European Journal of Foreign Language Teaching
ISSN: 2537 - 1754
ISSN-L: 2537 - 1754
Available on-line at: www.oapub.org/edu

doi: 10.46827/ejfl.v5i2.3344 Volume 5 | Issue 2 | 2020

DISFLUENCY IN ENGLISH LANGUAGE TEACHING?

Laura Loder Buechel[i]
Zurich University of Teacher Education,
Zurich, Switzerland

Abstract:
Desirable difficulties, coined by Bjork (1994), includes concepts such as spacing learning, interleaving, and disfluency, all of which can be practiced in the English as a Foreign Language (EFL) classroom. Sans Forgetica, a disfluent font developed at RMIT University in Australia, was specifically designed to enhance learning. Seventy-two preservice elementary school teachers in Switzerland participated in this study aimed at increasing awareness of desirable difficulties in general, and more specifically in the role of disfluency in reading comprehension. No significant differences between participants receiving a text on scaffolding and desirable difficulties in EFL classrooms in Sans Forgetica or Arial were found yet this study contributes to a larger discussion of alternative practices in English language classrooms around the world.

Keywords: desirable difficulties, scaffolding, English language teaching, disfluency, teacher education

1. Introduction

More relevant than thinking about scaffolding and providing crutches to our learners might be to consider taking away the crutches and making learning more difficult – slowing it down in order to make it more sustainable. Bjork (1994) coined the concept of desirable difficulties by asking whether the fastest route to "performance" is really the best route to "learning" or if throwing in stumbling blocks to learning on the way might actually promote deeper learning more. Teachers by nature may want to scaffold but as Jaffe and Bye (2011) state, "*Making learning too easy and straightforward can cause a misleading boost in the retrieval strength without causing the deeper processing that encourages the long-term retention afforded by higher storage strength*".

1.1 Desirable Difficulties in Language Learning

We often hear of learners cramming for tests or translating and memorizing words for a test only to forget them the next day. We can question the quality of the activity (is

[i] Correspondence: email laura.buechel@phzh.ch

translating a meaningful language learning activity?) but we can also encourage and teach our learners to study better. For example, in terms of vocabulary learning Bjork and Kroll (2015) provocatively summarize the result of many studies that actually support the idea of making learning harder for better retention of lexical items, e.g. through letting learners first guess, produce and make mistakes (as more beneficial than explaining lexical items before learners have had a chance to interact with them). General categories of desirable difficulties include taking learning and studying to unfamiliar places, spacing learning by breaking content and tasks up into smaller pieces and studying regularly, interleaving or going back and forth between tasks and subjects in one study session, reducing feedback to learners, using tests for learning, and disfluency (see Benjamin, 2011 for a comprehensive description of each of these).

Dunlosky et al's (2013) analyses provide support for many of the Bjork findings, especially for spacing learning and practice testing. Breaking content and tasks up into smaller pieces and studying regularly is thought (e.g. Miles, 2010) to be better for learning than blocking content into larger-size portions or studying (cramming) the night before a big test or performance assessment. Generating materials in the language classroom can include having learners create questions to quiz themselves, taking the provided materials and writing down what they think might be included on "the test" and essentially producing teaching materials on their own. Learners can create a "crossword puzzle" from the vocabulary in the unit, write their own questions from a provided text (even if they write these questions in the local language), create their own flash cards (e.g. using a Frayer model) or use interactive notebooks with foldables for self-testing. In essence, less is more – if teachers provide fewer pre-made activities and materials, and learners produce their own activities, then although this takes time, it might not take as much time as getting through all the provided materials and might be better for learning.

1.2 Disfluent fonts

Earlier research on generation effects, (see Gardiner, Smith, Richardson, Burrows, & Williams, 1985 or Jacoby, Craik and Begg, 1979) preceded the discussion of "desirable difficulties" by suggesting that the more effort or challenge involved in a study situation lead to better recall in a test situation. For example, in recalling word lists, injecting an obstacle such as missing letters into an input (e.g. the opposites 'h _ _' : 'cold') can increase retention for the test situation - indicating the desirability of generating challenge in the study condition (see Bertsch, Pesta, Wiscott & McDaniel, 2007 who go deeper into the various conditions). Disfluency, then, is essentially a disruption which slows down the decoding process, in this case we refer to written text – text that is perhaps blurred or lighter or slowly progressively lightened or missing end of lines or letters and the intention of using such fonts is thus to add a layer of focus and attention to the text.

In the search for simple instructional design tricks that might enhance learning, using disfluent fonts could be a way of deepening learning that is easily applicable in the mainstream classroom. Publishers of any textbook play with typographic features of text to (de)emphasize content and make it "more" memorable for learners and give learners tips on what to focus on within a text. Font size, spacing, and style can also play a role in

what the reader retains (e.g. Rello, Pielot, & Marcos, 2016 or Nikmah, 2018). Yet children, and learners, seem to enjoy cryptograms, snaked sentences, word scrambles which are not so simple and that are puzzle-like. In the language classroom, techniques such as those used in the Bewegte Geschichten project (https://www.bewegte-geschichten.ch/) where learners read aloud with various disruptions (text that gets fainter by the end of each line, snaked text with no spaces, or text grouped five words at a time to make it easier and then the whole text one dense line at a time to make it more difficult) are examples where disfluency is used for instructional purposes.

Disfluency has been considered a "desirable difficulty" in some research in that "*interfering with the perceptual processing of an item leads to additional, higher-level processing, which strengthens the associations among visual, semantic, and acoustic information in the perceptual system*" (Yue, Castel & Bjork, 2013, p. 230). The Diemand-Yauman, Oppenheimer, and Vaughan (2011) studies indicate that presenting materials in a font that is harder to read but not completely illegible can be advantageous to university and high school learners' retention of content. Kühl and Eitel (2016) found effects of disfluency on many levels: reasoning, recall of words, and retention of complex materials. Disfluency has as well been studied as a tool that affects judgement making. The Díaz-Lago & Matute (2019) research indicates that reducing fluency can reduce causality bias – readers given a disfluent font were less likely to jump to conclusions from written descriptions about whether a certain drug aided in patient recovery and similar research was conducted by Korn, Ries, Schalk, Oganian, & Saalbach (2018).

However, disfluency has not always been found to be so desirable. In Xie, Zhou, & Liu's (2018) meta-analysis of twenty-five empirical articles written on the topic, no differences between readers using or not using disfluent fonts in text-based activities was shown in the participants' recall of information found within the texts, and disfluent fonts were only shown to slow down the time needed to read, but not improve performance on transfer from or recall of texts afterwards. More recently, in attempting to recreate the original studies that made the basis for the development of Sans Forgetica font, Taylor, Sanson, Burnell, Wade, and Garry (2020) found no benefits of this font on recall of information. Yue, Castel, and Bjork (2013) also indicate that recall of words is not better with a disfluent condition, yet they do question how our own reading strategies affect our judgement of learning as we are often trained to disregard blurred or hard to read text. Whilst these and some other papers (e.g. Barley, 2016 or Geller, Davis & Peterson, 2020) indicate that disfluent fonts are neither better nor worse than other accepted fonts, that they even at times even impair recall, they do question at which point a font is actually disfluent, if there isn't a dynamic at play between the learner's motivation and performance, and in the words of Bjork and Yue (2016), "*does not have the background knowledge or skills to respond to them successfully*." Furthermore, there have been almost no studies of using disfluent fonts in a foreign language classroom.

2. Study: Disfluency in a foreign language? Desirable Difficulties as a topic of instruction?

2.1 Research question and purpose

The purpose of this research is multifold. Firstly, it is to delve into the use of the novel font, Sans Forgetica, in the classroom. It looks to answer the question of upon controlling for student-teacher language levels and prior awareness of the topic, is there a difference between a group receiving a text in Arial and a group receiving it in Sans Forgetica, a disfluent font, in their awareness of spelling, retention of specific words and expressions, and comprehension of content on a gist, detail and inference level? Due to the mixed findings in the research on disfluency, the hypothesis is that students having received the text in Sans Forgetica will not perform significantly better or worse than those receiving the text in Arial.

Secondly, this study was designed for students at the Zurich University of Teacher Education in Switzerland to make undergraduate students aware of the research process, the topic of "desirable difficulties" and of basic categories in testing reading comprehension. All of these points are relevant to their future careers as primary school teachers in Switzerland.

2.2 Context

This study was embedded into the course "Teaching English 1" at Zurich University of Teacher Education during a workshop on teaching reading to Swiss EFL primary school children with pre-service teachers. Thus, students had, in the hour prior to the study, been introduced to reading constructs such as gist, detail and inference, but not to the topic of desirable difficulties. As students have courses on education in general, they are aware of concepts around scaffolding, but had not yet had any input on scaffolding or reading in EFL.

2.3 Procedure

The participants had eight minutes to read a text on desirable difficulties written in the disfluent font, Sans Forgetica, which was subsequently taken away. Students then had fifteen minutes to answer questions. After some basic questions about their language level and familiarity with the topic, there were questions related to spelling words found within the text that were most likely difficult for German speakers of English, and definitions of words found within the text that are also not cognates for German speakers. The decision to use gist, detail and inference for further questions was because one topic emphasized in teacher training is that of testing reading skills and knowing these basic categories. Table 1 shows the constructs and the items tested. After the quiet reading, participants answered the questions via google forms, and then there was time for discussion about the research and the content. The texts were collected as well to see if students had highlighted or taken notes directly onto that paper.

Table 1: Categories of test questions

Construct	Questions
Comprehension of specific words	• Learners provided with definitions of 6 words found within the text (*discard, referent, accommodations, modifications, to coin, to interleave*) and had to write the word.
Phonology – spelling	• Comprehension words also spelled correctly. • 5 words from text read aloud, participants had to write them down (*deictic, temporariness, retrieval, advantageous, retrieval*)
Retainment of details – numbers, dates, authors	• 2 questions about authors and dates and specific words found in the text. (multiple choice)
Text comprehension – gist level	• Which sentence best summarizes the text as a whole? Choose one. (multiple choice) • Which of the following topics were mentioned in the text? Choose one. (multiple choice)
Inference – the author's tone	• Which statement best reflects the author's tone in the text? (multiple choice) • The author finds that "desirable difficulties" are implementable in the classroom. (true/false)

2.4 Materials

The text (Figure 1) used was written by the author and was purposefully constructed at a level that would most likely be slightly too difficult for many of the students on an English-language but not necessarily content level (students should theoretically almost be at the European Framework of Reference C1 level and have mother-tongue academic level knowledge of German). The Flesh-Kinkaid reading level measured 14 and the CEFR level was estimated to be a beginning C2 level (http://www.roadtogrammar.com/textanalysis/) but though the text contained many specialized words such as "interleaving", it also contained many German cognates and simple structures. Were this text in German (most students' mother tongue) it would be the expected level students should be accustomed to by the end of their studies (Swiss finish the academic baccalaureate (High School) in what would be considered the 12th grade, then enter the university for a Bachelor's at what one might call "Grade 15"). Thus, in English, the text was higher than the texts students are expected to deal with in English, yet as future teachers, the topic is relevant to them and if the text were too easy, there would be no variation in the results.

The font used for the this study are Arial and Sans Forgetica (Sans Forgetica, https://sansforgetica.rmit/) which has an eight degree backwards slant and gaps in the letters itself. The original aim of developing Sans Forgetica was to create a study tool for university students that would help to encourage deeper processing of content. According to the designers, Sans Forgetica is best used for information within a text and not necessarily for an entire text, but given the mixed research on the use of disfluent fonts in education, and the lack of research on disfluent fonts in foreign language education, it could have some value for an entire text.

Stop Spoon Feeding Your Kids!

One reoccurring theme in foreign language (FL) teaching is that of scaffolding. Teachers support comprehension via gestures - deictic (showing a referent like "that there") or symbolic (representing a specific object / action) ones, offering translations, and allowing for accommodations such as allowing more time or modifications by using graded texts. Teachers support learner production through prompting and providing concrete language the learners can use. Yet Hamidi and Bagherzadeh (2018) question the basic underpinnings of scaffolding in FL teaching: we cannot detect learners' knowledge well enough and temporariness - that teachers cannot predetermine how long learners will actually need assistance.

Thus, perhaps more relevant than thinking about scaffolding and providing crutches might be to think about discarding the crutches and making learning more difficult – slowing it down in order to make it more sustainable. The perfect concept here is that of "desirable difficulties" which was coined by Bjork in 1994. As Jaffe and Bye (2011) state, "Making learning too easy and straightforward can cause a misleading boost in the retrieval strength without causing the deeper processing that encourages the long-term retention afforded by higher storage strength" and desirable difficulties has been shown to support deeper processing.

Desirable difficulties can be operationalized in the FL classroom in several ways. Making text disfluent – less accessible - through fonts such as Sans Forgetica or unpredictable (less organized or where pictures are fuzzy or contradict or deepen the text to an unexpectedly high level) can create such a situation. A bit of shock treatment such as varying where teaching takes place ("let's practice this in the bathroom!") can also do this. Making learners generate their own quiz questions and materials as opposed to just having them read or being provided with pre-made questions can also be advantageous. Finally, teaching learners to practice interleaving, or switching back and forth from studying one subject to studying another within one short study session can contribute to making learning slightly more difficult but perhaps more sustainable! Thus, aggravating learning through desirable difficulties might be a way to scaffold learning in the right dose!

Figure 1: Text (in Sans Forgetica) provided to students

3. Results

3.1 Participants

Thirty-seven students received the text in Arial and thirty-five students received it in Sans Forgetica. The students were in their first or third year of a three-year teacher education program. They were asked if they had heard of scaffolding and its relationship to desirable difficulties during their studies thus far and all but two students responded that they had "heard of scaffolding but never of "desirable difficulties"", thus their amount of previous knowledge on the subject was not controlled for.

3.2 Differences between groups

To see if there was a difference in performance between learners having received the text in Sans Forgetica and those having received it in Arial, an ANCOVA was performed using SPSS. There was no significant effect of text type on learner performance after controlling for student language level, $F(1, 70) = .405$, $p = .53$. When scrutinizing the data, whilst there was a larger range of scores by participants having received the text in Sans Forgetica, the scores were slightly lower than those having received the text in Arial (see Figure 2).

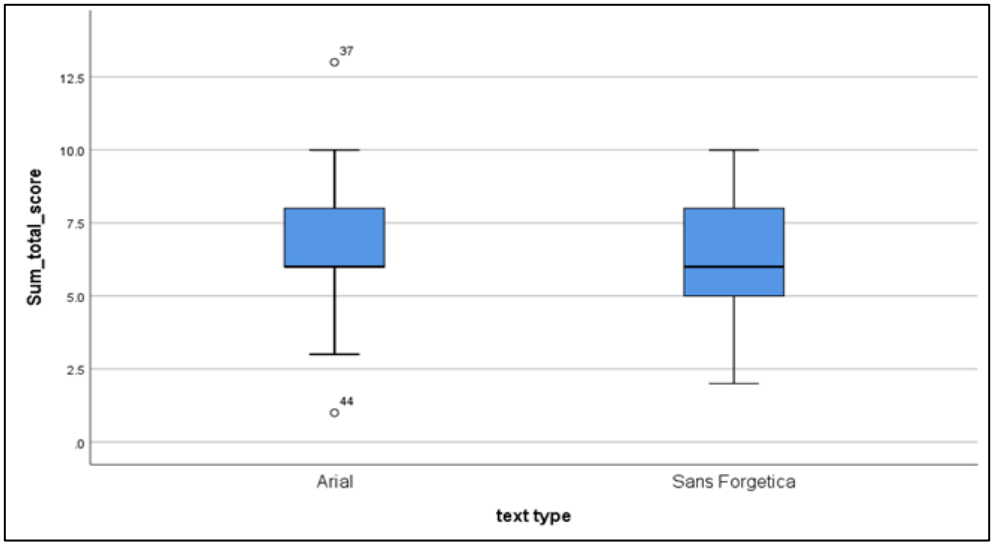

Figure 2: Distribution of scores grouped by assigned text type

Learners' reported language level provided more variation than text type (see Figure 3). The students who were mother tongue English (five total) and those who had the highest level of English also scored the highest whereas the range of scores did not vary as significantly re was no clear difference depending on the text received.

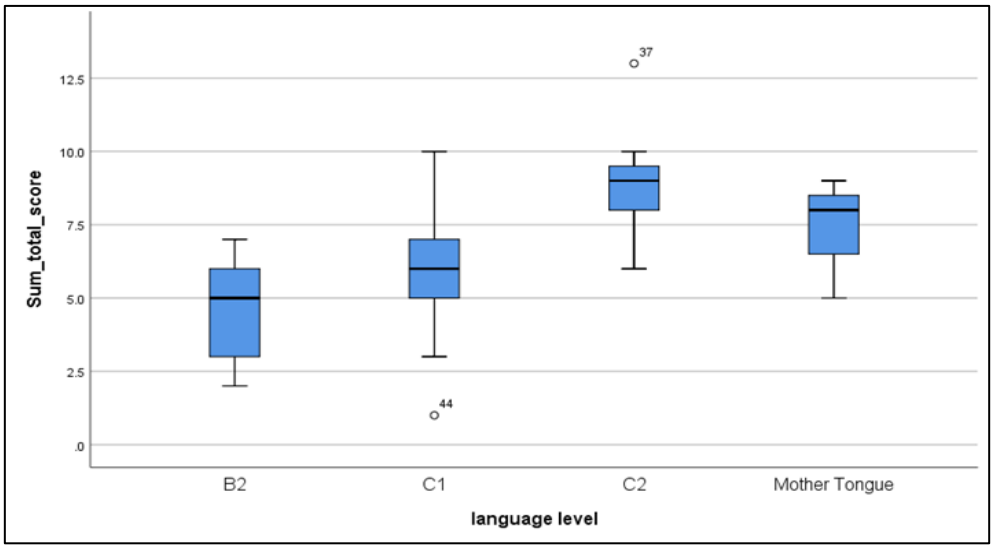

Figure 3: Distribution of total scores grouped by CEFR language level

3.3 Differences between groups on item types

Table 2: Difference in question types as related to text type provided

	Arial (mean score)	**Sans Forgetica (mean score)**	**ANOVA without controlling for language level**
Questions focused on gist	.62	.69	F: .197 p=.66
Questions focused on detail	2.11	2.63	F: 5.31 p<.05 (=.02)
Questions focused on inferencing	1	1.14	F: 1.267 p=.79

Looking at the means of the test items grouped by test construct did indicate some differences (Table 2) in performance between the two groups.

There was a significant difference in performance in the questions requiring the comprehension of details within the text (awareness of key terms mentioned and of spelling of authors' names and dates) where participants having received the text in Sans Forgetica performed better than those having received the text in Arial. However, for questions eliciting gist or having the participants infer, there was no difference and those with the Sans Forgetica font did not necessarily perform better.

3.4 Data collection - Informal observations

Students were asked to hand in their texts after having read them, and before answering the questions. Twenty students in the Sans Forgetica group and equally twenty students in the Arial group highlighted or underlined parts of the text itself, the remaining students just read and did not take notes or highlight. In the Arial group, 5 took additional notes alongside what the highlighted or underlined and 4 members of the Sans Forgetica group did so. Thus, the observable reading strategies used did not differ between the groups.

During the time provided to read the text, although students were asked to read the text quietly and then to answer the questions quietly, there was some noise during the data collection, mostly an initial shock "Oh, my, I think the printer wasn't working" and then some talking to oneself "what's this word?" After the data collection, students were handed back their texts so they could see each other's text. They were asked to speculate what the hypothesis might be. Many said that the Sans Forgetica would hinder comprehension in the short term because it was like reading out loud, you decipher the words, but don't understand the context. Others disagreed and said that if you had to focus so much on deciphering, you were also thinking.

When asked about how they felt reading a text in a disfluent text, there was a range of answers from "I loved the challenge, that was fun" to "frustrating, my English isn't that good and now THIS?". Students were also asked to recall the constructs of questions asked to their neighbors and they could do this. In a follow up activity, when provided with a text that might be used in a primary English language lesson, they all had ideas

for questions to be asked that would reflect different levels of text comprehension or form.

Students were asked to mention any parts of the text they thought were interesting and the majority were fascinated by "taking the learning elsewhere". Many said, however, that the aim of teaching was to bring the learning to the learners, not to make it more difficult, thus that scaffolding was essentially more important.

4. Discussion

The results of this study are in line with researchers such as Barley (2016) and Cacali (2016) in that the disfluent fonts did not generally increase comprehension of the text. This contradicts what the Bjork & Kroll (2015) research suggests for such use in the mother tongue. It could be that the foreign language in and of itself is enough challenge for learners and that, as Barley indicates, it's attractive to play with such fonts, they might increase motivation to read in some ways, but they do not necessarily lead to better uptake or deeper processing of the content. In this study, the questions with a focus on superficial features of text (authors names, dates and specific techniques mentioned) were more often correctly answered for learners with the text in Sans Forgetica indicating that perhaps there is a slight more attention to detail than with a smoother, more habitual font, though not enough to increase production of these same tokens or deeper comprehension of the text as a whole. As Bjork and Bjork (2011) state "*If, however, the learner does not have the background knowledge or skills to respond to them [the difficulties] successfully, they become undesirable difficulties*" (p. 58).

That there was no difference between the groups indicates that other factors or individual differences (e.g. as described in Eskenazi & Nix, 2020) might be more at play than the font itself. The learners' level of language might be the largest contributor as many students were really frustrated at having felt like they did not understand everything – the text was simply too hard. As Yue, Castel and Bjork (2013, p. 230) state "*if the task of visually interpreting and encoding distorted stimuli exceeds working memory limits, then we may see lower recall for those items; if, however, the task is within working memory limits but does not induce extra processing of the disfluent information, we may expect to see similar performance levels between fluent and disfluent conditions*." The motivation of having a challenge might be fun and encouraging for some learners, but not for others (e.g. Wenzel & Reinhard, 2019).

There are some ideas worth pursuing from the context of this study. For example, experimenting with future teachers can lead to valuable discussions about their own future classrooms such as when to task risks about "non-proven" techniques or how to teach reading. The study, the actual deciphering of a disfluent font and this in a foreign language was a real eye-opener for some as they put themselves in their own young learners' shoes. That very few students actually took notes or transferred what they had highlighted into their own words is in and of itself worth pursuing as there has been some novel research over the past years on the topic (e.g. Roediger & Pyc, 2012 or Dunlosky, et al., 2013). The discussion after the data collection about feelings of

frustration and coping with challenge can lead to more awareness of teaching strategies and motivation. The discussion of what often happens in schools that does not actually promote longer term learning is a topic that every instructor should take up with their class.

There are numerous limitations to this study. First and foremost, this study presented the entire passage (one page) in a specific font. Sans Forgetica were perhaps not intended for this purpose, but rather for features within a fluent text that are to be the focus of learning. Further research might also consider having a similar version of the test or follow up questions a week or two later to see if longer-term retention was different between the groups (e.g. Weissgerber and Reinhard, 2017). This was not possible in 2020 due to school closings during the pandemic. The question also remains about the depth of what was measured. It seems that numerous students went into the text with some knowledge of the topic, and from their answers, they left with general knowledge of the topic (e.g. many knew a little about scaffolding and wrote in "scaffolding" into the fields where the terms 'modification' or 'accommodation' was required, thus indicating they had a general idea but did not really pick up any nuances from the text), yet it cannot be said that they did not actually think about these differences – when teachers elicit specific vocabulary words, and do not get them, this does not mean that the concept is not there. Perhaps less of a focus on form and more questions with a focus on gist, detail or inferential understanding would have provided more insights into the learning which the author would actually liked to have seen. Finally, more systematic recording of individual differences via a survey or through the post-study discussion could have provided many more insights into what a disfluent font actually triggers.

This particular study leads to no conclusive results yet the topic of "desirable difficulties" is fascinating because it is implementable on many levels and encourages teachers to "think outside the box". This study leaves many questions unanswered that might be useful to pursue, those mentioned in the limitations, and on a more general level, do we scaffold instruction too much and spoon feed our learners even if a font change is not necessarily the way to go? How can performance best be measured – do teachers always emphasize in their lessons what they want to see in a particular performance? As Bjork and Yue state "*Importantly, in our opinion, and something that is not commented on in the present papers, the lack positive effects may be obscuring another potentially important finding—namely, the lack of negative effects*" (2016, p. 135). Therefore, following up on the topic does not necessarily mean that the future might not lead to some positive results. Even if using disfluent fonts is not necessarily the most important way of incorporating desirable difficulties into our lessons, the research on desirable difficulties and disfluency leads to other topics that might pan out to be valuable teaching tweaks.

About the author

Laura Loder Buechel is a teacher trainer at Zurich University of Teacher Education in Switzerland. Please visit https://phzh.ch/personen/laura.loder for more information.

References

Barley, N. (2016). Using Visually Disfluent Fonts for Cueing and Increased Reading.

Benjamin, A. S. (Ed.). (2011). Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork. Psychology Press.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & cognition*, *35*(2), 201-210.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. Psychology and the real world: Essays illustrating fundamental contributions to society, 2(59-68).

Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. The American journal of psychology, 128(2), 241.

Bjork, R. A., and Yue, C. L. (2016). Commentary: is disfluency desirable? Meta. Learn. 11,133–137. doi: 10.1007/s11409-016-9156-8

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.

Cacali, E. (2016). The Effects of Font on Vocabulary Memorization. Kwansei Gakuin University Humanities Review, 21, 111-115.

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the (): Effects of disfluency on educational outcomes. *Cognition*, *118*(1), 111-115.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. Psychological Science in the Public Interest, 14(1), 4-58.

Eskenazi, M. A., & Nix, B. (2020). Individual differences in the desirable difficulty effect during lexical acquisition. Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication. https://doi.org/10.1037/xlm0000809

Gardiner, J. M., Smith, H. E., Richardson, C. J., Burrows, M. V., & Williams, S. D. (1985). The generation effect: Continuity between generating and reading. *The American journal of psychology*, 373-378.

Geller, J., Davis, S. D., & Peterson, D. (2020). Sans forgetica is not desirable for learning.

Hamidi, E., & Bagherzadeh, R. (2018). The logical problem of scaffolding in second language acquisition. Asian-Pacific Journal of Second and Foreign Language Education, 3(1), 19.

Harris, A. (2018). Sans Forgetica: new typeface designed to help students study. Retrieved from https://www.rmit.edu.au/news/all-news/2018/oct/sans-forgetica-news-story

Jacoby, L. L., Craik, F. I., & Begg, I. (1979). Effects of decision difficulty on recognition and recall. *Journal of Verbal Learning and Verbal Behavior*, *18*(5), 585-600.

Jaffe, A., & Bye, J. (2011). Desirable difficulties in the classroom: When learning how matters as much as what. Psychology Today. Retrieved from

https://psychologytoday.com/blog/all-about-addiction/201105/desirable-difficulties-in-the-classroom.

Korn, C. W., Ries, J., Schalk, L., Oganian, Y., & Saalbach, H. (2018). A hard-to-read font reduces the framing effect in a large sample. *Psychonomic bulletin & review*, *25*(2), 696-703.

Kühl, T., & Eitel, A. (2016). Effects of disfluency on cognitive and metacognitive processes and outcomes. Metacognition and Learning, 11(1), 1-13.

Miles, S. (2010). Good Timing: The Spacing Effect and Grammar Acquisition. English Teaching, 65(2).

Nikmah, N. (2018). The Effects of Font Sizes on Students' Reading Comprehension (Doctoral dissertation, Universitas Negeri Makassar).

Rello, L., Pielot, M., & Marcos, M. C. (2016, May). Make it big! The effect of font size and line spacing on online readability. In Proceedings of the 2016 CHI conference on Human Factors in Computing Systems (pp. 3637-3648).

RMIT University. (2018). Sans Forgetica. Available at https://sansforgetica.rmit/.

Roediger III, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. Journal of Applied Research in Memory and Cognition, 1(4), 242-248.

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.

Weissgerber, S. C., & Reinhard, M. A. (2017). Is disfluency desirable for learning?. Learning and instruction, 49, 199-217.

Wenzel, K., & Reinhard, M. A. (2019). Relatively unintelligent individuals do not benefit from intentionally hindered learning: The role of desirable difficulties. Intelligence, 77, 101405.

Xie, H., Zhou, Z., & Liu, Q. (2018). Null effects of perceptual disfluency on learning outcomes in a text-based educational context: A meta-analysis.

Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. Memory & cognition, 41(2), 229-241.