Effects of disfluency on cognitive and metacognitive processes and outcomes

Tim Kühl¹ · Alexander Eitel^{2,3}

Received: 12 February 2016 / Accepted: 15 February 2016 / Published online: 10 March 2016
© Springer Science+Business Media New York 2016

Introduction

As counterpart of fluency, the concept of *disfluency* refers to the metacognitive experience of ease or difficulty associated with completing a mental task. If task completion is perceived as easy or fluent, one often uses heuristics and intuitions to process information. If task completion is perceived as difficult or disfluent, by contrast, one is more likely to engage in effortful and analytic processing (Alter et al. 2007). The metacognitive experience of difficulty or disfluency can be created by making information harder to process on a perceptual level, for instance, by presenting text in fonts that are slightly harder to read (Alter et al. 2007). Such disfluency on a perceptual level can be desirable to self-regulated learning, because it can function as a metacognitive cue that one may not have mastery over material, leading to more effortful and analytic processing, and in turn, to better performance. Two experiments by Diemand-Yauman et al. (2011) lend support to these assumptions, showing that using harder-to-read fonts to present word lists and text fostered learning – and not only in the lab but even in an applied school setting.

Unsurprisingly, the two experiments published by Diemand-Yauman et al. (2011) captured a lot of attention, which is reflected in 150 citations of the Diemand-Yauman et al. article in less than five years (source: Google Scholar®; 02/12/2016) or in several articles in the popular press (e.g., New York Times: http://www.nytimes.com/2011/04/19/health/19mind.html?pagewanted=all&_r=0). However, to the best of our knowledge, only a few published studies were able to confirm the basic effect in conceptual replications (French et al. 2013; Sungkhasettee et al. 2011; Weltman and Eakin 2014) whereas other studies found null effects or even negative effects of degraded text fonts on learning outcomes (e.g., Eitel et al. 2014;

✉ Tim Kühl
tim.kuehl@uni-mannheim.de

¹ Psychology of Education, University of Mannheim, P.O. Box 103462, 68131 Mannheim, Germany

² Leibniz-Institut für Wissensmedien, Tübingen, Germany

³ Department of Psychology, University of Freiburg, Freiburg, Germany

Miele et al. 2013; Yue et al. 2013). Given the high number of studies citing the Diemand-Yauman et al. article, the lack of published studies replicating the original effect is surprising. One explanation is that there exist several studies failing to replicate the disfluency effect so that they may end up in the file-drawer rather than in a scientific journal (cf. file-drawer problem; Rosenthal, 1979). This may result in a biased estimation of the overall disfluency effect when being based on published studies only. Thus, to be better able to estimate the real size of the (positive) effect of degraded fonts on learning outcomes, the first aim of this special issue was to accumulate empirical evidence in (dis-)favor of degraded fonts by circumventing a potential publication bias (i.e., easier to publish positive results than null effects). However, to make sure that null effects are not due to a flawed study methodology, but can be considered meaningful in the sense of pointing towards important boundary conditions, only studies with sound experimentation and sufficient sample size were accepted for this special issue (see Schüler et al. 2011, for a similar approach). Additionally, so far empirical research on disfluency largely revolved around the question of *whether*, instead of *when* and *how*, degraded fonts foster learning and understanding (Kühl et al. 2014a, b). Therefore, the second aim of the empirical studies of this special issue was to test potentially moderating as well as mediating variables of disfluency to uncover the cognitive and metacognitive processes associated with it.

Applying this rationale, we were able to assemble six manuscripts investigating effects of disfluency within this special issue that are commented by leading experts in the field (see commentaries by Bjork and Yue (2016), as well as Dunlosky and Mueller (2016)). The six manuscripts comprise 13 experiments with a total of more than 1000 participants, including attempts to replicate experiment 1 of Diemand-Yauman et al. (2011) both more directly and conceptually. Materials and measures of the studies of this special issue range from short syllogisms to complex expository text with pictures and transfer tests. Studies test the disfluency effect on processes and outcomes on both a cognitive and a metacognitive level. Moreover, studies investigate potential moderators (e.g., working memory) and mediators (e.g., confidence judgments or visual attention) of the disfluency effect to 1) assess whether disfluency could be recommendable as an instructional intervention under certain boundary conditions, and to 2) shed light on the mechanism behind the (missing) disfluency effect. To this end, research on disfluency will be categorized and interpreted in light of theories on metacognition and self-regulated learning. This rationale will be explained in the following.

Disfluency theory and metacognition

Disfluency theory is built upon considerations of James (1890/1950), who stated that humans possess two distinct processing systems: one that is quick, effortless, associative, and intuitive (System 1) and another one that is slow, effortful, analytic, and deliberate (System 2). Alter et al. (2007) applied this dual-process model to effects of processing fluency in decision-making. They argued that whether System 1 or System 2 is activated also depends on the perceived ease or difficulty associated with a cognitive task, which can be operationalized by printing text in fluent or disfluent fonts. If information processing is perceived as easy (fluent font), it is more likely that System 1 is activated, leading to an effortless and intuitive processing. If, on the other hand, information processing is perceived as difficult (disfluent font), System 2 will be more likely

activated, resulting in more invested mental effort and analytic processing, which affected social judgments and improved performance in reasoning tasks (e.g., Alter et al. 2007).

Importantly, activating deeper and more analytic processing (System 2) by printing text in harder-to-read fonts (disfluent) can be beneficial not only to judgments and reasoning tasks, but also to learning, as analytical processing is deemed desirable for education. Accordingly, Diemand-Yauman et al. (2011) found better recall of instructional contents by making text disfluent. These effects can be explained by referring to metacognition and self-regulated learning (Bjork et al. 2013). Particularly, following the classical metamemory framework of Nelson and Narens (1990) there is a continuous interaction between monitoring and control processes in self-regulated learning situations. Students monitor their difficulties and progress while learning in order to initiate adequate control processes (see also Serra and Metcalfe 2009). If learning seems easy and to progress well, students are confident and may invest relatively little mental effort. They believe that they can quickly reach their desired knowledge level, and terminate studying early (cf. discrepancy-reduction model; Thiede and Dunlosky 1999). If they face difficulties learning or understanding the instructional contents, they are less confident that they reach their desired knowledge level soon, and hence initiate control processes such as to invest more mental effort, to prolong studying or to restudy materials.

Aside from objective reasons such as a higher difficulty of instructional contents, the superficial appearance of instructional materials can influence how confident students are in their learning. That is, the easier or more fluently information is processed on a perceptual level, the more confident students are that the information is understood and can be recalled—even though this is not the case (e.g., Kornell et al. 2011; Yue et al. 2013). This is referred to as overconfidence (Koriat et al. 1980) – a metacognitive bias that can be detrimental to self-regulated learning (e.g., Dunlosky and Rawson 2012). Accordingly, disfluency effects can be explained, in that high fluency leads to overconfidence and System 1 processing, resulting in less mental effort and premature study termination. By contrast, lower perceptual fluency (i.e., disfluency) can influence monitoring because the contents are perceived as more difficult, leading to lower confidence (on meta-level), and in turn, control processes are initiated to activate System 2, thus increasing mental effort and prolonging studying, leading to better performance.

According to this rationale, perceptual disfluency could be assumed to foster performance across a broad range of cognitive tasks. The effects of disfluency have been found 1) for solving reasoning tasks (e.g., syllogisms), 2) for verbatim recall of word lists and even 3) for retention and comprehension of complex instructional materials. The following review of disfluency research is structured accordingly. It comprises studies on disfluency besides and within this special issue. In this review, most of the studies manipulated the perceptual fluency of written text by using fonts that are either easy or harder to read (e.g., Diemand-Yauman et al. 2011), by printing text in either bigger or smaller fonts (e.g., Rhodes and Castel 2008), or by presenting text in either a clear or a distorted or blurred manner (e.g., Eitel et al. 2014; Yue et al. 2013). However, it is important to note that this is not a comprehensive review of disfluency research, but rather serves as an introduction to frame and motivate the studies of this special issue (summarized in Table 1). To foreshadow, results from this review suggest that aside from many studies assessing disfluency effects on an outcome-level, fewer studies sought to uncover the processes and boundary conditions underlying the (missing) effect of perceptual disfluency, which is in the focus of studies within this special issue.

Table 1 Studies of this special issue, assessing potential mediators and moderators of disfluency

Study	Sample	Materials	Disfluency Manipulation	Assumed Moderators	Dependent Variables	Disfluency Effect? ^a	Moderated disfluency Effect?	Metacognitive Effect?
Sidi, Ophir, & Ackerman (Exp. 1,2)	University students (Exp. 1: 204; Exp. 2: 117)	3 reasoning tasks (from CRT; Frederick 2005)	Arial 18 pt., black (fluent) vs. Arial 9 pt., italicized light grey (disfluent)	medium (on screen vs. paper)	Task performance; Calibration ^b ; Resolution ^c	Exp. 1: No Exp. 2: No	Exp. 1: No Exp. 2: Yes, disfluent better on screen but worse on paper.	Exp. 2: Calibration fostered by disfluency on screen; Resolution fostered by disfluency on paper
Magrehan, Serra, Schwartz, & Narciss (Exp. 1–5)	University students (Exp. 1: 100; Exp. 2: 90; Exp. 3: 38; Exp. 4: 67; Exp. 5: 40)	Exp. 1: 30 unrelated, 30 related word pairs; Exp. 2: 60 unrelated pairs Exp. 3: 32 unrelated, 32 related word pairs; Exp. 4 & 5: 64 unrelated pairs	Arial 56 pt., black, bold (fluent) vs. Times New Roman, 32 pt., light gray, italicized on white background (disfluent) vs. on slightly lighter grey background (“even-more-disfluent”)	Item relatedness (to-be-learned word pairs semantically related or not); Manipulation of disfluency Within- (Exp. 3, 4, 5) vs. Between- subjects (Exp. 1, 2)	Recall performance JOL magnitude Study time	Exp. 1: No Exp. 2: No Exp. 3: No Exp. 4: No Exp. 5: No	Exp. 1: No Exp. 3: No	Exp. 1–3: JOL unaffected by disfluency Exp. 4, 5: lower JOL due to disfluency Across Exp.: Within- vs. Between-subjects moderates JOL--effect
Runner, Schweppe, & Schwede (Exp. 1–3)	University students (Exp. 1: 63; Exp. 2: 58; Exp. 3: 80)	5 lists of alien species (7 characters per species)	Arial, 16 pt., black (fluent) vs. Comic Sans MS, 12 pt., 60 % grayscale (disfluent)	Distinctiveness (one fluent and four disfluent lists vs. four fluent and one disfluent lists)	Recall performance Study time in Exp. 1	Exp. 1: No Exp. 2: No Exp. 3: No	Exp. 1: No Exp. 2: No Exp. 3: No	Exp. 1: Longer study time with more disfluent lists
Eitel & Kühl	97 university students	How an airplane achieves lift (coherent text, ca. 320 words)	Arial, 16 pt. (fluent) vs. Brush Script MT, 16 pt. (disfluent)	Test expectancy (high vs. low)	Recall and transfer performance; Perceived difficulty, mental effort, Study time	No	No	Test expectancy fostered performance
Strukelj, Scheiter, Nyström, & Holmqvist	60 university students	How an airplane achieves lift (coherent text, 280 words)	Arial, 48 pt. (fluent) vs. blurred Arial, 48 pt. (disfluent)	WMC (continuous; assessed by operation span task)	Recall performance; Visual attention (eye movements) Study time	No	No	Initially shorter reading of disfluent text; increased over time (adaptation)

Table 1 (continued)

Study	Sample	Materials	Disfluency Manipulation	Assumed Moderators	Dependent Variables	Disfluency Effect? ^a	Moderated disfluency Effect?	Metacognitive Effect?
Lehmann, Goussios, & Seufert	65 university students	Time & date differences on earth (coherent text, 1070 words)	Arial, 12 pt., black (fluent) vs. Haettenschweiler, 12 pt., grayscale 35 % (disfluent)	WMC (continuous; numerical memory updating subtest)	Recall, comprehension, and transfer performance Cognitive load		Yes, for recall and comprehension: disfluency better only with high WMC	no effects on cognitive load

^a Here the term *disfluency effect* only refers to better performance due to the disfluency manipulation

^b *Calibration* refers to the difference between overall JOL magnitude and mean performance (in recall or reasoning tasks). The closer the difference is to zero, the more accurate JOLs are on an absolute level (hence, the better “calibrated”)

^c *Resolution* reflects the extent to which JOL magnitude discriminates between learned and unlearned items. The better item X is learned and recalled, the higher the JOL should be for item X in relation to items Y or Z that are learned (and recalled) less well

Disfluency and reasoning tasks

Effects of disfluency have been investigated with reasoning tasks – for instance, with syllogisms or misleading problems from the Cognitive Reflection Test (CRT; Frederick 2005). One example item of the CRT is: “A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?” The first answer that comes to mind may be 10 cents, which is incorrect and is supposed to reflect System 1 processing. A more analytic (System 2) processing would lead to the correct answer, which is 5 cents. Studies suggest that making materials disfluent (e.g., by printing text smaller, in italics and greyscale) acts as a metacognitive cue to control task processing by activating more analytic reasoning, reflected in better performance in such tasks (Alter et al. 2007, Exp. 1; Song and Schwarz 2008). However, recent research challenges these findings. In three experiments of Thompson et al. (2013) as well as in 13 experiments of Meyer et al. (2015), null effects of disfluency for reasoning tasks occurred, which were also not moderated by cognitive ability (measured by SAT score). Pooling over all experiments including the one by Alter et al. (2007), there was an overall flat null effect of disfluency (Meyer et al. 2015), suggesting that disfluency (via font manipulation) does not activate analytic reasoning. Reasons for why disfluency did not foster performance may be that it did not act as a metacognitive cue to affect monitoring¹ (e.g., by reducing confidence), or because more accurate monitoring did not stimulate adequate control processes.

The study of Sidi et al. (2015) investigated, amongst other things, this issue by assessing metacognitive monitoring (i.e., calibration, resolution) as a dependent variable in addition to performance (in Exp. 2; see Table 1). Moreover they investigated whether the medium (on screen vs. paper) might act as a potential moderator of disfluency. Whereas Experiment 1 revealed no main effects or moderation, an interaction effect between medium and disfluency was found in Experiment 2. Results suggest that disfluency on screen indeed led to more accurate metacognitive monitoring (calibration), and in turn, to better performance. On paper, this was not the case, suggesting that the medium acts as a moderator for disfluency to work the way it should – that is, by fostering monitoring and control processes in reasoning tasks. However, this conclusion has to be drawn with caution, because Meyer et al. (2015) failed to find this moderation when comparing different studies using tasks either on screen or on paper (see also Sidi et al. 2015).

Disfluency and memory for words

Several studies have investigated disfluency effects with students having to memorize word lists (often assessing verbatim recall). These studies usually share a metacognitive perspective: Students judge during learning for each word how confident they are that they would be able to recall it later, and relations between judged and actual performance are investigated (e.g., Yue et al. 2013). Applying this procedure, effects of disfluency were investigated by manipulating the font size in which to-be remembered words are displayed – e.g., with words printed in Arial 48 pt. serving as the fluent condition, and words printed in Arial 18 pt. serving as the disfluent condition (font size effect; e.g., Rhodes and Castel 2008). This font size effect is

¹ Results of Thompson et al. (2013) speak rather in favor of disfluency not acting as a metacognitive cue, since they did not find evidence that disfluent fonts reduced respondents' confidence in their given answers.

usually investigated by presenting the different font sizes to university students in a within-subjects design (e.g., Kornell et al. 2011; Mueller et al. 2014; Rhodes and Castel 2008; Susser et al. 2013; Exp. 1). In all of these cited studies, there were *no* differences observable with respect to learning success (i.e., font size had no influence on memory). However, font size did have an impact on metacognitive monitoring: In almost all studies, JOLs were higher with large compared to small fonts, meaning that learners were more confident to remember words that were presented in a large font, even though actual performance was not better. Thus, it seems that disfluency (smaller fonts) affected monitoring by reducing confidence, which however did not lead to more adequate control processes (e.g., higher mental effort; more adaptive restudy selection), possibly leading to null effects for learning outcomes.

The studies of Magreehan et al. (2015) come to similar conclusions. Moreover, they show that the disfluency effect on monitoring (JOL magnitude) is also bound to specific conditions. First, as in Susser et al. (2013) or Yue et al. (2013), disfluency affected monitoring only when it was manipulated within-subjects, because only then, the learner recognized a difference between fluent and disfluent presentation. Second, even if disfluency was manipulated within-subjects, according to Magreehan et al. the disfluency effect on monitoring was only present when the to-be-learned word pairs were unrelated (e.g., “DOG – FORK”) rather than related (e.g., “TABLE – CHAIR”), because otherwise item relatedness was used as a cue for monitoring and therefore replaced disfluency. These results indicate that the effects of perceptual fluency on JOLs are not that robust and occur only under specific conditions. In addition, different fonts did not aid memory performance in the five experiments of Magreehan et al., suggesting that the disfluency effect on performance is bound to even stricter conditions than initially assumed.

Disfluency and learning with complex materials

Research has shown that students often have problems learning about complex topics in a self-regulated manner (Azevedo 2005). Hence, disfluency research aimed at fostering monitoring and control processes to result in better recall and comprehension of complex texts (and pictures). However, results are mixed – with studies finding positive, negative or null effects.

Positive disfluency effects have been found in two experiments of Diemand-Yauman et al. (2011). The first experiment was carried out in a laboratory setting, where university-students had to learn 21 alien features presented either in a fluent or in disfluent font, with students receiving disfluent fonts recalling more features than students receiving the fluent font. Experiment 2 was carried out in a school setting across several weeks. Instructional materials from the curriculum were made disfluent or not, and again an advantage of disfluent materials on learning outcomes was observed, hence suggesting that disfluency is beneficial across contexts and materials. This research attracted a lot of attention, but no direct (or near to direct) replication has been published since. The article of Rummer et al. (2015) comprises three attempts to replicate study 1 of Diemand-Yauman et al. (2011) using quite similar materials and manipulations. Moreover, Rummer et al. investigate a potential confound in the original study of Diemand-Yauman et al., which refers to the unusualness and distinctiveness of the disfluent (i.e., Comic Sans MS, greyscale) compared to the fluent font (i.e., Arial, black). To disentangle disfluency and distinctiveness, participants either received one fluent and four disfluent lists or they received four fluent lists and one disfluent list (Experiment 1 and 2). Using this within-subjects design, both experiments showed neither an effect of disfluency nor

of distinctiveness on recall performance. In a third experiment, Rummer et al. tried to replicate the Diemand-Yauman et al. study more directly using a between-subjects design, but there were again no effects on recall, hence questioning the generality of the disfluency effect with respect to learning outcomes.

On the other hand, the disfluency effect was replicated on a conceptual level in a study by French et al. (2013), in which pupils (age 13–16) recognized more of a short text when they learned with a disfluent than with a fluent font, irrespective of their learning prerequisites. Similarly, Weltman and Eakin (2014) were able to find a positive disfluency effect by investigating the influence of different font types on statistics learning. Participants learned in pairs and received either a fluent or a disfluent font. Results revealed that learners with the disfluent font reported lower feelings of confidence regarding their learning and performed better in a knowledge test. Hence, in line with models of metacognitive monitoring and regulation (e.g., Nelson and Narens 1990), these studies suggest that disfluency reduced confidence, and in turn, initiated adequate control processes to foster performance.

By contrast, other studies found that even when disfluency led to lower confidence it did not foster performance, but was actually detrimental to it. In particular, in a study by Miele and Molden (2010; Exp. 3), university students who read expository text in a disfluent font were less confident that they understood the contents, and accordingly performed worse in a subsequent comprehension test. Miele et al. (2013; Exp. 1) found similar results for pupils that were either at 3rd or at 5th grade. Irrespective of their grade, pupils were less confident (lower JOL²) and had lower comprehension scores when reading text in a disfluent compared to a fluent font. Moreover, in the study of Miele and Molden (2010), and of Miele et al. (2013), students and pupils took longer to read text in a disfluent font, suggesting that disfluency worked the way it should work on a process-level: Disfluency reduced confidence (on meta-level), which led to longer studying on the object-level. This did, however, not translate into better performance. One may thus conclude that additional time was needed to decipher disfluent text on a superficial level rather than to process the contents more deeply. This notion is supported by studies of Gao et al. (2011) as well as of Gao et al. (2012). In their studies, making text harder-to-read by adding visual noise increased reading times, but did not positively affect recall performance. This suggests that study times may be a too indirect and vague means to assess deeper processing of the contents, because they (also) comprise the time for the increased need to decipher disfluent text.

Hence, other studies directly asked students about the mental effort they invested for processing the study contents, which was, however, not affected by disfluency in a study of Kühl et al. (2014a, b). Eitel et al. (2014) conducted four experiments and found a positive disfluency effect only in the first of four experiments, so that the overall pattern of results was clearly in favor of a null effect of disfluency (shown by Bayesian analysis). Interestingly, though, results for reported mental effort paralleled results for performance in the studies of Eitel et al. (2014). Only in the first of four experiments, more mental effort was reported for disfluent texts, suggesting a relation between mental effort and performance. However, the effort rating is only a subjective measure for the assumed control process of effort investment. A more objective means to assess control processes is eye tracking, which has been implemented in the study of Strukelj et al. (2015). In particular, Strukelj et al. presented expository text to students either in a clear (fluent) or in a blurred (disfluent) manner. They assessed eye

² The results concerning JOL were qualified by an interaction with “naïve theory of intelligence”; that is, disfluency affected JOLs only for “entity theorists” (who believe that intelligence is fixed).

movements during learning as well as student's working memory capacity. Learning outcomes (retention and comprehension) revealed no differences between the fluent and disfluent font. Eye tracking data revealed that learners in the disfluent conditions initially spent less time on the instructional material compared to learners in the fluent conditions, and that this pattern reversed during learning. These results may suggest that students altered their reading strategies over time, with a more effortful processing occurring later in the course of instruction. This would mean that students need some time to adapt to disfluency, and then use it as a metacognitive cue for better monitoring and regulation during learning.

Aside from potential mediators such as mental effort ratings or eye movements, moderators can provide new insights and potentially reconcile inconsistent findings from the literature on disfluency (cf. Kühl et al. 2014a, b). For instance, disfluency effects might only occur under specific task conditions, as was subject to the study of Halin et al. (2014). In their study, the authors argued that a more difficult task (e.g., a task in a disfluent font compared to a fluent font) will help in concentrating on that task, so that learners will be less distracted by irrelevant background speech. The results confirmed their assumption: When background speech was present, memory for written prose was indeed better in a disfluent than in a fluent font, whereas under "normal" (silent) conditions, a disfluent font did not lead to better, but to (at least descriptively) worse performance (Halin et al. 2014). Eitel and Kühl (2015) also studied the influence of specific task conditions on the disfluency effect. They reasoned that under "normal" learning conditions, students engage in deeper processing to perform well in an upcoming test (high test expectancy). Hence, particularly when students are not instructed to learn in preparation of a knowledge test (low test expectancy) they are endangered to process superficially; only then, disfluency should lead to a change in processing behavior, and by extension, to better performance. Even though the disfluent font led to longer study times, it neither led to a higher perceived difficulty nor to more mental effort nor to better learning outcomes (for both retention and transfer). High test expectancy, by contrast, led to longer learning time and better learning outcomes, but there was no moderation between test expectancy and disfluency.³

Another potential moderator for disfluency effects to occur is the learner's cognitive ability such as intelligence (IQ) or working memory capacity (WMC). The idea is that a disrupted (disfluent) superficial quality of materials requires additional cognitive resources devoted to deciphering text etc. (similar to extraneous cognitive load; Sweller et al. 1998). Thus, fewer resources are free to be spent on processing of the contents so that learners with lower cognitive abilities might not benefit from disfluency. A study of Dickinson and Rabbitt (1991) seems to support this claim, because free recall of prose text was poorer when the text was harder to read (blurred), particularly for learners with a lower IQ. A similar moderating effect of learner's cognitive abilities was found by Lehmann et al. (2015). Lehmann et al. investigated the potential moderating role of working memory when learning with a fluent compared to a disfluent expository text. The authors found that working memory capacity was related to retention and comprehension outcomes in the disfluent but not in the fluent font conditions. Their results indicated (at least descriptively) that for university-students with lower working memory capacity, a disfluent font was disadvantageous compared to a fluent font, whereas for

³ Recently, the factor test expectancy on learning with a fluent compared to a disfluent font was also investigated in a study by Rosner et al. (2015), Exp. 1 and 2, where participants had to learn isolated words. Similarly, test expectancy did not play a moderating role.

university-students with higher working memory capacity, a disfluent font was advantageous. Hence, a disfluent font may only stimulate deeper processing of complex text when working memory capacity is large enough to counteract the higher demands that arise from working with disfluent texts. However, in contrast to the results by Lehmann et al. (2015), the disfluency effect was not moderated by working memory capacity in the study of Strukelj et al. (2015) – which might be caused by a lower complexity/length of the instructional material, by a rather subtle font manipulation or by the way working memory capacity was assessed.

To conclude, it seems that although disfluency reduced confidence and increased study times in several studies using more complex materials, most of these studies did not reveal better performance for retention or comprehension, suggesting that control processes were initiated more to deal with the superficial quality of materials than to process the contents with more effort. And although there are hints for moderators of the effect, the pattern of results is not entirely clear.

Conclusions and outlook

The first goal of this special issue was to assess whether disfluency could still be recommendable as an instructional intervention by reporting a multitude of studies testing the disfluency effect – along with its potential moderators. Results are clear-cut. All 13 studies of this special issue failed to show overall better performance due to disfluency (see Table 1), suggesting the effect either to be marginal or to be bound to specific conditions. Hence, results further question the positive impact of disfluency interventions in educational settings, and we would not recommend implementing them.

Given the overall null effects of the disfluency manipulations on performance, the potential of metacognitive monitoring and control processes to mediate the (missing) effect was limited. Nonetheless, results from this special issue provide hints as to where in the process of metacognitive monitoring and control, disfluency fails to exert its hypothesized effect. Referring to more basic research, disfluency affected monitoring (i.e., JOLs) under specific conditions – that is, when disfluency was manipulated within-subjects (with unrelated word lists; Magreehan et al. 2015) or on a computer screen rather than on paper (Sidi et al. 2015). However, the majority of results from this issue suggest that disfluency (as a metacognitive cue) did not trigger more effective control processes such as longer and more effortful studying that, in turn, fostered performance. Although more disfluent materials increased study times (e.g., Eitel and Kühl 2015; Rummer et al. 2015, Exp. 1), overall performance or the mental effort invested studying the contents was not higher (e.g., Eitel and Kühl 2015). Hence, longer study times seemed to rather reflect more intense processing on the level of decoding words than to concentrate more on the contents behind them.

Even though disfluency in terms of perceptual difficulties did not impact upon performance as supposed to, this does not necessarily generalize to effects of other difficulties on performance. Whereas the promise of fostering learning by introducing subtle disfluency manipulations in complex materials seems questionable, stronger disfluency manipulations that actually produce higher difficulties at encoding (e.g., deleting words from a passage) were found to be beneficial in more basic experimental research (Maki et al. 1990). Such difficulties can be desirable, because they require more generative processing at encoding, producing stronger memory traces and hence fostering later recall (e.g., Bjork 1994). With respect to

educational interventions, it might hence be promising to shift disfluency research from perceptual difficulties to difficulties that ask learners to generate information - and test for whether this pays off in the long run. In this respect, disfluency research could be more strongly incorporated into the framework of desirable difficulties (Bjork 1994; Bjork et al. 2013), where it may stimulate further innovations on how to improve learning.

Acknowledgments We would like to thank Roger Azevedo, Editor-in-Chief of *Metacognition and Learning*, for the opportunity to organize this special issue. Moreover, we thank all the reviewers who provided invaluable comments and feedback to improve the individual manuscripts. Last but not least, we would like to thank Katharina Scheiter for her very helpful advice concerning the handling of this special issue.

References

- Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, 136, 569–576. doi:10.1037/0096-3445.136.4.569.
- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, 40, 199–209. doi:10.1207/s15326985Sep4004_2.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. E. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Boston: The MIT Press.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. doi:10.1146/annurev-psych-113011-143823.
- Bjork, R. A., & Yue, C. L. (2016). Is disfluency desirable? *Metacognition and Learning*. doi:10.1007/s11409-016-9156-8.
- Dickinson, C. V. M. & Rabbitt, P. M. A. (1991). Simulated visual impairment: Effects on text comprehension and reading speed. *Clinical Vision Sciences*, 6, 301–308.
- Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. *Cognition*, 118, 114–118. doi:10.1016/j.cognition.2010.09.012.
- Dunlosky, J., & Mueller, M. L. (2016). Recommendations for exploring the disfluency hypothesis toward establishing whether perceptually degrading materials impacts performance. *Metacognition and Learning*. doi:10.1007/s11409-016-9155-9.
- Dunlosky, J. & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22, 271–280. doi:10.1016/j.learninstruc.2011.08.003.
- Eitel, A. & Khl, T. (2015). Effects of disfluency and test expectancy on learning with text. *Metacognition and Learning*. doi:10.1007/s11409-015-9145-3
- Eitel, A., Khl, T., Scheiter, K., & Gerjets, P. (2014). Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? *Applied Cognitive Psychology*. doi:10.1002/acp.3004.
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19, 25–42. doi:10.1257/089533005775196732.
- French, M. M. J., Blood, A., Bright, N. D., Futak, D., Grohmann, M. J., Hasthorpe, A., et al. (2013). Changing fonts in education: How the benefits vary with ability and dyslexia. *The Journal of Educational Research*, 106, 301–304. doi:10.1080/00220671.2012.736430.
- Gao, X., Levinthal, B. R., & Stine-Morrow, E. A. (2012). The effects of ageing and visual noise on conceptual integration during sentence reading. *The Quarterly Journal of Experimental Psychology*, 65, 1833–1847. doi:10.1080/17470218.2012.674146.
- Gao, X., Stine-Morrow, E. A., Noh, S. R., & Eskew Jr., R. T. (2011). Visual noise disrupts conceptual integration in reading. *Psychonomic Bulletin & Review*, 18, 83–88. doi:10.3758/s13423-010-0014-4.
- Halin, N., Marsh, J. E., Hellman, A., Hellstrm, I., & Srqvist, P. (2014). A shield against distraction. *Journal of Applied Research in Memory and Cognition*, 3, 31–36. doi:10.1016/j.jarmac.2014.01.003.
- James, W. (1950). *The principles of psychology*. New York: Dover (Original work published 1890).
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 107–118. doi:10.1037/0278-7393.6.2.107.

- Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and stability bias: Dissociating memory, memory beliefs, and memory judgment. *Psychological Science*, 22, 787–794. doi:10.1177/0956797611407929.
- Kühl, T., Eitel, A., Damnik, G., & Körmde, H. (2014a). The impact of disfluency, pacing, and students' need for cognition on learning with multimedia. *Computers in Human Behavior*, 35, 189–198. doi:10.1016/j.chb.2014.03.004.
- Kühl, T., Eitel, A., Scheiter, K., & Gerjets, P. (2014b). A call for an unbiased search for moderators in disfluency research: Reply to Oppenheimer and Alter (2014). *Applied Cognitive Psychology*, 28, 805–806. doi:10.1002/acp.3030.
- Lehmann, J., Goussios, C., & Seufert, T. (2015). Working memory capacity and disfluency effect: An aptitude-treatment-interaction study. *Metacognition and Learning*. doi:10.1007/s11409-015-9149-z
- Magreehan, D. A., Serra, M. J., Schwartz, N. H., & Narciss, S. (2015). Further boundary conditions for the effects of perceptual disfluency on judgments of learning. *Metacognition and Learning*. doi:10.1007/s11409-015-9147-1
- Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 609–616. doi:10.1037/0278-7393.16.4.609.
- Meyer, A., Frederick, S., Burnham, T., Guevara Pinto, J. D., Boyer, T. W., Ball, L. J., Pennycook, G., Ackerman, R., Thompson, V., Schuldt, J. P. (2015). Disfluent fonts don't help people solve math problems. *Journal of Experimental Psychology: General*, 144, 16–30. doi:10.1037/xge0000049.
- Miele, D. B. & Molden, D. C. (2010). Naive theories of intelligence and the role of processing fluency in perceived comprehension. *Journal of Experimental Psychology: General*, 139, 535–557. doi:10.1037/a0019745.
- Miele, D. B., Son, L. K., & Metcalfe, J. (2013). Children's naive theories of intelligence influence their metacognitive judgments. *Child Development*, 84, 1879–1886. doi:10.1111/cdev.12101.
- Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*, 70, 1–12. doi:10.1016/j.jml.2013.09.007.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G.H. Bower (Ed.), *The psychology of learning and motivation: Vol. 26* (pp. 125–173). New York: Academic Press. doi:10.1016/S0079-7421(08)60053-5
- Rhodes, M. G. & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, 137, 615–625. doi:10.1037/a0013684.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641. doi:10.1037/0033-2909.86.3.638.
- Rosner, T. M., Davis, H., & Milliken, B. (2015). Perceptual blurring and recognition memory: A desirable difficulty effect revealed. *Acta Psychologica*, 160, 11–22. doi:10.1016/j.actpsy.2015.06.006.
- Rummer, R., Schweppe, J., & Schwede, A. (2015). Fortune is fickle: Null-effects of disfluency on educational outcomes. *Metacognition and Learning*. doi:10.1007/s11409-015-9151-5
- Schüler, A., Scheiter, K., & Schmidt-Weigand, F. (2011). Boundary conditions and constraints of the modality effect. *Zeitschrift für Pädagogische Psychologie*, 25, 211–220. doi:10.1024/1010-0652/a000046.
- Serra, M. J. & Metcalfe, J. (2009). Effective implementation of metacognition. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 278–298). New York: Psychology Press.
- Sidi, Y., Ophir, Y., & Ackerman, R. (2015). Generalizing screen inferiority - does the medium, screen versus paper, affect performance even with brief tasks? *Metacognition and Learning*. doi:10.1007/s11409-015-9150-6
- Song, H. & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, 26, 791–799. doi:10.1521/soco.2008.26.6.791.
- Strukelj, A., Scheiter, K., Nyström, M., & Holmqvist, K. (2015). Exploring the lack of disfluency effect: evidence from eye movements. *Metacognition and Learning*. doi:10.1007/s11409-015-9146-2.
- 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, 973–978. doi:10.3758/s13423-011-0114-9.
- Susser, J. A., Mulligan, N. W., & Besken, M. (2013). The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory & Cognition*, 41, 1000–1011. doi:10.3758/s13421-013-0323-8.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296. doi:10.1023/A:1022193728205.
- Thiede, K. W. & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1024–1037. doi:10.1037/0278-7393.25.4.1024.

- Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, 128, 237–251. doi:[10.1016/j.cognition.2012.09.012](https://doi.org/10.1016/j.cognition.2012.09.012).
- Weltman, D. & Eakin, M. (2014). Incorporating unusual fonts and planned mistakes in study materials to increase business student focus and retention. *INFORMS Transactions on Education*, 15, 156–165. doi:[10.1287/ited.2014.0130](https://doi.org/10.1287/ited.2014.0130).
- 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, 229–241. doi:[10.3758/s13421-012-0255-8](https://doi.org/10.3758/s13421-012-0255-8).