Learning and Instruction 49 (2017) 199–217

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Is disfluency desirable for learning?

Sophia Christin Weissgerber*, Marc-André Reinhard

Institute of Psychology, University of Kassel, Holländische Str. 36-38, 34127 Kassel, Germany

ARTICLE INFO

Article history:
Received 11 May 2016
Received in revised form
14 February 2017
Accepted 24 February 2017

Keywords:
Desirable difficulties
Disfluency
Learning
Memory
Performance assessment

ABSTRACT

While some difficult learning conditions can improve learning, the findings regarding the contribution to learning of disfluent, hard-to-read text materials have been inconsistent. We identified test delay and disfluency manipulations as factors potentially contributing to these discrepancies. We tested students' immediate and delayed memory performance (2 weeks later) on a course text that was presented between-subjects ($N = 134$) either as perceptually disfluent with a hard-to-read-font, as lexically disfluent with 20% scrambled letters, or in its original format. By distinguishing between short-term and long-term learning, our expectations were supported; an illegible font reduced forgetting, thereby producing delayed memory benefits. We also tested whether lexical disfluency would have similar memory effects as perceptual disfluency, as the meta-cognitive perspective suggests, or whether different disfluency manipulations would have different memory effects, as ideas from a contextualized framework on desirable difficulties suggests. The findings supported the latter. The results are discussed regarding the generalizability of the disfluency effects and the implications for when disfluency is desirable.

1. Introduction

What are effective learning approaches for improving study outcomes? Students and teachers alike mistake easy learning conditions as being effective for the learning process (Bjork, Dunlosky, & Kornell, 2013; Koriat & Ma'ayan, 2005). Common sense advises them to avoid difficult-to-read text materials. This advice is contradictory to the counterintuitive intervention of enhancing learning with less fluent and harder-to-read text materials (Diemand-Yauman, Oppenheimer, & Vaughan, 2011). Support for the lay-view comes from cognitive theories emphasizing that low burdens on mental resources during studying improve learning (Sweller, Ayres, & Kalyuga, 2011).

However, a growing body of exciting research on systematically implemented learning challenges, termed *desirable difficulties*, has accumulated with a different message (Bjork, 1994): certain difficult learning conditions can foster long-term learning, although short-term performance may not be enhanced. These difficulties are desirable because overcoming their challenges facilitates beneficial encoding and retrieval processes, which results in *durable learning improvements later on*. For example, generating and practicing to retrieve new information, in contrast to the effects of reception and re-reading, enhanced long-term retention, despite the initial effect that retrieval practice was not superior to re-reading (generation-effect: DeWinstanley & Bjork, 2004; testing-effect: Roediger & Karpicke, 2006).

Other well-researched desirable difficulties exist, and meta-analytic reviews generally have described their effectiveness and their boundary conditions (e.g., spacing effect: Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; generation/testing-effect: Bertsch, Pesta, Wiscott, & McDaniel, 2007; Rowland, 2014). However, replicating to enhance learning with disfluent learning materials has been taxing. Various conceptual replication attempts of Diemand-Yauman and colleagues' (2011; abbreviated DYC) seminal disfluency findings have led to a remarkably inconsistent body of research. Supportive evidence exists (e.g., French et al., 2013; Weltman & Eakin, 2014), though some studies have found a lack of support or even counter-evidence (e.g., Eitel, Kühl, Scheiter, & Gerjets, 2014; Kühl, Eitel, Damnik, & Körndle, 2014). Discrepant findings are characteristic of research on disfluency and learning outcomes, which have employed various disfluency manipulations, including DYCs' illegible font manipulation (e.g., Besken & Mulligan, 2013; Carpenter, Wilford, Kornell, & Mullaney, 2013; Corley, MacGregor, & Donaldson, 2007; Katzir, Hershko, & Halamish, 2013; Kornell, Rhodes, Castel, & Tauber, 2011;

* Corresponding author.
E-mail addresses: scweissgerber@uni-kassel.de (S.C. Weissgerber), reinhard@psychologie.uni-kassel.de (M.-A. Reinhard).

http://dx.doi.org/10.1016/j.learninstruc.2017.02.004
0959-4752/

Sungkhasettee, Friedman, & Castel, 2011). The empirical discrepancies call into question whether disfluency actually functions as a difficulty that desirably improves learning.

Consequentially, the mixed findings have stimulated a controversial debate, which has focused on the robustness, generalizability, and practical effectiveness for education purposes of findings regarding disfluency (e.g., Rummer, Schweppe, & Schwede, 2016). This deliberation has extended to unknown moderators (Kühl, Eitel, Scheiter & Gerjets, 2014; Oppenheimer & Alter, 2014). For example, in a recently published special issue of Metacognition and Learning, multiple studies examined moderators, such as test expectancy (Eitel & Kühl, 2016), item-relatedness (Magreehan, Serra, Schwartz, & Narciss, 2016), distinctiveness (Rummer et al., 2016), and working memory capacity (Lehmann, Goussios, & Seufert, 2016). Working memory and medium of presentation (Sidi, Ophir, & Ackerman, 2016) moderated disfluency effects with an illegible font. Apparently, the perceptual disfluency effect has specific boundary conditions, is less robust, and is potentially smaller than can be inferred from DYCs' original work. In the following, we suggest two other factors that may have contributed to the inconsistencies in the literature.

Parallel to the well-established desirable difficulties, we propose that test delay is a factor and argue that disfluency may have beneficial long-term effects on learning outcomes. These long-term outcomes are determined by the role of the (fading) retrieval strength of stored information over time, and manipulations that make encoding more difficult have a larger effect after a delay because they strengthen retrieval pathways. Encoding processes should not only be relevant for delayed effects of disfluency but also for proximate effects; the kind of disfluency manipulation (e.g., perceptual vs. lexical disfluency) that is utilized may influence what kind of encoding processes are stimulated during learning and thus affect disfluency's effects on memory outcomes.

1.1. When disfluency acts as a desirable learning difficulty

We will now apply the time-dependent distinction between performance and learning that is prominent in desirable difficulty research on disfluency. *Performance* refers to impermanent changes in knowledge that are observed during and/or immediately after acquisition. In contrast, *learning* encompasses durable alterations in knowledge that are observable later in time (Soderstrom & Bjork, 2015). Importantly, immediate performance is not a good indicator of long-term learning (Bjork, 1994). In the classroom experiment (Study 2) of DYCs' (2011) controversial paper, the studying time period of the disfluent learning materials with the hard-to-read fonts ranged from about a week and a half to a month. After the instructional units and final exams were completed, the results showed enhanced learning for the disfluent materials. The important detail was the implicit time delay between the disfluency manipulations and the assessment of their effects, mirroring an intervention that tests learning, not performance. When directly looking at those replication studies that tested the impact of an *illegible font on retention*, an interesting pattern emerges. A subset of these replication studies measured only immediate performance after the manipulation, without a major time delay (see Table 1: Eitel & Kühl, 2016; Eitel et al. (2014); Lehmann et al. (2016); Magreehan et al., 2016; Rummer et al., 2016), and they found no effect. The other subset of studies (see Table 1: French et al., 2013; French, unpublished note; Weltman & Eakin, 2014) used at least some delay and found an effect. Although these studies implemented a short delay (approximately 40 min), they support the potential benefit of disfluency effects *later* in time.

We do not wish to over-interpret this apparent pattern, given that it is based on a small number of studies that represent the closest replication attempts to the original by using an illegible font manipulation rather than another disfluency manipulation, such as inverted words (Sungkhasettee et al., 2011) and by measuring effects on retention rather than another cognitive outcome, such as critical thinking (e.g., Sidi et al., 2016). Nevertheless, we entertained and systematically tested the idea that disfluency, as suggested by the desirable difficulty perspective, may show retention effects after a major time delay. The implications of the previous follow-up studies regarding the desirability of disfluency, which addressed only immediate performance, are limited. Difficulties, such as disfluent fonts that make encoding more difficult during learning, should have a larger effect after a delay because they strengthen retrieval pathways (Bjork, 1994). The question may not be whether or not an illegible font or disfluency has beneficial memory effects but when and under which circumstances.

1.2. Disfluency effects beyond meta-cognitions

These circumstances are difficult to illuminate using theories describing the (unknown) mechanism by which disfluency exerts positive effects on memory outcomes. Presumably, the experienced subjective difficulty that is associated with disfluent information processing functions as a general meta-cognitive cue to allocate more cognitive resources. In other words, the evoked meta-cognitions foster control processes that lead to deeper processing and elaboration (Alter, Oppenheimer, Epley, & Eyre, 2007), which in turn improves memory performance (Diemand-Yauman et al., 2011). A related assumption holds that disfluency may improve meta-cognitive accuracy because it leads to more appropriate control processes that improve performance (Pieger, Mengelkamp, & Bannert, 2016). From this perspective, disfluency should be desirable across various contexts, and its effects should be independent of the applied disfluency manipulation.

Accordingly, different representations of disfluency have evoked the same general meta-cognitive experience; visually perceptually disfluent items, like blurred words (Yue, Castel, & Bjork, 2013), smaller fonts (Kornell et al., 2011; Rhodes & Castel, 2008), masked items (Besken & Mulligan, 2013, 2014), auditory perceptually disfluent content (Carpenter et al., 2013), and linguistic phonological disfluent materials (Jia et al., 2016; Miele, Finn, & Molden, 2011) all evoked lower judgments of learning (abbreviated JOLs), despite being qualitatively distinctive disfluency manipulations (Alter & Oppenheimer, 2009).

Surprisingly, however, and not in line with the meta-cognitive explanation of disfluency effects, some of these different representations of disfluency failed to produce or lead to negative memory effects, reflecting a dissociation between their effects on meta-cognitions and memory. For example, smaller fonts lead to lower JOLs than larger fonts in a word-learning task, but actual memory was equal for both font types (Kornell et al., 2011). Intact words, compared to backward masked words, produced higher JOLs, but actual memory was either better for perceptually distorted words (Besken & Mulligan, 2013; Exp. 1) or was equal (Exp. 2). Blurred words, compared to clear words, led to lower JOLs, but given sufficient processing time, actual memory performance was the same (Yue et al., 2013). Inverted words, compared to upright words, resulted in similar JOLs, but inversion actually improved memory performance (Sungkhasettee et al., 2011).

Within the perspective of meta-cognitions, these dissociations from memory effects have been explained (post hoc) as a failure to evoke adequate meta-cognitively induced regulation processes. It is important to note that this framework would not predict that differences in the applied disfluency manipulation would have different memory effects. In stark contrast, the empirical evidence on the disfluency effect emphasizes its context-dependent nature