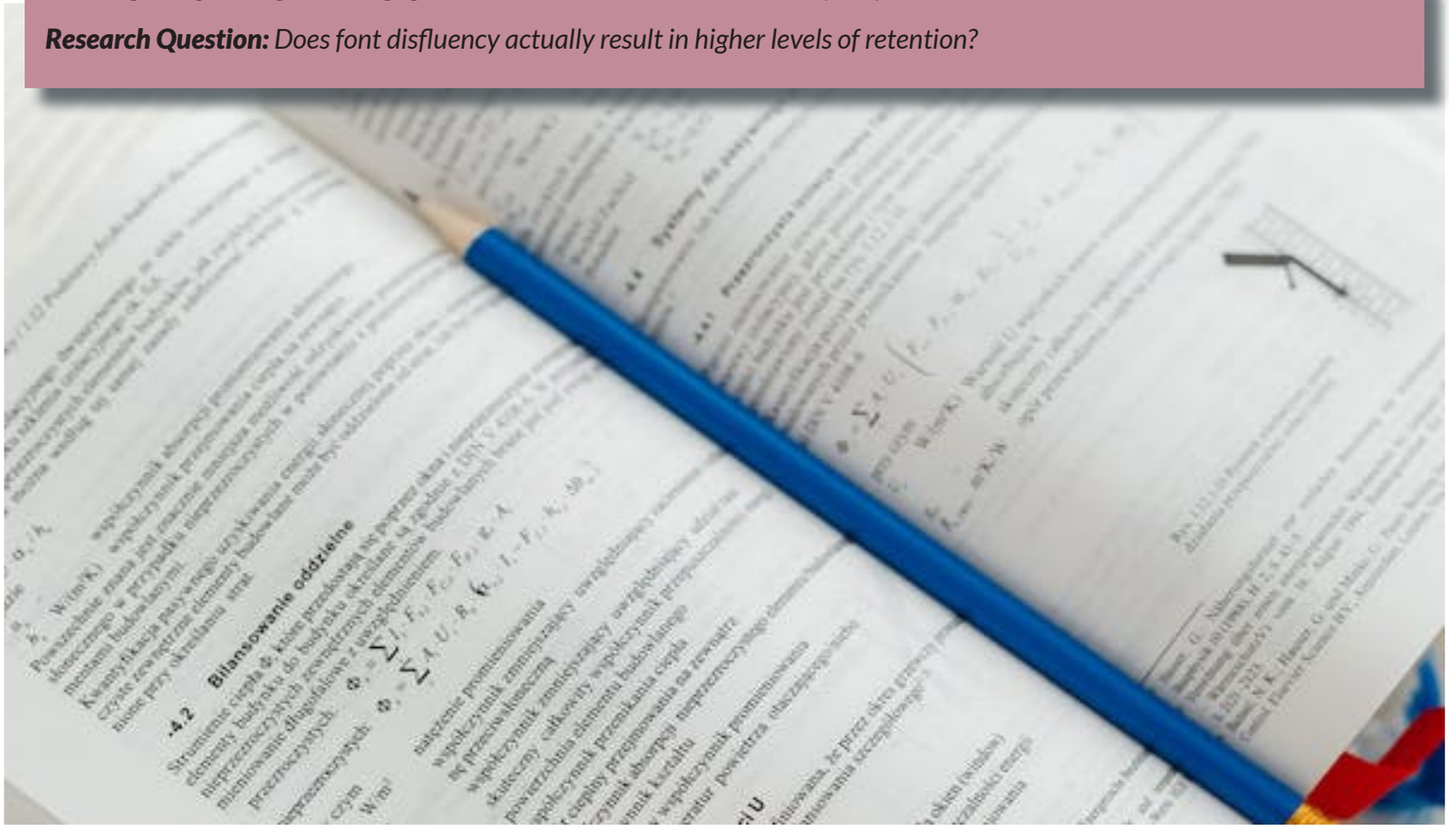
Marissa Bailey, Lauren Duty, Austen Greenway *December 7th, 2021*

Abstract

Some earlier research into the relationship between reading ease and reading retention suggests that higher reading difficulty promotes higher cognitive engagement, which increases how much readers retain (Bjork, 1994). If the difficulty level is too high, it frustrates readers and decreases engagement, but if the difficulty is just high enough or "desirable," then reading retention will improve (Bjork, 1994). Some researchers believe that this theory can be applied to font choice. Hard-to-read fonts may create a desirable difficulty and increase how much a reader retains. This theory is known as the font disfluency effect. If valid, the font disfluency effect could impact a wide range of fields, including education, marketing, and design. However, while several studies have shown font disfluency to be effective (Bjork, 1994; Oppenheimer et al., 2010; Sungkhasettee et al., 2011), several other studies have shown it to be ineffective (Eitel & Kühl, 2016; Rummer et al., 2016; Taylor et al., 2020). In an attempt to learn more about the effects of font disfluency on reading retention, we conducted a study involving 64 participants. We administered a timed reading test in four different fonts styles to evaluate font disfluency and rank reading difficulty. We then gave participants differing versions of a multiple-choice reading retention test to compare participant scores to font styles and difficulty rankings. Lastly, we administered a post-test interview to assess participant perceptions of font and performance. Our results may indicate that there is a correlation between the legibility of a font style and how much content readers retain; however, the usefulness of font disfluency still remains in question.

Research Premise: *The* ***font disfluency effect*** *theorizes that harder to read fonts provide a level of "desirable difficulty" resulting in higher cognitive engagement. Therefore, disfluent fonts may help students retain more information.*

Research Question: *Does font disfluency actually result in higher levels of retention?*



Introduction

Important Terms and Concepts

Desirable difficulty *- the level at which readers work just hard enough to engage with text at a higher cognitive level, but not so much that they become frustrated*

Font disfluency *- suggests that that hard-to-read fonts can create desirable difficulty*

Reading retention *- how much information readers retain from a text*

Textual communication has increasingly become an integral part of our society. We most commonly see it used for marketing; corporations and interest groups use it to reach, inform, or persuade a target audience to engage with a service or buy a product. For text to fulfill these purposes, the information it communicates must be retained. One important variable in creating more memorable text may be font style. A 2015 study, "The Taste of Typeface," explored some of the ways in which people associate taste with different shapes and fonts (Velasco et al., 2015). The researchers found that people tend to associate fonts with certain ideas, emotions, and experiences. Choice of typeface can also impact cognitive engagement. One 2020 study tested how handwritten text and typed text promote cognitive engagement (Izadi & Patrick, 2020). The study concluded that fonts which mimic handwriting elicit the action of approach and therefore haptic engagement (Izadi & Patrick, 2020). Building on these critical studies and others, our study seeks to understand whether font disfluency has any effect on how much content readers remember.

"The researchers found that people tend to associate fonts with certain ideas, emotions, and experiences."

What is font disfluency?

The font disfluency effect is based on an earlier theory called the disfluency effect. This theory posits that hard-to-read text promotes higher cognitive engagement, therefore increasing content retention at a "desirable difficulty" (Bjork, 1994). The desirable difficulty is the level at which readers work just hard enough to engage with text at a higher cognitive level, but not so difficult that readers become frustrated and lose retention. The font disfluency effect builds on this research by suggesting that that hard-to-read fonts can create desirable difficulty. If this is true, the application of font disfluency theory could have wide-ranging implications for several fields, including education, marketing, and design.

In two related studies on disfluency, researchers found that harder-to-read fonts increased retention rates, leading them to conclude that perceptual disfluency can successfully function as a desirable difficulty (Oppenheimer et al., 2010). Further research on the theory of desirable difficulty has shown the potential benefits of applying font disfluency. A small 2011 study tested font disfluency with 20 undergraduate students from the university of California (Sungkhasettee et al., 2011). Researchers asked each participant to study lists of words. These lists were presented in two different formats: upright and inverted. Researchers found that recall performance was better for inverted words across all lists (Sungkhasettee et al., 2011).

"...harder-to-read fonts increased retention rates, leading them to conclude that perceptual disfluency can successfully function as a desirable difficulty"

Although some studies have shown promising results for the use of font disfluency, there is still doubt surrounding the validity of this theory. Some researchers argue that there is a difference between disfluent difficulty and desirable difficulty. A recent 2020 study on the relationships between fonts and memory noted that, "Of course, not all difficulties are desirable, and desirable difficulties are notoriously fickle" (Taylor et al., 2020). Several other research studies agree that applying desirable difficulties is not generally effective. One 2016 study hypothesized that disfluent text paired with high test expectancy would prompt more mental effort, resulting in increased retention and better test scores (Eitel & Kühl, 2016). However, the researchers found that

disfluency was not effective and could even be a drawback under those experimental conditions (Eitel & Kühl, 2016).

In fact, several researchers have found flaws in studies that support disfluency. One flaw is that many disfluency-supporting studies tested their participants using word lists rather than paragraphs, which does not mimic real-world contexts. Additionally, it has been noted that the test content in certain studies was not only disfluent but also unusual. The test included words, phrases, or concepts which were so unusual that their peculiarity may have made them more memorable. In 2016, a study was conducted in response to this flaw and, using multiple fluent and disfluent word lists, produced opposing results (Rummer et al., 2016). Researchers found that the use of disfluent text in educational settings did not produce learning advantages (Rummer et al., 2016).

> ***"...many disfluency-supporting studies tested their participants using word lists rather than paragraphs, which does not mimic real-world contexts."***

The conflicting results of many of the previously mentioned studies makes the effectiveness of font disfluency unclear. To make matters even more unclear, some businesses and institutions have begun using font disfluency in their marketing campaigns with varying results. For example, the Royal Melbourne Institute of Technology (RMIT) created a font they call Sans Forgetica that, based on an unpublished study they did in 2018, supposedly increases reading retention (RMIT, 2018). The Sans Forgetica font was supposedly created to reach the ideal desirable difficulty in reading. While this study may have some validity, it seems to have been part of a targeted publicity campaign.

This is Sans Forgetica

In our study, we aim to determine if there is a reliable relationship between font legibility and reading retention. To produce valid results, we have drawn methods and best practices from past studies to eliminate as many flaws as possible. Among our primary considerations, we found that font selection and content may have imposed flaws on previous studies. We chose to address these concerns by using a structured font selection process, using paragraphs instead of word lists, controlling for reading level, and focusing on less memorable sentence components. We believe that these considerations have allowed us to address many of the flaws in earlier studies.

Methods

Important Terms and Concepts

Serif font *- font style with small extensions or extra strokes protruding from the ends of letters (these protrusions are called serifs)*

Sans-serif font *- font style without serifs*

Script font *- font style designed to mimic handwriting*

Display font *- font style that is designed to be used at large sizes for display; usually eccentric, eye-catching, and decorative*

In preparation for our study, we set up a structured process for selecting our test fonts. We chose four fonts, one to represent each of the four main typeface styles: serif, sans serif, script, and display. Three of the fonts we used were chosen from the population of fonts on Google Fonts. We chose these fonts as ideal representations because they possessed the highest frequency of the characteristics of their style. Old Standard TT was chosen as our serif font because it had the highest number of

This is Old Standard TT

This is Zen Maru Gothic

This is Cherish

This is Sans Forgetica

serifs per letter. Zen Maru Gothic, our sans serif font, had the fewest number of letter extensions. Cherish, our script font, had the highest frequency and longest length of flourishes per letter. For our fourth font, a display font, we chose Sans Foregtica. We were intrigued by the Sans Forgetica font study and wanted to test the validity of RMIT's claim.

We also focused on addressing flaws from earlier studies. In previous studies, researchers used word pairs or highly unusual words when designing their reading retention tests. These tests do not mimic natural conditions. Instead, we chose to use a paragraph to mimic the style of reading that participants would normally engage in. We considered if perceptual difficulty from the font may add to content difficulty, thereby increasing the overall difficulty above desirability. To overcome this, we created a paragraph geared towards a low reading level (it rates between third and fifth grade depending on the readability scale used) to control for the difficulty of the content. Because we chose to use a paragraph, we also considered the difference between content retention and concept retention. Readers are more likely to remember concepts than specific content. Therefore, we chose to base our reading comprehension questions on adjectives and adverbs to test for content retention as opposed to using nouns and verbs, which may have only been effective for testing concept retention.

> ***"In previous studies, researchers used word pairs or highly unusual words when designing their reading retention tests"***

Our research team utilized convenience sampling for recruitment due to the time and locational constraints of our study. The main method we used with our test participants was quantitative experimental research, but we also conducted post-test interviews. Our methods are described as follows:

1 We recruited participants in the Atrium building on campus. We tested during four different sessions which took place across varying times but were mostly conducted during midday. Once recruited, our participants were seated in a controlled environment (a quiet, well-lit room) and given a consent cover letter that discussed the following sections:

- Title of Research Study
- Researcher's Contact Information
- Description of Project
- Explanation of Procedures
- Risks or Discomforts
- Benefits
- Compensation
- Confidentiality.

2 Upon agreeing to the terms of the cover letter, participants began the timed reading section of our test. Participants were informed that their reading would be timed and then were sequentially given four different printed paragraphs, each containing 60 words. Our researchers used a phone app to record their reading speeds.

Each paragraph was printed in a different style from one of the following four font styles: Old Standard TT (serif font), Zen Maru Gothic (sans serif font), Sans Forgetica (display font), and Cherish (script font). Each paragraph was different; they gave a description and history of the font they were written in. This controlled for content familiarity. The order of the font styles received by the participants was varied in aggregates of eight to ensure that the reading order of the fonts did not affect the participants' performance. Participants read each paragraph separately and consecutively until their individual reading times had been documented for each font style. This process gave us each individual's baseline to compare later results with.

3 After completing the timed reading section, participants immediately began the reading retention section of the test. Participants were informed that they would receive a sheet of paper containing a paragraph printed in one of the four font styles from the previous section. The order of the font styles received by the participants in the second section was operated by a schedule to ensure that our team gathered equivalent data points on each of the four font styles. Researchers explained to participants that this portion of the test was untimed and encouraged participants to take their time in reading.

4 When participants had finished the second section, researchers removed the paragraph gave participants a multiple-choice questionnaire. The questionnaire contained five questions designed to test participants' ability to recall certain adjectives and adverbs from the paragraph they just finished reading. At the end of the multiple-choice section, the participants were given a demographic questionnaire so that the research team might recognize existing patterns in the data found based on their personal information such as age, self-identification, ethnicity, and education level.

5 After participants completed the reading test, we conducted our second research method: a post-test interview. Our researchers followed an interview schedule containing five questions regarding the test that the participants had just completed. These questions were designed to measure participants' perceptions of the fonts. The interview questions were asked in a funnel sequence, beginning with broad questions first, followed by specific, closed-ended questions. Our researchers also noted any additional comments participants made about their feelings towards specific fonts. The nature of the interview schedule allowed our researchers to ask follow-up questions and gain greater insights into our findings.

Summary of Test Design

- Chose four fonts; one representing each major style

- Selected fonts from Google Fonts based on having the highest frequency of ideal characteristics

- Used paragraphs with common words instead of word lists with peculiar or uncommon words

- Controlled for content difficulty by using a low reading level (between 3rd and 5th grade)

- Considered how difficult a font style was for each individual

- Used individual rankings to determine difficulty, not font style alone

- Focused retention test on adjectives and adverbs

- Administered tests based on a test schedule to ensure that each font style had equal tests

6 After completing our quantitative experimental research and post-test interviews for our 64 participants, our team gathered and organized the data to review trends and assess our findings.

Summary of Research Methods

- Recruited participants in the Atrium building on the Marietta Campus across four testing sessions

- Brought participants to a private, quiet, well-lit room on the second floor of the Atrium

- Gave participants a consent cover letter to read and confirmed consent

- Conducted four timed reading tests with participants; one on each of the four different font styles

- Removed timed reading test from participant

- Presented participants with a fifth paragraph in one of four font styles determined by a test schedule so each font received an equal number of tests

- Informed participants that this part of the test was not timed and encouraged them to read at their own pace

- Removed paragraph when the participant indicated they were finished

- Presented participants with a 5-question multiple-choice reading retention quiz

- Collected the quiz when participants indicated they were finished

- Interviewed participants for demographic data as well as feelings and attitudes towards font styles and tests

Results and Discussion

The initial results of our study do seem to indicate that there may be a link between the reading difficulty of a font style and the reader's retention of content, as shown in the chart below.

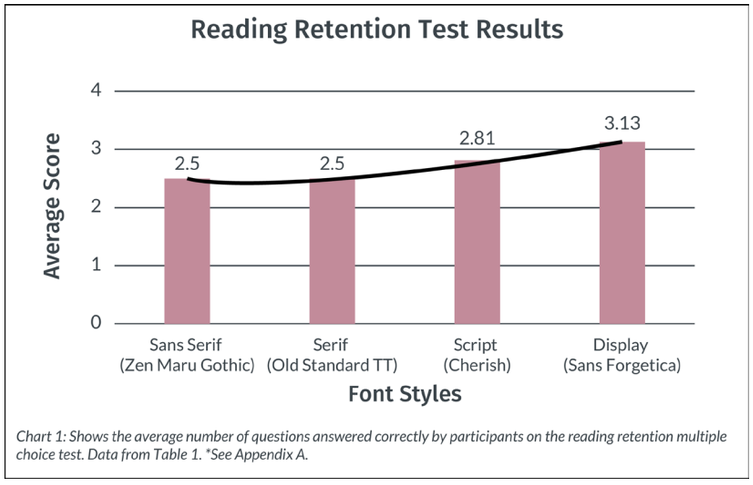


*Chart 1: Shows the average number of questions answered correctly by participants on the reading retention multiple choice test. Data from Table 1. *See Appendix A.*

The results in the table above also seem to indicate that Sans Forgetica did have the highest retention rate among the chosen fonts, which is consistent with the study conducted at RMIT that led to the creation of Sans Forgetica. These findings are in opposition with other studies conducted that specifically involved Sans Forgetica and font disfluency (Geller & Peterson, 2021; Taylor et al., 2020), but this may be due to differences between those studies and our study.

Key Differences

- The other studies used word pairs (Taylor et al., 2020) and word lists (Geller & Peterson, 2021) while our study used full paragraphs.

- Both other studies were conducted online while our study was conducted in a physical testing space with a controlled environment.

- Both other studies only compared Sans Forgetica with one other font style while our study was conducted using four font styles of varying disfluency levels.

During the timed reading test, we varied the reading order. It seemed to have no impact on participant reading times. Our findings do indicate that font difficulty levels varied based on the participant, therefore the fonts were shown to be more or less difficult for different individuals, as shown in the chart below.

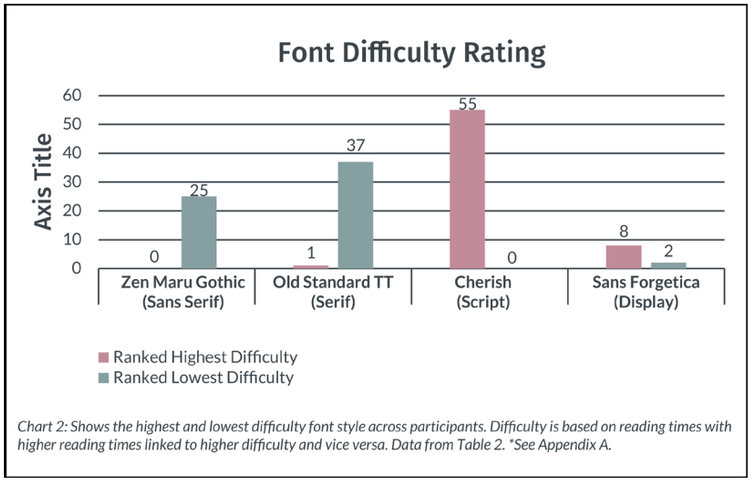


*Chart 2: Shows the highest and lowest difficulty font style across participants. Difficulty is based on reading times with higher reading times linked to higher difficulty and vice versa. Data from Table 2. *See Appendix A.*

For most participants, the highest difficulty font was Cherish (script) and the lowest difficulty font was Old Standard TT (serif). Sans Forgetica (display) was the second highest difficulty font and Zen Maru Gothic (sans serif) was the second-lowest difficulty font. These ratings were not consistent for all participants, and due to these inconsistencies, the font difficulty ratings varied for each participant. Previous studies conducted on font disfluency have not accounted for individual differences in font difficulty rankings, which may possibly have confounded results.

> ***"Previous studies conducted on font disfluency have not accounted for individual differences in font difficulty rankings"***

To account for varying difficulty rankings among participants, we also analyzed our results by correlating the individual participant's font difficulty ranking with the version of the test they received so that we could review the scores based on how the test version matched up with each participant's personal difficulty ratings. The results

are shown in the chart below.

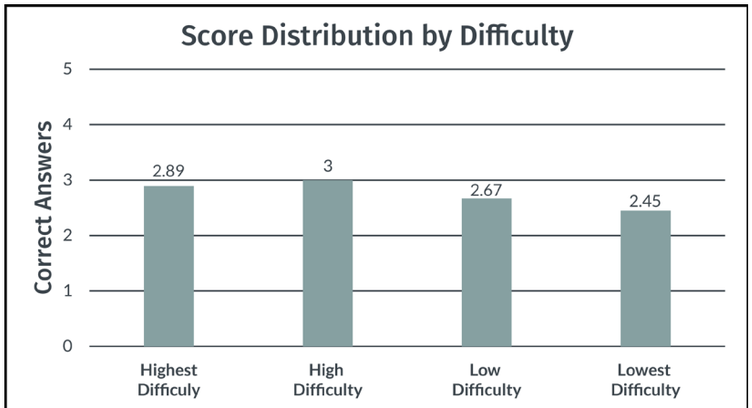


*Chart 3: Shows the average number of questions answered correctly by participants on the reading retention multiple choice test distributed based on font difficulty. Data from Table 3. *See Appendix A.*

When disregarding the font style itself and considering font difficulty among individual participants, the results show that the participants typically scored higher when given the version of the test in the font that was rated as the second most difficult for them. Likewise, participants typically scored lower when given the version of the test in the font that was ranked lowest difficulty for them.

"...participants typically scored higher when given the version of the test in the font that was rated as the second most difficult for them."

Several other data points were also collected to either validate or invalidate the findings of our timed reading and reading comprehension tests. During the interview portion of our test, we measured the participant's perception of which font was the most difficult, the participant's perception of which font was most enjoyable to read, and the participant's perception of how much information they retained based on the font style. These factors did not correlate to their test results, which may imply that the participant's perception of the font did not impact their retention. Many participants who perceived Cherish as having the highest difficulty rating had higher reading times on Sans Forgetica and vice versa. Participants who received the version of the test in the font they perceived as most enjoyable to read did not do any better on average than other participants who received the same test version.

We also collected data on age, gender, and education. Education and age both showed a correlation with test scores, which indicates that older participants and those with higher education levels tended to get higher test scores; however, these results should be interpreted with caution. One reason to interpret our findings with caution is that higher education is inherently linked to higher age due to the time it takes participants to obtain higher education. On average, participants who had the highest levels of education fell into higher age ranges. Furthermore, many participants self-identified that their highest level of education obtained was "high school graduate" despite the fact that they were college students and "some college" was the most accurate response. Some participants who were graduating seniors struggled to choose between "some college" and "4-year degree" because, although they had intellectually obtained a "4-year degree" education level, they had not received their diploma. Therefore, it could be that students who had obtained "some college" self-identified at a lower education level because they were not academically confident. Our findings for education and age should also be considered in the context of our study's skewed sampling bias, which is discussed in further detail later in this section.

While our results would seem to indicate a correlation between the reading difficulty of a font style and the reader's retention of content, there were several issues with our study. The largest potential issue with our study is that participants knew what they were being tested for. The information given about the study during the consent process may have impacted their expectations for the subsequent reading tests and therefore primed them to respond to the tests in certain ways. Many of the participants made comments during their interviews to explain why they thought they did or did not remember more, and some of those comments were oddly similar to the phrasing in the consent letter. This could be coincidence, but the possibility exists that this information may have influenced how participants responded to the test. Due to the limitations of our research context, we were not able to use deception while conducting our research, but future researchers may consider using some deception during the testing process to avoid priming participants.

"Many of the participants made comments during their interviews... and some of those comments were oddly similar to the phrasing in the consent letter. This could be coincidence, but the possibility exists that this information may have influenced how participants responded to the test."

Another potential issue with our study is that majority of our participants were white male college students, and because most of our participants were chosen from a similar geographic location, many of them shared the same areas of study: computer science or engineering. Both of these sample characteristics may have skewed our results. For example, we noticed that three themes continued to appear during our interviews, especially among the majority demographic of our sample.

Reoccuring Themes

- Participants self-identified as poor readers or did not do much reading outside of school-related settings.

- Participants thought Sans Forgetica was "cool" or "interesting" because it reminded them of certain font styles from video games.

- Participants did not like the script font Cherish because they did not know how to read cursive.

These interview responses indicate that our study may have achieved different results if our sample group had contained equal representations for age, gender, and education as opposed to our heavily skewed white male STEM-major college student sample. Sample groups who identify as avid readers, spend more time reading, don't play video games, or know how to read cursive may have responded to the font styles differently. While some of our interview data seems to indicate that the participant's perception of the font did not influence their performance on the test, further research with other sample groups is recommended to confirm if our results are valid.

"Sample groups who identify as avid readers, spend more time reading, don't play video games, or know how to read cursive may have responded to the font styles differently."

Regardless of whether our results are shown to have external validation across other populations, or at least external validation for other populations of college students, it is still questionable as to whether or not it would benefit readers to engage with texts that intentionally promote font disfluency. During our interviews, many participants discussed the two highest difficulty fonts, Cherish and Sans Forgetica, with mostly very negative descriptions.

Descriptions of Fonts

- "Felt like every hole in a letter was a hole in my brain" (Sans Forgetica)

- "Looked like a bunch of shapes" (Sans Forgetica)

- "Weird. Made my brain want to fill in the letter so I got stuck on every sentence." (Sans Forgetica)

- "Too fancy. My brain kind of filled in as I was reading but it was still too difficult" (Sans Forgetica)

- Almost illegible" (Cherish)

- "Miserable" (Cherish)

- "Didn't like it" (Sans Forgetica)

- "I don't really do cursive" (Cherish)

- "[Looked like] fake cursive" (Cherish)

- "No word space" (Cherish)

- "My brain rebelled against me" (Sans Forgetica & Cherish)

Our participants were tasked with reading very short samples of these fonts, just one 60-word paragraph in each style and an additional 72-word paragraph in one of the font styles. The paragraphs were very short and had an average reading level between third and fifth grade, so participants did not read for long and did not struggle to read or understand the material itself, only the font style. In a real-world setting, such as a textbook in which readers may read for hours and may struggle to memorize new vocabulary terms or comprehend new concepts, would it be beneficial to their learning to increase the difficulty of the text? Would increasing the text difficulty level make the material so difficult for readers to understand that their "brain would rebel" as one participant put it? Further research with longer and more difficult material may give better insight into the potential applications of our findings and confirm whether readers would actually benefit from reading materials in disfluent fonts or if comprehension would decline due to frustration.

"Would increasing the text difficulty level make the material so difficult for readers to understand that their "brain would rebel" as one participant put it? Further research with longer and more difficult material may give better insight into the potential applications of our findings."

Appendix A - List of Tables

List of Tables

Table 1: Participant Score Distribution by Test Version

Test# SANS (Zen Maru Gothic)	Score	Test # SERF (Old Standard TT)	Score	Test # SCRP (Cherish)	Score	Test # FORG (Sans Forgetica)	Score
SANS-03	1	SERF-05	0	SCRP-04	1	FORG-02	2
SANS-05	1	SERF-11	0	SCRP-06	1	FORG-03	2
SANS-16	1	SERF-07	1	SCRP-05	2	FORG-10	2
SANS-07	2	SERF-04	2	SCRP-07	2	FORG-12	2
SANS-10	2	SERF-06	2	SCRP-11	2	FORG-15	2
SANS-14	2	SERF-08	2	SCRP-14	2	FORG-01	3
SANS-15	2	SERF-14	2	SCRP-16	2	FORG-06	3
SANS-01	3	SERF-15	2	SCRP-01	3	FORG-09	3
SANS-04	3	SERF-01	3	SCRP-02	3	FORG-11	3
SANS-06	3	SERF-02	3	SCRP-10	3	FORG-13	3
SANS-08	3	SERF-09	3	SCRP-12	3	FORG-04	4
SANS-09	3	SERF-10	3	SCRP-03	4	FORG-05	4
SANS-11	3	SERF-03	4	SCRP-08	4	FORG-08	4
SANS-12	3	SERF-12	4	SCRP-09	4	FORG-14	4
SANS-02	4	SERF-16	4	SCRP-13	4	FORG-16	4
SANS-13	4	SERF-13	5	SCRP-15	5	FORG-07	5
Average Score	2.50	Average Score	2.50	Average Score	2.81	Average Score	3.13

Table 1: SANS tests were administered in the sans serif font (Zen Maru Gothic), SERF tests were administered in the serif font (Old Standard TT), SCRP tests were administered in the script font (Cherish), and FORG tests were administered in the display font (Sans Forgetica). Columns are ordered by average score from lowest to highest.

Table 2: Font Difficulty Rank by Individual Participant Reading Times

	Zen Maru Gothic (Sans Serif)	Old Standard TT (Serif)	Cherish (Script)	Sans Forgetica (Display)
Ranked Highest Difficulty	0	1	55	8
Ranked Lowest Difficulty	25	37	0	2

Table 2: Numbers in each section represent the number of participants whose scores reflected the difficulty rating in column one for the font in row one.

Table 3: Participant Score Distribution by Ranked Difficulty

Difficulty: Highest	Score	Difficulty: High	Score	Difficulty: Low	Score	Difficulty: Lowest	Score
FORG-07	5	FORG-04	4	SANS-02	4	SERF-13	5
SCRP-15	5	FORG-14	4	SANS-13	4	SERF-12	4
FORG-05	4	FORG-16	4	SERF-03	4	SERF-16	4
FORG-08	4	SCRP-03	4	SANS-01	3	SANS-04	3
SCRP-09	4	SCRP-08	4	SERF-02	3	SANS-06	3
SCRP-13	4	FORG-01	3	SERF-10	3	SANS-08	3
SCRP-01	3	FORG-06	3	SANS-07	2	SANS-09	3
SCRP-02	3	FORG-09	3	SANS-14	2	SANS-11	3
SCRP-10	3	FORG-11	3	SERF-04	2	SANS-12	3
SCRP-12	3	FORG-13	3	SERF-14	2	SERF-01	3
FORG-15	2	FORG-02	2	SERF-15	2	SERF-09	3
SCRP-05	2	FORG-03	2	SANS-16	1	FORG-12	2
SCRP-07	2	FORG-10	2			SANS-10	2
SCRP-11	2	SANS-03	1			SANS-15	2
SCRP-14	2					SERF-06	2
SCRP-16	2					SERF-08	2
SCRP-04	1					SANS-05	1
SCRP-06	1					SERF-07	1
						SERF-05	0
						SERF-11	0
Average Score	2.89	Average Score	3.00	Average Score	2.67	Average Score	2.45

Table 3: There are uneven numbers of participants in each category because tests were distributed based on font, not difficulty.

Appendix B - Test Materials

Section 1: Timed Reading in Each Font Style

Font A: Zen Maru Gothic (Sans Serif)

This font is called Zen Maru Gothic. It's a sans serif style font. Sans serif means "without serifs." Serifs are the small decorations on the ends of certain letters in serif font styles. One of the most popular sans serif fonts is Helvetica. You may be more familiar with Calibri, which is currently the default font used by Microsoft Word.

Font B: Old Standard TT (Serif)

This font is called Old Standard TT. It's a serif font. Serifs are small decorative marks added to the ends of some letters. You can see this very well on the ends of the "T" or "S." The most popular serif font is Times New Roman, which you may be familiar with if you've ever written a paper for school.

Font C: Cherish (Script)

This font is called Cherish. It's a decorative font known as a script font. Script fonts were created to imitate handwriting, as you may have guessed. A font called Blackletter is often credited with being the first typeface ever created. Blackletter is considered to be a script font, so it may have also been the first script font ever created.

Font D: Sans Forgetica (Display)

This font is called Sans Forgetica. It's a display font. Display is a large category that includes many different types of non-standard fonts. Designers may use it for titles, headings, logos, brand names, or other decoration, but it is rarely used for body text. Display fonts may have features such as rounded serifs, stencil shapes, irregular strokes, icons, and more.

Section 2: Reading Retention in One Font Style (Random)

Test Version A

Once, there was a hungry fox. It was early morning when the fox snuck into the hen house. The hens were all sound asleep, but the rooster was not. The rooster crowed loudly and lunged at the fox. Terrified, the fox leapt from the hen house and ran straight into the farmer's old, fat dog. The quick brown fox jumped over the lazy dog and escaped. The hungry fox would have to try again another day.

Test Version B

Once, there was a hungry fox. It was early morning when the fox snuck into the hen house. The hens were all sound asleep, but the rooster was not. The rooster crowed loudly and lunged at the fox. Terrified, the fox leapt from the hen house and ran straight into the farmer's old, fat dog. The quick brown fox jumped over the lazy dog and escaped. The hungry fox would have to try again another day.

Test Version C

Once, there was a hungry fox. It was early morning when the fox snuck into the hen house. The hens were all sound asleep, but the rooster was not. The rooster crowed loudly and lunged at the fox. Terrified, the fox leapt from the hen house and ran straight into the farmer's old, fat dog. The quick brown fox jumped over the lazy dog and escaped. The hungry fox would have to try again another day.

Test Version D

Once, there was a hungry fox. It was early morning when the fox snuck into the hen house. The hens were all sound asleep, but the rooster was not. The rooster crowed loudly and lunged at the fox. Terrified, the fox leapt from the hen house and ran straight into the farmer's old, fat dog. The quick brown fox jumped over the lazy dog and escaped. The hungry fox would have to try again another day.

Reading Retention Test Questions (correct answers in red)

1. Below are four similar sentences. One of these sentences appeared in the paragraph you just read. Please choose the sentence that appeared in the paragraph you just read.

a) It was early evening when the fox slipped into the hen house.

b) It was late morning when the fox crept into the hen house.

c) It was early morning when the fox snuck into the hen house.

d) It was early morning when the fox slipped into the hen house.

2. Below are four similar sentences. One of these sentences appeared in the paragraph you just read. Please choose the sentence that appeared in the paragraph you just read.

a) The hens slept silently, but the rooster did not.

b) The hens all slept soundly, but the rooster was awake.

c) The hens were sound asleep, but the rooster was awake.

d) The hens were all sound asleep, but the rooster was not.

3. Below are four similar sentences. One of these sentences appeared in the paragraph you just read. Please choose the sentence that appeared in the paragraph you just read.

a) The rooster squawked and lunged wildly at the fox.

b) The rooster crowed loudly and lunged at the fox.

c) The rooster crowed wildly and lunged at the fox.

d) The rooster squawked loudly and lunged at the fox.

4. Below are four similar sentences. One of these sentences appeared in the paragraph you just read. Please choose the sentence that appeared in the paragraph you just read.

a) Terrified, the fox leapt from the hen house and ran straight into the farmer's old, fat dog.

b) Frightened, the fox leapt from the hen house and landed right in front of the farmer's old, fat dog.

c) Terrified, the fox leapt from the hen house and landed right in front of the farmer's old, fat dog.

d) Frightened, the fox leapt from the hen house and ran straight into the farmer's old, fat dog.

5. Below are four similar sentences. One of these sentences appeared in the paragraph you just read. Please choose the sentence that appeared in the paragraph you just read.

a) The swift brown fox jumped over the lazy dog and got away.

b) The quick brown fox jumped over the lazy dog and got away.

c) The swift brown fox jumped over the lazy dog and escaped.

d) The quick brown fox jumped over the lazy dog and escaped.

Demographics Questions

1. What is your age?

o 18 – 24

o 25 – 34

o 35 – 44

o 45 – 54

o 55 – 64

o 65 – 74

o 75 – 84

o 85+

2. How do you self-identify?

o Male

o Female

o Non-binary/third gender

o Prefer not to say

3. What is your ethnicity? (Mark all that apply)

o White

o Black or African American

o American Indian or Alaska Native

o Asian

o Native Hawaiian or Pacific Islander

o Other

4. What is your highest level of education attained?

o Less than high school

o GED

o High school graduate

o Some college

o Certificate

o 2-year degree

o 4-year degree

o Professional degree

o Masters

o Doctorate

Interview Questions

Q1: How well do you think you did on section two?

A: _____

Q2: Were you surprised by how much you did or did not remember from the paragraph?

A: _____

Q3: Which font did you find the hardest to read?

A: _____

Q4: Which font did enjoy reading the most?

A: _____

Q5: Think back to the first four paragraphs you read. Do you feel like you remember more details from one of those paragraphs better than the others?

A: _____

References

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). Cambridge, MA: MIT Press.

Eitel, A., & Kühl, T. (2016). Effects of disfluency and test expectancy on learning with text. Metacognition & Learning, 11(1), 107–121. **https://doi.org/10.1007/s11409-015-9145-3**

Geller, J., & Peterson, D. (2021). Is this going to be on the test? Test expectancy moderates the disfluency effect with sans forgetica. Journal of Experimental Psychology: Learning, Memory, and Cognition. **https://doi.org/10.1037/xlm0001042**.supp (Supplemental)

Izadi, A., & Patrick, V. M. (2020). The power of the pen: Handwritten fonts promote haptic engagement. Psychology & Marketing, 37(8), 1082–1100. **https://doi.org/10.1002/mar.21318**

Oppenheimer, D., Diemand-Yauman, C., & Vaughan, E. (2010). Fortune Favors the Bold (and the Italicized): Effects of Disfluency on Educational Outcomes. Proceedings of the Annual Meeting of the Cognitive Science Society, 32. **https://escholarship.org/uc/item/4wd1s7hj**

Royal Melbourne Institute of Technology. (2018). Sans Forgetica. Sans Forgetica-RMIT. **https://www.rmit.edu.au/news/all-news/2018/oct/sans-forgetica-news-story**

Rummer, R., Schweppe, J., & Schwede, A. (2016). Fortune is fickle: null-effects of disfluency on learning outcomes. Metacognition & Learning, 11(1), 57–70. **https://doi.org/10.1007/s11409-015-9151-5**

Sungkhasettee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: illusions of competency and desirable difficulties. Psychonomic Bulletin & Review, 18(5), 973–978. **https://doi.org/10.3758/s13423-011-0114-9**

Taylor, A., Sanson, M., Burnell, R., Wade, K. A., & Garry, M. (2020). Disfluent difficulties are not desirable difficulties: the (lack of) effect of Sans Forgetica on memory. Memory (Hove, England), 28(7), 850–857. **https://doi.org/10.1080/09658211.2020.1758726**

Velasco, C., Woods, A. T., Hyndman, S., & Spence, C. (2015). The Taste of Typeface. I-Perception, 6(4), 1–10. **https://doi.org/10.1177/2041669515593040**