

# Sans Forgetica is Really Forgettable

Jason Geller<sup>1</sup>, Sara D. Davis<sup>2</sup>, & Daniel Peterson<sup>2</sup>

<sup>1</sup> University of Iowa

<sup>2</sup> Skidmore College

## Abstract

Do students learn better with material that is perceptually harder-to-process? While evidence is equivocal on the matter, recent claims suggest that placing materials in Sans Forgetica font, which is perceptually hard-to-process, has positive effects on student learning. Given the weak evidence for perceptual disfluency effects, this led us to examine the mnemonic effects of Sans Forgetica more closely. In two preregistered experiments, we tested if Sans Forgetica is really unforgettable. In Experiment 1 ( $N = 233$ ), participants studied weakly related cue-target pairs with targets presented in either Sans Forgetica or with missing letters (e.g., G\_\_RL). Cued recall performance showed a robust generation effect, but no Sans Forgetica memory benefit. In Experiment 2 ( $N=528$ ), participants read a passage about ground water with select sentences presented in either Sans Forgetica, yellow highlighting, or unchanged. Cued recall for select words were better for pre-highlighted information than when unmodified. Critically, presenting sentences in Sans Forgetica did not produce better cued recall than pre-highlighted sentences or sentences presented unchanged. Our findings suggest that Sans Forgetica is really forgettable.

*Keywords:* Disfluency, Recall, Desirable Difficulty, Learning and Memory

Word count: X

Students want to remember more and forget less. Decades of research have put forth the paradoxical idea that making learning harder (not easier) should have the desirable effect of improving long-term retention of material—called the desirable difficulty principle (Bjork, 1994). Notable examples of desirable difficulties include having participants generate

---

Jason Geller, Department of Psychology and Brain Sciences, University of Iowa; John E. Scofield, Department of Psychology, University of Missouri, Columbia, MO, 65211; Nathan Nunley, University of Mississippi, P.O. Box 1848, University, MS, 38677.

Correspondence concerning this article should be addressed to Jason Geller, Postal address. E-mail: jason-geller@uiowa.edu

information from word fragments instead of passively reading intact words (e.g., Slamecka & Graf, 1978 (NEWER REFERENCE)), spacing out study sessions instead of massing them (e.g., Carpenter, 2017), and having participants engage in retrieval practice after studying instead of simply restudying the information (Kornell & Vaughn, 2016). Another simple strategy that has gained some attention is to make material more perceptually disfluent. This can be done by changing the material’s perceptual characteristics (Diemand-Yauman, Oppenheimer, & Vaughan, 2011; French et al., 2013). Visual material that is masked (Mulligan, 1996), inverted (Sungkhasette, Friedman, & Castel, 2011), presented in an atypical font (Diemand Yauman et al., 2011), blurred (Rosner, Davis, & Milliken, 2015), or even in handwritten cursive (Geller, Still, Dark, Carpenter, 2018) have all been shown to produce memory benefits. The desirable effect of perceptual disfluency on memory is called the disfluency effect (Bjork, 2016)

Although appealing as a pedagogical strategy due to the relative ease of implementation, there have been several experiments that failed to find memorial benefits for perceptually disfluent materials (e.g., Magreehan, Serra, Schwartz & Narciss, 2016; Rhodes & Castel, 2008, 2009; Rummer, Scheweppe, & Schewede, 2016; Yue, Castel, & Bjork, 2013), casting doubt upon the robustness of the disfluency effect. Corroborating this, A recent meta-analysis by Xie, Zhou, and Liu (2018) with 25 studies and 3,135 participants found a small, nonsignificant, effect of perceptual disfluency on recall and ( $d = -0.01$ ) and transfer ( $d = 0.03$ ). Despite having no mnemonic effect, perceptual disfluency produced longer reading times ( $d = 0.52$ ) and lower judgments of learning ( $d = -0.043$ ). Experimentally, Geller et al.(2018) and Geller & Still (2018) manipulated several boundary conditions (e.g., level of degradation, type of judgement of learning, retentional interval, and testing expectancy) and found you can get mnemonic benefits from perceptual disfluent materials (in recognition), but it is rather fickle and not at all robust. Taken together, the evidence is weak for perceptual disfluency being a desirable difficulty.

Despite the weak evidence, perceptual disfluency is still being touted as a viable learning tool, especially in the popular press. Recently, reputable news sources like Washington Post (<https://www.washingtonpost.com/business/2018/10/05/introducing-sans-forgetica-font-designed-boost-your-memory/>) and NPR (<https://www.npr.org/2018/10/06/655121384/sans-forgetica-a-font-to-remember>) claimed that a new font called Sans Forgetica can enhance memory. Since the release of those articles, the Sans Forgetica font is available on all operating systems (all you have to do is download the font file), some browsers (e.g., Chrome), and as a phone application. As of this writing no peer-reviewed research has been released that supports the assertions that Sans Forgetica enhances memory.

## What do we know about Sans Forgetica?

There is not a lot information on Sans Forgetica. What we do know is that the typeface itself is a variation of a sans-serif typeface. Sans Forgetica is a typeface that consists of intermittent gaps in letters that are back slanted (see fig. 1). As it pertains to the empirical validation of the claims made, the website does offer some information about

52 Sans Forgetica and how the original results were obtained, but not enough information to  
53 replicate the studies.



Figure 1. Example of Sans Forgetica font.

54 According to an interview conducted by Earp (2018), In the first experiment ( $N=96$ ),  
55 the Sans Forgetica team had participants read 20 word pairs (e.g., girl - guy) in three new  
56 fonts (one of them being SF) and a typical or common font. The font pairs were presented  
57 in was counterbalanced participants. What this means is that all fonts were shown, but  
58 the same pairs were never presented in more than one type of font. Each word pair was  
59 presented on the screen for 100 ms. For a final test, they were given the cue (e.g., *girl*) and  
60 had to respond with the target (*guy*). According to the interview, targets were recalled 68%  
61 of time when presented in a common font. For cue-target pairs in Sans Forgetica, targets  
62 were recalled 69% of the time—a negligible difference.

63 In the second experiment ( $N = 300$ ), participants were presented passages (250 words  
64 in total) where one of the paragraphs was presented in Sans Forgetica font. Each participant  
65 saw five different texts in total. For each text they were asked one question about the part

written in SF and another question about the part written in standard Arial. Participants remembered 57% of the text when a section was written in Sans Forgetica, compared to 50% of the surrounding text that was written in a plain Arial font.

## Current Studies

Given the negative evidence for perceptual disfluency, we thought it pertinent to empirically examine whether Sans Forgetica leads to positive learning outcomes. The question of whether Sans Forgetica produces a mnemonic benefits has clear practical implications. In the educational domain, it would be relatively quick and easy to use Sans Forgetica. However, in order for the Sans Forgetica to be useful, it is important to note and understand both its successes and failures. To the authors' knowledge, the experiments contained herein are the first to test the replicability of the claims made about Sans Forgetica. In two high-powered preregistered experiments, we examine whether information is better remembered in Sans Forgetica, but also how it compares with other notable learning techniques—generation (Experiment 1) and pre-highlighting (Experiment 2). Comparing Sans Forgetica to other study techniques allows us to examine the mechanisms underlying the effect, if any.

## Experiment 1

In Experiment 1 we were interested in answering two questions. First, is Sans Forgetica more memorable than a normal, fluent, font (e.g., Arial)? Second, is the Sans Forgetica effect on memory similar in magnitude to the generation effect? While very little is known about Sans Forgetica, one of the most intuitively appealing theories for why Sans Forgetica font benefits memory is that of mental effort. It is believed that reading materials in Sans Forgetica requires more effort than simply reading a normal font. Essentially, the intermittent gaps of Sans Forgetica requires readers to generate or fill in the missing pieces producing a memory advantage. This mechanism of action is similar to that of the generation effect, wherein information is better remembered when generated or filled-in compared to if it is simply read. In Experiment 1 we examined the mnemonic benefit of Sans Forgetica and generation by looking at cued recall performance with weakly related pairs. If Sans Forgetica does produce a mnemonic benefit we should find that cued recall is higher for Sans Forgetica compared to normal font. Further, if it is similar to the generation effect, the magnitude of the memory benefit between the two should be similar.

## Method

**Participants.** Two-hundred and thirty people from Amazon's Mechanical Turk Service participated for money. Sample size was based on a priori power analyses conducted using PANGAEA v0.2 (Westfall, 2016). Sample size was calculated based on the smallest effect of interest (SEOI; Lakens & Evers, 2014). In this case, we were interested in powering our study to detect medium-to-large effect sizes ( $d \geq .35$ ). We choose this effect size as our SEOI due in part to the small effect sizes seen in actual classroom studies (Bulter et

al., 2014). Therefore, assuming an alpha of .05 and a desired power of 90%, a sample size of 270 is required to detect whether an effect size of .35 differs from zero. After excluding participants who 1) did not complete every phase of the experiment, 2) started the experiment multiple times, 3) reported experiencing technical problems did not indicate that they were fluent in English [^2]: This question was not asked during the experiment., or 5) reported seeing our stimuli before, we were left with 115 participants per group.

**Materials.** The preregistration for Experiment 1 can be found here: <https://aspredicted.org/3ai98.pdf>. All materials, data, and analysis scripts for both Experiment 1 can be found here (<https://osf.io/d2vy8/>). The results contained herein are computationally reproducible by going to the primary author’s github repository for the paper ([https://github.com/jgeller112/SF\\_Expt2](https://github.com/jgeller112/SF_Expt2)) and clicking on the binder button.

Participants were presented with 22 weakly related cue-target pairs taken from Carpenter, Pashler, & Vul, 2012[^1]: Two cue-target pairs (e.g., range-rifle and train-plane) had to be thrown out as they were not preseted due to a coding error. The cue-target pairs were all nouns, 5–7 letters and 1–3 syllables in length, and high in concreteness (400–700) and frequency (at least 30 per million). Free association norms (Nelson, McEvoy, & Schreiber, 1998) were used to create 22 weakly associated pairs of similar forward and backward strength.

**Design and Procedure.** Disfluency (fluent vs. disfluency) was manipulated within-subjects and within-items and difficulty type (Generation vs. Sans Forgetcia) was manipulated between participants. For half the participants, targets were presented in Sans Forgetica while the other half were presented in Arial font; for the other half of participants, targets were presented with missing letters (vowels were replaced by underscores) and the other half were intact (Arial font). After a short 2 minute distractor task (anagram generation), they completed a cued recall test. During cued recall, participants were presented 24 cues one at a time and asked to provide the target word. After they were thanked and debriefed.

Participants completed the experiment on-line via the Qualtrics survey platform hosted on Amazon Mechanical Turk. The experiment began with the presentation of 22 word pairs, shown one at a time, for 2 seconds each. The cue word always appeared on the left and the target always on the right. Immediately proceeding this, participants did a short 2 minute distractor task (anagram generation). Finally participants completed a cued recall test. During cued recall, participants were presented 24 cues one at a time and asked to provide the target word. Responses were self-paced. Once completed participants clicked on a button to advance to the next question. After they were asked several demographic questions.

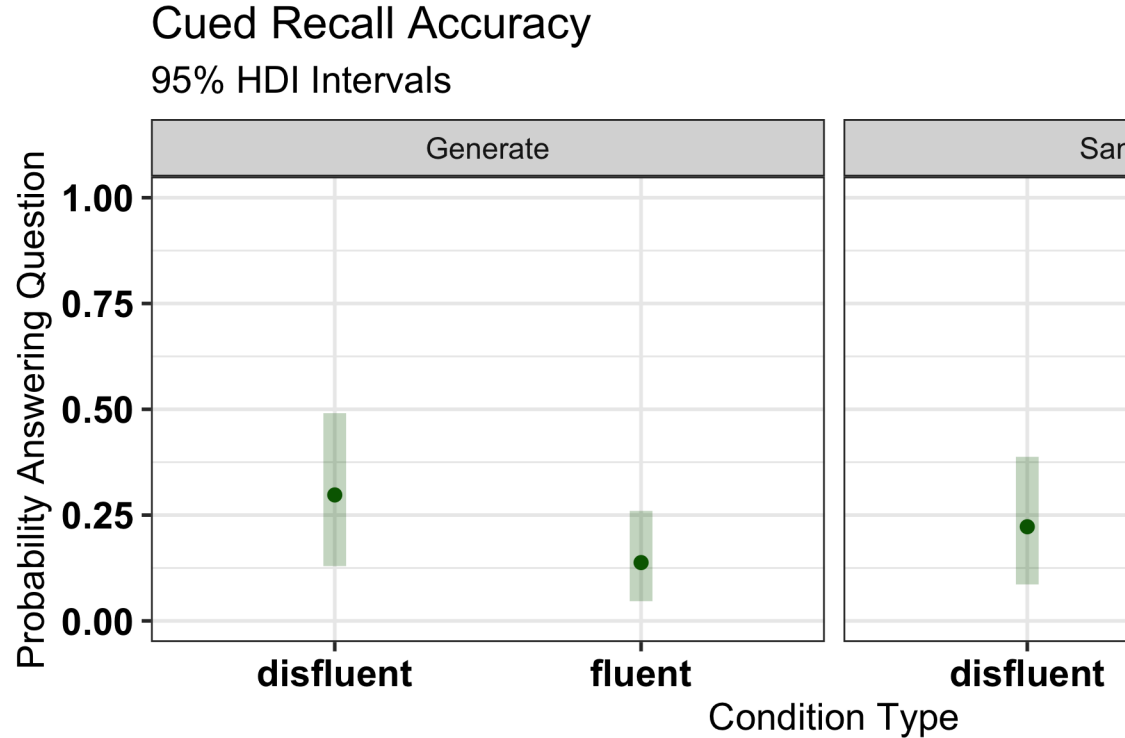
**Scoring.** Spell checking was automated with the same procedure as Experiment 1. the hunspell package in R (Ooms, 2018) using spellCheck.R. At the next step we manually examined the output to catch incorrect suggestions and to add their own corrections. Because participants were recruited in the United States, we used the American English dictionary. A nice walkthrough on how to use the package can be found in Buchcamam, De Deyne, and Montefinese (2019). Using the package, each response was corrected for misspellings. Corrected spellings are provided in the most probable order, therefore, the

first suggestion is selected as the correct answer. In the package, As a second pass, we went through and made sure the program selected the correct spelling. If the response was close to the correct response, it was marked as correct.

## Results and Discussion

Models were fit in R (vers. 3.5.0; R Core Team, 2019) with the lme4 package (vers. 2.3.1; Bates, Machler, Bolker, & Walker, 2015). We fit a logistic mixed model to predict cued recall accuracy with difficulty type (generation vs. sans forgetcia) and disfluency (fluent vs. disfluency). We fit the maximal model (formula: “brm(acc~difftypedisflu + (1+disflu|ResponseID) + (1+disflu difftype|target), family=bernoulli, data=data”). Effect sizes were labelled following Chen’s (2010) recommendations. The model’s total explanatory power is substantial (conditional  $R^2 = 0.60$ ) and the part related to the fixed effects alone (marginal  $R^2$ ) is of 0.01. The effect of difficulty was nonsignificant ( $\beta = -0.09$ ,  $SE = 0.11$ , 95% CI [-0.30, 0.13], std.  $\beta = -0.09$ ,  $p = 0.431$ ,  $d = 0.05$ ). Individuals performed better on disfluent vs. fluent conditions,  $b = 0.21$ ,  $SE = 0.06$ , 95% CI [0.09, 0.33], std.  $\beta = 0.22$ ,  $p < .001$ ,  $d = .12$ ). The interaction between difficulty type and disfluency was significant ( $\beta = 0.22$ ,  $SE = 0.04$ , 95% CI [0.14, 0.30], std.  $\beta = 0.21$ ,  $p < .001$ ,  $d = .11$ ).

To examine the strength of the interaction we examined the full model against the main effects model using the brms package (vers. 2.3.1). We used normal priors on all fixed effects. These are uninformative in terms of direction—both positive and negative effects are equally likely—but they are informative in terms of magnitude. The prior indicated that a model with the interaction term was strongly preferred over the model without the interaction ( $BF > 100$ ; Jeffreys, 1961). This suggests that the magnitude of the generation effect is larger than the Sans Forgetica effect. This can be clearly seen in Fig. 2.



171       \begin{figure}  
 172       \caption{Accuracy on Cued Recall Test. Error bars are 95% HDI derived from the brms  
 173       model.} \end{figure}

174       The results for Experiment 1 are clear-cut. Cued recall performance for cue-target  
 175       pairs presented intact and in Sans Forgetica font were equivocal. That is, we did not  
 176       observe a memory benefit for Sans Forgetica. We did, however, observe better cued recall  
 177       performance for generated compared to read targets, which replicates decades of literature  
 178       (Bertsch et al., 2007). This suggests that (1) presenting materials in Sans Forgetica does  
 179       not lead to better memory and (2) the Sans Forgetica effect is most likely not generated by  
 180       the same mechanisms that give rise to the generation effect.

## 181       Experiment 2

182       Experiment 1 failed to find a memory benefit for Sans Forgetica effect. A limitation  
 183       of Experiment 1 is that simple paired associate learning lacks educational realism. To  
 184       remedy this, Experiment 2 tested the mnemonic effects of Sans Forgetica using more realistic  
 185       materials. Whereas Experiment 1 tested whether Sans Forgetica is driven by generation,  
 186       Experiment 2 examined whether the Sans Forgetica effect might exert its mnemonic  
 187       benefit by making material more distinctive. Specifically, Sans Forgetica may make the  
 188       marked portion of text more memorable because it stands out from the surrounding text.  
 189       This is similar to the effects of pre-highlighting on learning. Indeed, some evidence supports  
 190       this type of role for highlighting: When students read pre-highlighted passages, they recall  
 191       more of the highlighted information and less of the non-highlighted information compared  
 192       to students who receive an unmarked copy of the same passage (Fowler and Barker 1974;

Silvers and Kreiner 1997). To this end, Experiment 2 compared cued recall performance on a passage where some of the sentences were either presented in: SF, pre-highlighted in yellow, or unmodified. If the Sans Forgetica effect is mainly driven by distinctiveness, words presented in Sans Forgetica should benefit more from disfluency than the passage presented unmodified. Further, the benefit for Sans Forgetica should be similar in magnitude to the pre-highlighting condition.

The pre-registration form for Experiment 2, which includes hypotheses, planned analyses, exclusion criteria, and sample size justification, can be found at: <https://aspredicted.org/3jz3z.pdf>.

## Participants

Five hundred and twenty-eight undergraduates participated for partial completion of course credit. Sample size was based on a priori power analyses conducted using PANGAEA v0.2. Sample size was calculated based on the smallest effect of interest (Lakens & Evers, 2014). Similar to Experiment 1, we were interested in powering our study to detect a medium-sized effect size ( $d = .35$ ). Therefore, assuming an alpha of .05 and a desired power of 90%, a sample size of 170 per group is required to detect whether an effect size of .35 differs from zero. After excluding participants based on our preregistered exclusion criteria, we were left with unequal group sizes. Because of this, we ran six more participants per group, giving us 176 participants in each of the three conditions.

Participants were 528 undergraduates who participated for partial completion of course credit. Sample size was based on a priori power analyses conducted using PANGAEA v0.2. Sample size was calculated based on the smallest effect of interest (Lakens & Evers, 2014). In this case, we were interested in powering our study to detect a medium-sized effect size ( $d = .35$ ). Therefore, assuming an alpha of .05 and a desired power of 90%, a sample size of 170 is required to detect whether an effect size of .35 differs from zero. After excluding participants based on our preregistered exclusion criteria, we were left with unequal group sizes. Because of this, we ran six more participants per group, giving us 176 participants in each of the three conditions.

## Materials

Participants read a passage on ground water (856 words) taken from the U.S. Geological Survey (see Yue et al., 2014). Eleven critical phrases<sup>1</sup> each containing a different keyword, were selected from the passage (e.g., the term *recharge* was the keyword in the phrase: Water seeping down from the land surface adds to the ground water and is called recharge water.) and were either presented in SF, highlighted, or unmodified. Then, 11 fill-in-the-blank questions were created from these phrases by deleting the keyword and asking participants to provide it on the final test (e.g., Water seeping down from the land surface

<sup>1</sup>originally we had 12 critical phrases but a pilot test showed that one of the questions was repeated twice so we removed one of them and also added a manipulation check question to sure participants were paying attention



adds to the ground water and is called \_\_\_\_\_ water). There was 1 manipulation check question: “What was the passage you read on?”

## Design and Procedure

Participants were randomly assigned to either the pre-highlighted condition, sans forgetica condition, or unmodified condition. Our design manipulated three difference types of passages between-subjects: pre-highlighting, sans forgetica, and normal.

Participants completed the experiment on-line via the Qualtrics survey platform. Participant read a passage on ground water. Participants were given 10 minutes to read the passage. Participants in the pre-highlighted condition received some of the passages in yellow highlighting. Participants in the sans forgetica condition were presented some of the sentences in the sans forgetica font. Participants in the normal passage condition were presented sentences with no changes. All participants were instructed to read the passage as though they were studying material for a class.

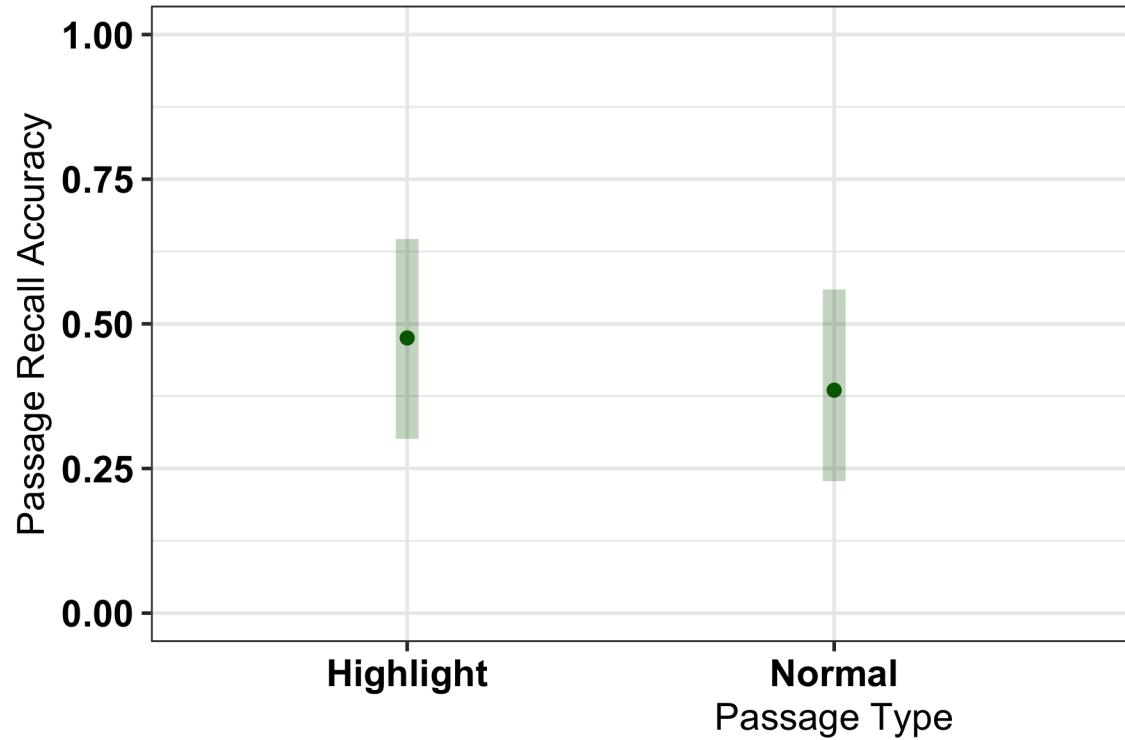
After 10 minutes, all participants were given a brief questionnaire (2 questions) asking them to indicate their metacognitive beliefs after reading the passage. The two questions were: “Do you feel that the presentation of the material helped you remember” and “How likely is it that you will be able to recall material from the passage you just read on a scale of 0 (not likely to recall) to 100 (likely to recall) in 5 minutes?” Participants were then given a short distractor task (anagrams) for 3 minutes. Finally, all participants were given 12 fill-in-the-blank test questions, one at a time.

## Scoring

Spell checking was automated with the same procedure as Experiment 1.

## Results and Discussion

We fit a logistic mixed model in a similar fashion to Experiment 1. We fit a model with passage type as a fixed effect and random intercepts for participants ( $n=528$ ) and questions ( $n=11$ ): (formula: `acc=glmer(auto_acc~passage_type+(1|Participant) + (1|Question), data=data, family="binomial")`). We hypothesized that recall for pre-highlighted and sans forgetica sentences would be better remembered than normal sentences and that there would be no recall differences between the highlighted and sans forgetica sentences. Our hypotheses were partially supported (see Fig. 2). Results indicated that pre-highlighted sentences were better remembered than sentences presented normally,  $b = 0.38$ ,  $SE = 0.17$ , 95% CI [0.05, 0.71], std. beta = 0.38,  $p < .05$ ,  $d = .21$ , and were marginally better remembered than sentences presented in Sans Forgetica,  $b = -.317$ ,  $exp(B) = 1.37$ ,  $SE = .168$ ,  $z = -1.89$ ,  $p = .059$ ,  $d = .76$ . Critically, there was no difference between sentences presented normally and in sans forgetica,  $b = 0.06$ ,  $SE = 0.17$ , 95% CI [-0.26, 0.39], std. beta = 0.06,  $p = 0.700$ ,  $d = .03$ . A Bayes factor using the brms package (Burkner, 2015) was computed and there is moderate evidence that there is no difference between the two conditions (BF = 7.47).



\begin{figure}  
 \caption{Passage accuracy. Error bars are 95% HDI dervied from the brms model}  
 \end{figure}

## Exploratory Analysis

In Experiment 2 we also asked students about their metacognitive awariness. Specifically we asked participants: “How likely is it that you will be able to recall material from the passage you just read on a scale of 0 (not likely to recall) to 100 (likely to recall) in 5 minutes?” Initial analyses suggest that the normal passage was given higher JOLs ( $M = 57.4$ ,  $SE = 1.97$ ) than the pre-highlighted passage ( $M = 50.3$ ,  $SE = 1.97$ ),  $t(525) = -7.08$ ,  $p = .023$ . There were no reliable differences between the pre-highlighted passage and Sans Forgetica ( $M = 53.8$ ,  $SE = 1.97$ ),  $t(525) = -3.52$ ,  $p = .415$  or between the passage in Sans Forgetica and the passage presneted normally,  $t(525) = 3.56$ ,  $p = .406$ .

contrast	estimate	SE	df	t.ratio	p.value
Highlight - Normal	-7.079546	2.7792	525	-2.547332	0.0299152
Highlight - Passage	-3.517046	2.7792	525	-1.265488	0.4153929
Normal - Passage	3.562500	2.7792	525	1.281844	0.4060534

NULL

Words presented in Sans Forgetica did not lead to better recall than words left unmodified or pre-highlighted. We did, however, observe better memory for pre-highlighted information compared to words presented unmodified or in a Sans Forgetica font.

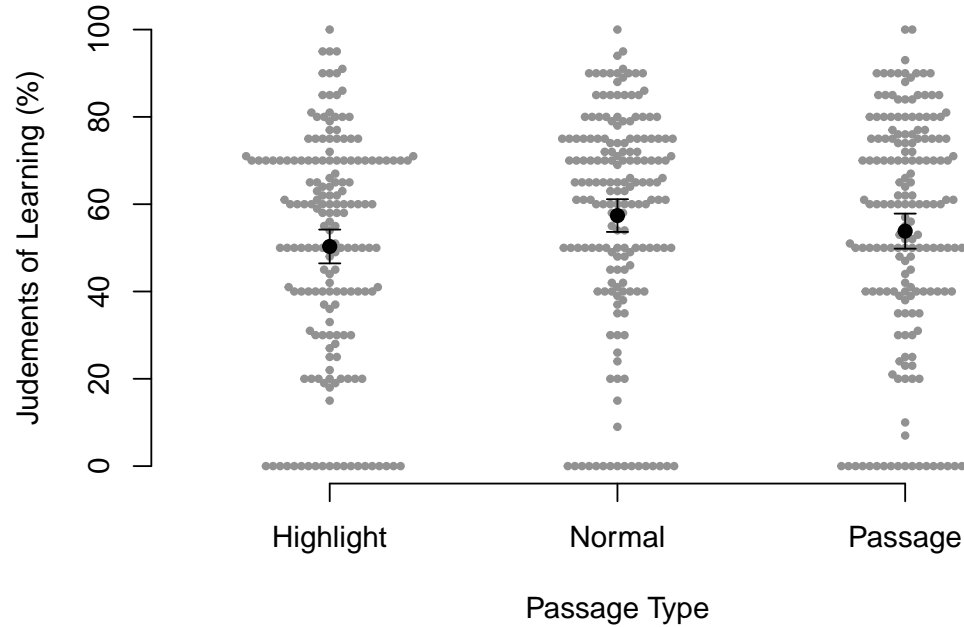


Figure 2. Judgements of learning as a function of passage type.

Table 1

contrast	estimate	SE	df	t.ratio	p.value
Highlight - Normal	-7.08	2.78	525.00	-2.55	0.03
Highlight - Passage	-3.52	2.78	525.00	-1.27	0.42
Normal - Passage	3.56	2.78	525.00	1.28	0.41

Examining metamemory judgments, we showed that a passage in Sans Forgetica does not produce lower judgements of learning compared to an unmodified or pre-highlighted font. Interestingly, individuals gave lower JOLs to prehighlighted information compared to materials presented in a normal font. One potential reason for pre-highlighted information receiving lower JOLs than the normal passage is that pre-highlighted information served to focus participants attention specific parts of the passage. Given the question, participants might have thought this would hinder them if tested over the passage as a whole.

Taken together, these results suggests that Sans Forgetica might not be a desirable difficulty.

## Discussion

While it has been reported that Sans Forgetica font can enhance memory, we report results from two high-powered experiments arguing against this claim. Specifically, we demonstrated that Sans Forgetica does not enhance cued recall for cue-target pairs (Experiment 1) or words embedded in sentences from a passage (Experiment 2). This adds to the increasing literature showing that perceptual disfluency has very little impact on actual memory performance (e.g., Magreehan et al., 2016; Rhodes & Castel, 2008, 2009; Rummer et al., 2016; Xie et al., 2018; Yue et al., 2013). While Sans Forgetica did not produce a memory benefit, we did observe a memory advantage for items that had to be generated (Experiment 1) and that were pre-highlighted.

## Limitations

**Materials-appropriate processing.** In both experiments we looked at cued recall. A recent transfer-appropriate processing (TAP) framework has contextualized when difficulties are desirable and when they are not (McDaniel & Butler, 2010). Its essence emphasizes the qualitative mismatch of the evoked encoding processes of the applied difficulty (and by the material) with respect to the required retrieval processes of the memory test. Thus, one important aspect postulated by this framework denotes the specific encoding processes stimulated by the type of difficulty applied. For example, generating incomplete word-fragments within a text intensifies the processing of the word cue and the word surroundings that help to identify the word, thus enhancing proposition-specific encoding. In contrast, creating sentence coherency in a text with randomized sentences intensifies the processing of the relationships of information in the text, thus enhancing relational encoding (McDaniel, Hines, Waddill, & Einstein, 1994). Consequently, the generation-task, which required word-generation, led to improved verbatim recall, but it was not desirable for relational test questions (and vice versa). These differently evoked encoding processes (proposition-specific versus relational) by the generation task predicted different memory effects. It is thus possible that the Sans Forgetica effect arises only under certain memory paradigms (e.g., free recall or recognition). It is hard to test this however, as the mechanism(s) that give rise to the effect are unclear and currently there is not strong evidence that the Sans Forgetica effect is reliable.

**Processing Difficulty.** One criticism put forth when examining perceptual disfluency is that studies do not objectively test (e.g., by using RTs) that stimuli are in fact perceptually disfluent (see Geller et al., 2018). Given that the two experiments contained herein were presented on-line, it is difficult to test this assumption. In Experiment 2 we used metamemory judgements as a proxy for disfluency, but we did not find that Sans Forgetica font produced lower judgements of learning.

## Conclusion

The two experiments herein represent the first tests of the claims put forth by its creators and the media. We concede that our conclusions of no effect might be a bit

332 premature. It is possible that there is an effect of Sans Forgetica, but the effect size might  
333 be smaller than we could detect across our two studies. Our SESOI was  $d = .35$ . Once more  
334 research is published, a meta-analysis can be conducted to determine the effect size and any  
335 moderating factors of the Sans Forgetica effect. Regardless it is our conclusion that Sans  
336 Forgetica lives up its name and it is really forgettable. Students looking to remember more  
337 and forget less should use other power tools () shown to enhance learning.

## References