Surprise! Low Testing Expectancy Moderates the Sans Forgetica Effect

Jason Geller[1,2]

[1] University of Iowa

[2] Rutgers Center for Cognitive Science

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

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

11   Abstract

12   One or two sentences providing a **basic introduction** to the field, comprehensible to a

13   scientist in any discipline.

14   Two to three sentences of **more detailed background**, comprehensible to scientists

15   in related disciplines.

16   One sentence clearly stating the **general problem** being addressed by this particular

17   study.

18   One sentence summarizing the main result (with the words "**here we show**" or their

19   equivalent).

20   Two or three sentences explaining what the **main result** reveals in direct comparison

21   to what was thought to be the case previously, or how the main result adds to previous

22   knowledge.

23   One or two sentences to put the results into a more **general context**.

24   Two or three sentences to provide a **broader perspective**, readily comprehensible to

25   a scientist in any discipline.

26   *Keywords:* keywords

27   Word count: X

₂₈   Surprise! Low Testing Expectancy Moderates the Sans Forgetica Effect

₂₉   The highly influential desirable difficulty principle suggests that making learning

₃₀ harder not easier, such as having students engage in retrieval practice, can have noticeable

₃₁ and lasting impacts on student achievement (Bjork & Bjork, 2011; see Sotola & Crede,

₃₂ 2020 for a recent meta-analysis). Recently, the concept of desirable difficulties has been

₃₃ extended to include subtle perceptual manipulations that are difficult to encode (e.g.,

₃₄ atypical fonts, blurring, handwritten cursive; **???**; **???**; Geller et al., 2018). One perceptual

₃₅ disfluency manipulation garnering increased attention from news outlets (NPR and

₃₆ Washingtton Post) and researchers alike is the Sans Forgetica typeface. Sans Forgetica is a

₃₇ typeface developed by a team of psychologists, graphic designers, and marketers, consisting

₃₈ of intermittent gaps and black-slanted letters (**???**). The perceptual characteristics of Sans

₃₉ Forgetica are purported to stave off forgetting and enhance learning. However, as Carl

₄₀ Sagan once said, "Extradionary claims require extradionary evidence." []

₄₁   In two independent attempts, Taylor, Sanson, Burnell, Wade, and Garry (2020) and

₄₂ Geller, Davis, and Peterson (2020) set out to examine whether Sans Forgetica is *really* a

₄₃ desirable difficulty. In the first conceptual replications of the Sans Forgetica effect, Taylor

₄₄ et al. (2020), found (in a sample of 882 people across 4 experiments) that while Sans

₄₅ Forgetica was perceived as more disfluent by participants (Experiment 1) there was no

₄₆ evidence that Sans Forgetica yielded a mnemonic boost in cued recall with highly related

₄₇ word pairs (Experiment 2) compared to a fluent typeface (Arial) or when learning simple

₄₈ prose passages (Experiments 3-4). Extending these findings, Geller et al. (2020) conducted

₄₉ three pre-registered experiments with over 800 participants, and found, similar to (**???**),

₅₀ that Sans Forgetica does not enhance learning for weakly related word pairs (Experiment

₅₁ 1), a complex prose passage on ground water (Experiment 2), or when the type of test was

₅₂ changed to a recognition memory test (Experiment 3). Taken together, across two

₅₃ independent replication attempts, and over a 1000 participants, there is weak evidence for

54   a Sans Forgetica memory effect.

55   Despite these findings, some evidence for the effectiveness of the Sans Forgetica

56   typeface does exist. For instance, Eskenazi and Nix (2020) found that Sans Forgetica can

57   enhance learning. Using eye-tracking, Eskenazi and Nix (2020) had participants learn the

58   spelling and meaning for 15 low-frequency words each presented in the context of two

59   sentences. Both orthographic discriminabity (i.e., choosing the correct spelling of a word)

60   and semantic acquisition (i.e., retrieving the definition of a word) were assessed. The

61   authors reported a memory benefit for both orthographic discrimnability and semantics for

62   words presented in Sans Forgetica compared to a normal (Courier) typeface, but only for

63   participants that were good spellers.

64   The mixed findings suggest that the Sans Forgetica may be fickel, with positive

65   effects potentially bounded by specific conditions. Probing into Eskenazi and Nix (2020),

66   an important difference between their study and (**???**) and Geller et al. (2020), is testing

67   expectancy. In Eskenazi and Nix (2020), they did did not tell their participants about the

68   upcoming orthographic and semantic discriminability test. Thus, one common design

69   feature that may moderate whether we see a Sans Forgetica effect is high testing

70   expectancy. Indeed, Eitel and Kuhl (2016) posited that testing expectancy may be an

71   important moderator of the perceptual disfluency effect. They reasoned that if the

72   disfluency effect arises because of deeper, more effortful, processing, telling participants

73   about a memory test should eliminate the effect. This occurs because testing expectancy

74   would countervail the effects of perceptual disfluency by eliciting additional processing for

75   both fluent and disfluent stimuli. In contrast, low testing expectancy is less likley to

76   impact processing of individual items,leaving effects of processing difficulty intact. Whuile

77   Eitel and Kuhl did not find evidence for this, Geller and Still (2018), using a masking

78   disfluency manipulation, demonstrated in a yes/no recognition memory test that only

79   under low testing expectancy did a disfluency effect occurr–high testing expectancy elicited

80   no disfluency effect. Given this, it is possible, then, that a Sans Forgetica effect might arise

81 when participants are not told about an upcoming memory test.

Experiment 1

83 Experiment 1 examined whether the positive effects of Sans Forgetica (as seen in
84 Eskenazi & Nix, 2020) were moderated by testing expectancy. Using a yes/no recognition
85 memory test, we manipulated whether individuals were told about an upcoming memory
86 test. In addition, we examined participants study times and judgments of learning (JOLs)
87 to Sans Forgetica stimuli. We preresgitered that the Sans Forgetica effect be influenced by
88 testing expectancy insofar when particpants were not told about a memory test we would
89 see the positive effect, but not if they were told abouyt a mmeory test.

Participants

91 We preregistered a sample size of 230. All participants were recruited through prolific
92 (prolific.co), and completed the study on the gorilla platform [www.gorilla.sc;
93 Anwyl-Irvine2020]. Participants completed the experiment in return for U.S.$8.00 an
94 hour.The sample size was based off a previous experiment (Geller et al. (2020), Experiment
95 1), wherein they calculated power to detect a medium sized interaction effect ($d = 0.35$)
96 using a similar design to the current study. After data collection had ended we had a total
97 of 231 participants. No participants met our pre-registered exclusion criteria (i.e., did not
98 complete the experiment, started the experiment multiple times, experienced technical
99 problems, or reported familiarity with the stimuli).

Method

101 Sample size, experimental design, hypotheses, outcome measures, and analysis plan
102 for Experiment 1 were can be found on the Open Science Framework (Experiments 1 and
103 2: https://osf.io/d2vy8/; Experiment 3: https://osf.io/dsxrc/). All raw and summary

data, materials, and R scripts for preprocessing, analysis, and plotting can be found at https://osf.io/d2vy8/.

Across three experiments presenting material in Sans Forgetica did not enhance memory. A common design feature of all three experiments (and the Taylor et al., 2020 replication) was high testing expectancy. That is, participants were told there would be an upcoming memory test. Eitel and Kuhl (2016) posited that testing expectancy may be an important moderator of the disfluency effect. They reasoned that if the disfluency effect arises because of deeper, more effortful, processing, telling participants about a memory test should eliminate the effect. This occurs because testing expectancy would countervail the effects of perceptual disfluency by eliciting additional processing for both fluent and disfluent stimuli. In contrast, incidental instructions are less likely to impact processing of individual items,leaving effects of processing difficulty intact. In their study, Eitel and Kuhl found that testing expectancy lead to better learning, overall, but they failed to find a disfluency effect, which makes it difficult to make inferences about potential moderators. Following up on this, Geller and Still (2018) using a masking disfluency manipulation, demonstrated in a yes/no recognition test that under incidental encoding instructions, a disfluency effect occurred, but when intentional instructions were used, no disfluency effect occurred. Given this, it is possible, then, that a Sans Forgetica effect might arise when participants are not told bout an upcoming memory test.

Method

Participants. One hundred and forty four participants ($N = 144$) participated on Prolific for U.S. \$2.43. All participants were native English speakers with normal or corrected-to-normal vision. A sensitivity analysis conducted with the R package pwr indicated that our sample size provided 90% power to detect a small effect size (d = 0.16) or larger.

129   **Materials, Procedure, and Design.**   Experiment 4 was in all respects identical

130   to Experiment 3, except for the encoding instructions given to participants at the start.

131   Participants were told that they would be reading words presented in different typefaces.

132   No memory test was mentioned.

133   **Results and Discussion**

134   **Recognition Memory.**   With intentional instructions, performance was better

135   when words were presented in Sans Forgetica compared to Arial ($M_{\text{diff}} = 0.14$), $t(143) =$

136   3.16, $SE = 0.046$, $p = .002$, $d = 0.26$. There was strong evidence for an effect, ($\text{BF}_{10} =$

137   10.55).2

138   **Reaction Times.**   Self-paced reading times for Sans Forgetica were longer than

139   reading times for Arial, ($M_{\text{diff}} = 0.14$), $t(143) = 3.16$, $SE = 0.046$, $p = .002$, $d = 0.26$.

140   There was strong evidence for an effect, ($\text{BF}_{10} = 10.55$).

141   **JOLs.**   There were no JOL differences between Sans Forgetica and Arial typefaces,

142   $t(141) = 0.220$, $SE = 1.634$, $p = .826$, $\text{BF}_{01} = 11.11$).

143   Contrary to Experiments 1-3, when testing expectancy was low, we observed better

144   memory for materials in Sans Forgetica. This provides a potential boundary condition for

145   the Sans Forgetica effect. That is, when testing expectancy is high (e.g., Experiments 1-3)

146   we do not see a Sans Forgetica effect. However, we do when testing expectancy is low. This

147   might offer a potential explanation for why there is mixed evidence on the effectiveness of

148   Sans Forgetica to enhance memory (See Eskenazi & Nix, 2020). Indeed, probing into

149   Eskenazi and Nix (2020), it appears they did not tell their participants about the

150   upcoming orthographic and semantic discriminability test. The results herein might

151   explain why they did find a positive effect for Sans Forgetica in a subset of their

152   participants. Despite this, given the small effect size and the fact that studying is almost

153   always done intentionally, their is really no evidence that it should be used as a study tool.

154   RTs (one possible is optimal study hypothesis switching from harder stimuli to

155 stumuli they know). JOLs would contradict this.

References

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In *Psychology and the real world: Essays illustrating fundamental contributions to society.* (pp. 56–64). New York, NY, US: Worth Publishers.

Eskenazi, M. A., & Nix, B. (2020). Individual Differences in the Desirable Difficulty Effect During Lexical Acquisition. *Journal of Experimental Psychology: Learning Memory and Cognition.* https://doi.org/10.1037/xlm0000809

Geller, J., Davis, S. D., & Peterson, D. J. (2020). Sans Forgetica is not desirable for learning. *Memory.* https://doi.org/10.1080/09658211.2020.1797096

Geller, J., & Still, M. L. (2018). Testing expectancy, but not judgements of learning, moderate the disfluency effect. In J. Z. Chuck Kalish Martina Rau & T. Rogers (Eds.), *CogSci 2018* (pp. 1705–1710).

Geller, J., Still, M. L., Dark, V. J., & Carpenter, S. K. (2018). Would disfluency by any other name still be disfluent? Examining the disfluency effect with cursive handwriting. *Memory and Cognition*, *46*(7), 1109–1126. https://doi.org/10.3758/s13421-018-0824-6

Sotola, L. K., & Crede, M. (2020). Regarding Class Quizzes: a Meta-analytic Synthesis of Studies on the Relationship Between Frequent Low-Stakes Testing and Class Performance. *Educational Psychology Review*, 1–20. https://doi.org/10.1007/s10648-020-09563-9

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*, 1–8. https://doi.org/10.1080/09658211.2020.1758726