



INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Incorporating Unusual Fonts and Planned Mistakes in Study Materials to Increase Business Student Focus and Retention

David Weltman, Mark Eakin



To cite this article:

David Weltman, Mark Eakin (2014) Incorporating Unusual Fonts and Planned Mistakes in Study Materials to Increase Business Student Focus and Retention. INFORMS Transactions on Education 15(1):156-165. <https://doi.org/10.1287/ited.2014.0130>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2014, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Incorporating Unusual Fonts and Planned Mistakes in Study Materials to Increase Business Student Focus and Retention

David Weltman

Neeley School of Business, Texas Christian University, Fort Worth, Texas 76129, d.weltman@tcu.edu

Mark Eakin

University of Texas at Arlington, Arlington, Texas 76019, eakin@uta.edu

New studies indicate that incorporating planned difficulties such as unusual font styles in study material forces students to focus and concentrate more. In our experiment 155 students under different instructors learn a core topic in operations management in an active learning environment. One group of students in each class used a standard version of the study materials while two other groups used a modified version incorporating an unusual (difficult-to-read) font or planned mistakes. Students with the unusual font materials scored highest on a post-workshop comprehension test. Students who had workshop materials with pedagogical mistakes scored second followed by students in the control group. Further, it was observed that students in the control group deceived themselves with respect to sense of material mastery.

Keywords: active learning; teaching operations management; incorporating planned difficulties; self-deception;

T-tests; chi-square tests

History: Received: August 2013; accepted: July 2014.

1. Introduction

New learning tools and techniques, which have the potential to enhance the educational experience of the learner, continue to be of particular interest to researchers (Lee (2007), Barak et al. (2006), Hansen (2006), Raelin and Coghlan (2006)). In the quest for enhanced active learning methods, not all are universally effective (Weltman and Whiteside (2010)). University students expect, and in some cases, even demand a more experiential learning environment in core business courses (Auster and Wiley (2006)). This research continues an exploration of how experiential or active learning tools can be more effectively applied, particularly in university operations management and management science core courses. The study is the first to use difficult-to-read fonts and planned mistakes to analyze effects on comprehension in a university business school environment. Post-workshop comprehension test scores from three groups are compared. We also analyze differences in perception of mastery among the groups.

Some counterintuitive thinking may help educators design better experiential learning environments. Such environments might include student material with difficult-to-read fonts or planned errors. In our experiment, students learn a core topic (decision

making using Monte Carlo simulation) in operations management in an active learning workshop environment. One group of students used a standard version of active learning materials to study the topic. Another group used the exact same workshop material content but with a difficult-to-read font. The third group of students used a slightly modified version of the materials that incorporated two planned mistakes. It was proposed that the experimental groups would score higher on the same post-workshop test because the modified material would cause the students to focus more intently and this added effort would improve student comprehension. A final part of our experiment suggests that students' perceptions regarding their comprehension of material can be inaccurate.

2. Background

Our research applies the ideas of using unusual fonts or planned mistakes in active learning workshops for business school curricula, ascertaining if statistically significant results are obtained, and further, determining whether or not these students deceive themselves with respect to mastery of the material. We seek to find out if ideas proposed primarily in liberal arts domains

apply to operations management curricula, and if so, what types of modifications may be more effective.

Fluency refers to a subjective feeling regarding the ease or difficulty associated with completing a mental task (Oppenheimer (2008)). In many instances, one associates mental ease of processing (i.e., a feeling of fluency) with understanding and comprehension when in fact the reverse is the case. Several researchers, primarily in liberal arts domains, have ongoing projects exploring the benefits of *disfluency* in training materials. Disfluency (Diemand-Yauman et al. (2011)) refers to the subjective experience of difficulty associated with cognitive operations. Disfluency may be realized by introducing planned difficulties into training materials. Our study explores the effects of disfluency in a business school environment. Alter et al. (2007) showed that disfluency can lead to improved comprehension and better results. Bjork (1994) called this disfluency effect a “desirable difficulty.” In the political domain, Hernandez and Preston (2013) found that slowing people down with a challenging or difficult-to-read font significantly reduced polarization on certain political topics. When subjects read about topics with the difficult-to-read materials and had to struggle a bit, they became more moderate in their thinking and were able to understand an opposing view better. Our research explores the idea of increasing effort either in a superficial or pedagogical manner to see if improved comprehension occurs in a core university business school course. Adding such efforts to materials may be fairly easy for educators and provide significant benefits. Research in this area has also shown a subject’s natural instincts regarding comprehension may be wrong (Diemand-Yauman et al. (2011), Sungkhasettee et al. (2011), Yue et al. (2012)).

Informally we have asked colleagues and associates about the nature of workshop materials that will yield the best results. Almost universally they answer that “materials should be presented clearly,” “straightforward,” “with large and bold fonts.” The contrary may instead be true as many studies in liberal arts domains suggest (Carey (2011), Kornell et al. (2011), Diemand-Yauman et al. (2011)). In a modification to Shane Frederick’s classic three-question cognitive reflection test (Frederick (2005)), Alter et al. (2007) demonstrate the powerful effects of using a barely legible font. In this experiment three questions, all of which have an obvious quick intuitive answer that is incorrect and a more thoughtful correct answer, are given to subjects. Half of the subjects were given the test questions in an easy-to-read, clear font and the other half were given the test in more difficult-to-read, lighter font. The participants with the easy-to-read font questions “answered an average of 63% of the questions correctly compared to an average of 82% correct” for the participants with the lighter, italicized, smaller font.

Introducing cognitive strain into materials, even in such a superficial way, has a benefit. In our research of the literature we found little or no consistency as to the font styles that were selected for various experiments. Most used an “unusual” font that was lightened to some degree. These styles included bodoni, comic sans, Dakota, mistral, and others. We found no guidelines available regarding font selection for this type of research, which is not surprising since it is new, especially in university business school environments. Thus, for our experiment, we used our judgment and word processor capabilities to select a font style we believed would cause some cognitive strain. We selected a Dakota 12-point font that was 50% lighter than a normal, unadjusted font.

A number of studies demonstrate a relationship between materials that cause cognitive strain and confidence levels in comprehension of material and recall. Diemand-Yauman et al. (2011, p. 114) found that “larger fonts led participants to believe that they would have better recall” when in fact the participants with the large fonts did not perform better than the group with the smaller font materials. Additionally, Castel et al. (2007) found that fluency, or ease of processing, was a critical factor in subjects’ overestimating their recall abilities. Confidence and fluency are highly related in many studies. The more fluent the material, the more confident the subject is in his or her assimilation of the material.

Thus, in our research we focused on two questions: was there a difference in post-workshop test scores among the groups (H1A and H1B), and did students deceive themselves about their mastery of material (H2). Since we believed students needed to concentrate harder with materials that had either a difficult-to-read font or planned mistakes, we proposed the following hypotheses:

HYPOTHESIS 1A (H1A). *Mean scores for students with font-modified materials will be greater than mean scores for students in the control group.*

HYPOTHESIS 1B (H1B). *Mean scores for students with materials that have planned pedagogical mistakes will be greater than mean scores for students in the control group.*

There is a great deal of research demonstrating the link between maximizing legibility and perceived accuracy in judgment later (Rhodes and Castel (2008), Yue et al. (2012), Diemand-Yauman et al. (2011), Kahneman (2011)). “Representational richness produces confidence because it enhances fluency” (Gill et al. (1998)). Confidence in potential answers is based in part, on cognitive ease (Kelley and Lindsay (1993)). In general, if one perceives information to be easy to assimilate, one also believes the information is comprehended and can easily be recalled. This phenomena seems to be intuitive and a natural part of our cognitive processing (Alter et al. (2007)). It follows that

perceived difficulty (or cognitive strain) should instill a lack of confidence in material comprehension. When material is difficult to process, one's confidence in understanding the material is reduced and triggers are activated that more deliberate and effortful thought is needed (Botvinick et al. (2001), Liberman et al. (2002)).

To explore these feelings of mastery (confidence) among groups we established three groups. A control group that utilized a "cognitive easy" set of materials, and two experimental groups that each had a different type of cognitive difficulty. Specifically, the following hypothesis was proposed:

HYPOTHESIS 2 (H2). *The proportion of students that feel they have mastered workshop content will be different among the three groups (group with planned pedagogical mistakes materials, group with font-modified materials, and the control group with accurate and easy-to-read materials).*

3. Experimental Approach

Three class sections with a combined total of 155 students were involved in this study. Forty-nine students were in the control group, 54 students were in the group that had workshop materials with pedagogical mistakes, and 52 students were in the group that had workshop materials created with an unusual font. In a particular class section, approximately one-third of the class received each type of workshop material set as shown in Table 1.

Overall the proportion of students that received each material set was closer to one-third than for a particular class section. In the entire study, 0.316, 0.348, and 0.335 were the proportions for each group: control, pedagogical mistakes, and unusual fonts, respectively. Slightly different proportions of students in a particular class section received each workshop material set. These proportions are shown in Appendix A. In our experiment a total of 70 copies of each workshop material set was made. On the day of the workshop the precise number of students participating was not known. Each class section had a different number of students. We attempted to distribute materials randomly by mixing them with approximately one-third of each material type being distributed in a class section. Our distribution was slightly disproportionate (by section) as shown in Appendix A. A chi-square test for differences among these proportions was performed. These differences were not statistically significant (Appendix A).

Table 1 Sample Sizes for Each Workshop Group

Material set\class section	Section 1	Section 2	Section 3	Total
Control	18	15	16	49
Pedagogical mistakes	25	17	12	54
Unusual fonts	19	19	14	52

The experiment took place in a computer lab where students worked in pairs. Students received a set of the workshop learning materials and were given 60 minutes to complete the activity. In the workshop, students developed and built probability distributions for several situations. The final part of the workshop included an interactive Excel Monte Carlo simulation spreadsheet. The spreadsheet allowed students to use three probability distributions (the uniform, normal, and triangular) to model various profit scenarios where demand and cost components varied via one of these distributions. Students entered assumptions into the spreadsheet and viewed a portfolio of potential profit outcomes along with their associated frequency of occurrence.

In the lab, undergraduate business students worked in pairs with a set of workshop materials. Each student in a pair had the exact same material as his or her partner. Workshop handouts were in the form of a Microsoft PowerPoint slide set. Three slide sets were developed: a control set, a difficult-to-read font set, and a set with planned pedagogical mistakes. Each slide set had several slides of reading passages, figures, examples, and fill-in-the blank practice exercises. The idea of the workshop was to give students an introduction to decision making under uncertainty using three basic probability distributions: the uniform, normal, and triangular.

Students learned about the uniform distribution through reading short passages and then building a simple uniform distribution from example data. Twenty trials from flipping a coin were shown with their associated outcomes ("heads" or "tails"). Students used these data to complete a partially built histogram of the corresponding frequency distribution. Also, at the end of this section students calculated the mean of a discrete uniform distribution (its minimum value plus its maximum value divided by two) and the expected frequencies for each outcome given a specified total number of observations. For example in rolling a single die 600 times, students calculated the mean outcome to be 3.5 and expected individual number of outcomes of "1," "2," "3," "4," "5," and "6" to be 100 each. Similarly, students read short passages about the normal distribution for a basic understanding. From this reading they should have understood that the normal distribution is symmetric around its center (mean or median). Both words and figures were used to explain how a Monte Carlo simulation could be applied to a normal distribution to generate trial values from the distribution. It was explained to participants that more values in a given number of trials should be near the center point than toward the outer values. The triangular distribution was presented in a similar manner. The mean of a triangular was described as the average of its three distribution

parameters: minimum value, most likely value, and maximum value. All three distribution descriptions were placed in a business decision context. Finally, student teams downloaded a prebuilt Microsoft Excel spreadsheet to decide whether or not to introduce a new product in the market. The unit revenue for the product was fixed. Two cost components (variable costs per unit and overall administrative costs) varied via previously described distributions (variable costs via a uniform distribution and administrative costs via a triangular distribution). Demand for the product varied according to a normal distribution with mean and standard deviation parameters provided. Student teams worked with the spreadsheet to analyze potential profit or loss scenarios regarding product introduction.

Three sets of active learning materials were involved in the experiment: a base or control set, a font-modified set, and a pedagogically modified set. All three workshop-learning sets contained exactly the same content, the probability distributions and the Monte Carlo simulation approach to decision making. The base or control set of materials used a standard font (Gill sans 16 point) that is easy to read. An example of this font style is shown below:

Monte Carlo simulation is a technique in which many values of a random variable are generated from a probability distribution.

This font-modified set used a text style that was more difficult to read, i.e., handwriting Dakota 12 and 14 point, 50% lighter. An example of this font style is shown in Figure 1.

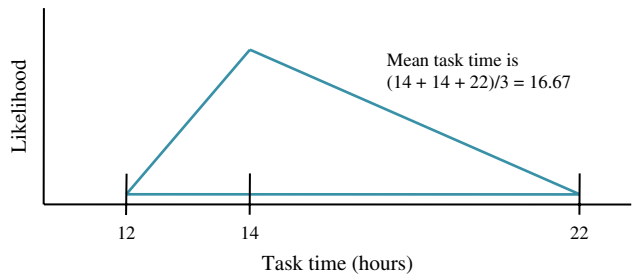
The pedagogically modified set was the same as the base set but with two planned content errors. One planned content error was an incorrect calculation of the mean for a triangular distribution. In the example presented to this group of students the mean task time should be 16 (as the smallest task time is 12 not 14). An incorrect mean task time of 16.67 was presented as shown in Figure 2.

Another pedagogical error incorrectly showed the mean of a normal distribution to be at the extreme right tail of the distribution not its center. The mean should be specified as 80 not 90 as was shown in Figure 3.

Figure 1 Modified Font Style

Monte Carlo simulation is a technique in which many values of a random variable are generated from a probability distribution.

Figure 2 Triangular Distribution with Incorrect Mean Task Time Calculation



Note. The task time is modeled below using the triangular distribution.

In the latter part of the workshop materials Monte Carlo simulation was explained. Students read that the cost of a particular part could vary between dollar amounts depending on current market conditions with all values between the two amounts equally likely (the idea of a uniform distribution). Administrative costs to introduce this item could be between three amounts but analysts believed the middle amount was most likely (the idea of a triangular distribution). Projected demand for the item was expected to follow a normal distribution. Given these parameters in the workshop, students developed a mean profit or loss. Students used an Excel worksheet to see 500 trials of possible outcomes from these distributions. Students could view individual outcomes

Figure 3 Normal Distribution with Incorrect Mean Class Grade Displayed

The normal distribution

Another very popular distribution is the normal distribution. Like the triangular distribution, the normal distribution is a continuous distribution, meaning the random variable can take on any value in its range. However, the normal distribution is always symmetric, with its mean always at the center value. Values on each side of the center are equally likely.

For example, below is a normal distribution of class grades with a mean of 90. A score around 80 is the most likely. Scores around 70 and 90 are less likely than a score around 80, but are equally likely to each other.

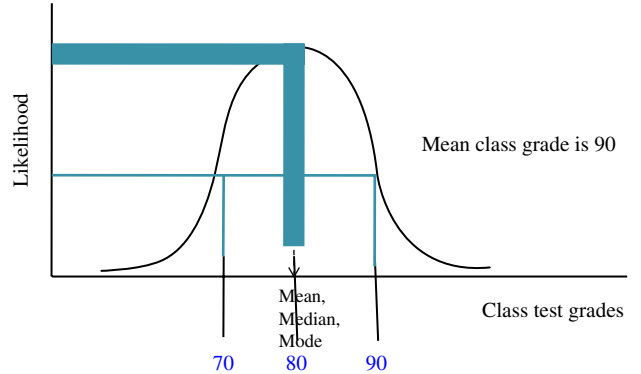
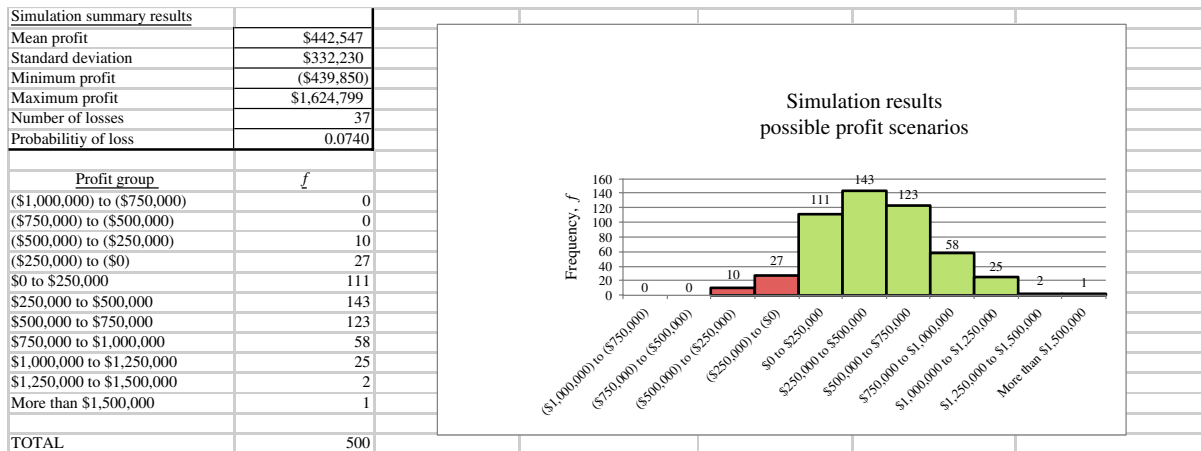


Figure 4 Student Summary of Monte Carlo Outcomes

or see a summary of all possible trials as shown below in Figure 4.

Using the simulation trial results (summary statistics and a plot of possible profit outcomes), students were asked to observe and try to understand possible profit and loss consequences resulting from introduction of this product. Students also computed the mean profit for their scenario, without reliance on the simulation results. Feedback regarding the accuracy of this computation was not provided.

A 20-minute multiple-choice quiz followed the workshop. Multiple-choice quizzes are commonly used to assess the effectiveness of various teaching methods. For example, in evaluating the effectiveness of teaching a business statistics course using “technology centric” and “technology light” methods, Gorman (2008) used a multiple-choice test to evaluate the effectiveness of these methods. A number of studies in the business education domain have used similar approaches in gauging performance (Hembrooke and Gay (2003), Rienzo and Han (2011), Coy and Adams (2012), Ueltschy (2001), Leon et al. (2006)), for example.

A classic and popular reference model that can be used for question development and to study the effectiveness of various teaching methods is Bloom’s Taxonomy of the Cognitive Domain (Bloom (1956)). Bloom’s model is well established and remains a very popular tool for development and assessment of teaching and training tools, especially in educational environments, which is the context for this study. The framework is hierarchical so that basic knowledge is assessed at the lowest level and more complex learning can be assessed at the highest levels. For a complete description and use of the taxonomy readers are encouraged to see (Bloom (1956), Bonwell and Sutherland 1996). Questions developed using this taxonomy as a guide were designed to assess

a fairly wide variety of skills obtained as shown in Appendix B. Questions were designed to test relatively simple skills such as the ability to recall and define as well as much more complex skills such as the ability to compare, apply, and employ techniques appropriately. Students were assessed immediately after a workshop session to measure the effectiveness of the technique on the subject. Performance was measured by the percent of questions answered correctly.

A common 5-point Likert scale was used for students to self-assess feelings of material mastery. Similar to Castel et al. (2007), we were interested in exploring the relationship between fluency and confidence. We wondered whether or not business school students would exhibit tendencies of self-deception based on types of materials used.

4. Methodology and Analysis of Results

4.1. Difference in Means Between Groups

We used the pooled-variance *t*-test and the Welch test that assumes unequal variances to compare mean score differences between groups. Welch’s *t*-test (Welch (1947)) is preferred when variances between samples are not equal. The Welch test overcomes potential unequal variance issues, and is also fairly robust to nonnormality population assumption violations (Rasch et al. (2011)). Sample sizes between groups need not be equal. In the present research, sample sizes in the three groups are approximately the same: control group $n = 49$, pedagogical mistakes group $n = 54$, and unusual front group $n = 52$. Both tests (pooled-variance *t*-test and Welch test) produced results with essentially the same level of significance for our data.

Table 2 Results of *t*-Tests for Difference in Means

	Font modified	Pedagogical errors	Control
<i>n</i>	52	54	49
Mean test score	0.73	0.69	0.67
Standard deviation	0.16	0.18	0.19
<i>p</i> -value (versus control group)	0.04*	0.23	Na

*Hypothesis 1A is statistically supported.

Overall the font-modified material produced the highest mean test score (0.73) followed by material with pedagogical mistakes (0.69). The lowest mean test score (0.67) was produced from the base or non-modified material control group. The *t*-test and Welch test both showed a significant difference between mean test scores for the group with font-modified materials and the control group. Hypothesis 1A regarding the use of the challenging font materials was supported for both tests at a 0.05 level of significance. Hypothesis 1B regarding the materials with planned pedagogical mistakes was not statistically supported. A summary of our *t*-test results is shown in Table 2.

Of the questions in the assessment test, three were classified as “low” level, seven as “medium” level, and three as “high” level using Bloom’s Taxonomy of the Cognitive Domain (Bloom (1956)) as a guideline (Appendix B). Medium and high level questions assess deeper understanding and application of material to new scenarios. Table 3 shows a breakdown of mean scores for the three levels of questions by experimental group. The most notable differences are between the control and font-modified groups for the medium and higher-level questions. These results suggest that introducing a significant difficulty into materials, i.e., font modification in our experiment, is especially helpful in higher levels of comprehension. For the low level questions, only three percentage points separate control and font-modified material groups, whereas for the medium and high level questions the spread is twice as much with a six percentage point separation.

4.2. Distribution of Feelings of Mastery

Our business school results were consistent with the literature. The students with the materials that were the most “representationally rich or easiest cognitively” reported the greatest feelings of confidence in

their learning. Specifically, in the control group of 49 students, 30 students (61%) reported that they either “learned the material well” or “completely mastered the material.” In the group of 54 students that had the materials with pedagogical mistakes, 17 students (31%) reported that they either “learned the material well” or “completely mastered the material.” In the group of 52 students with the font-modified materials, 27 students (52%) reported that they either “learned the material well” or “completely mastered the material.”

To determine whether or not there was evidence of student self-deception we performed a chi-square test for difference in proportions. Students with the base set of materials had the largest percentage of high feelings of mastery, whereas students that utilized the materials with pedagogical errors had the lowest percentage with high feelings of mastery (Figure 5). Our findings indicate that students in this experiment did in fact deceive themselves. Feelings of mastery among the (3) workshop material groups were distributed differently. The control group had the greatest feeling of material mastery, followed by font-modified and pedagogical error group materials, respectively. This result had a level of significance of 0.002. Hypothesis 2 was supported. Students with the font-modified material scored the highest mean comprehension score overall but this group’s percentage of high feelings of mastery was lower than the control group. Just the opposite was true for students in the control group; they had the greatest percentage of high feelings of mastery but the lowest mean test score. Students in this experiment clearly deceived themselves about what they knew. When students used workshop materials with clear bold fonts they believed they were best at assimilating the material but the converse was true. The group

Figure 5 Student Self-Deception

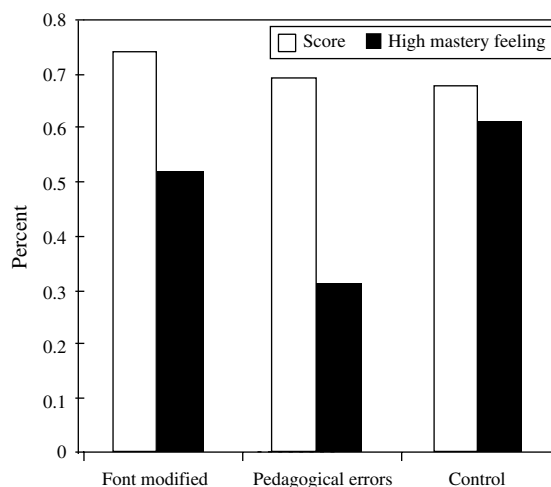


Table 3 Group Means Broken Down by Bloom Comprehension Level

	Low	Medium	High
Font-modified group	0.83	0.77	0.35
Control group	0.80	0.71	0.29

with lowest scores overall had the most confidence of mastery. Figure 5 graphically illustrates this disfluency effect along with the misalignment of student confidence.

5. Conclusions and Discussion

Introducing planned difficulties into learning materials seemed to work. In our experiment, involving 155 undergraduate business school students, mean scores were higher for subjects learning from materials that had complications. Students who used materials with an unusual font scored six points higher on average than students who used the standard materials in a post workshop test. This difference was statistically significant. Students who used the pedagogically altered material also had a higher mean post-workshop test score than students in the control group: however with our sample size of 155 students this result was not statistically significant. We believe that in the experiment our planned pedagogical errors were not obvious enough and likely overlooked by most students. Had we made these planned errors more clear, we strongly suspect the majority of students would have noticed these mistakes and slowed down to reflect on the material further. In short, our planned mistakes were too subtle. The first planned mistake presented showed the mean task time for a project activity incorrectly computed. The planned mistake was that the first number used in the computation was incorrect (the minimum value of 12 should have been used instead of the displayed value of 14). In retrospect this change was not obvious and easily missed by students. Similarly, another planned mistake showed an incorrect mean for a normal distribution. This change was also very minor and of little consequence. Had the pedagogically altered material been changed more substantially, we believe a larger, more statistically significant difference between the mean scores of the two groups would have been observed. Alter (2013) references a number of studies that indicate a “disfluency premium,” which shows that when disfluent effects are introduced, subjects answer questions correctly between 16 and 35 percent more often. When the disfluent effect is either not present or overlooked (as we believe was the case in our pedagogically altered material group) subjects do not slow down to process the information more deeply and come to an incorrect conclusion more often. Stronger support may have been obtained for H1B if more obvious planned mistakes that force students to slow down and think are incorporated into materials. In our workshop, for example, we would recommend presenting a mean task time result that is far from the correct mean, even beyond the boundaries of the distribution depicted. This type of content error would likely not be overlooked.

The disfluency effect associated with improvement in a subject’s comprehension and application of concepts seems to hold not only in liberal arts contexts but also in business contexts. Our results provide quantitative evidence demonstrating potential benefits of incorporating interesting complications into student active learning materials in a university business school environment. Providing materials that required students to slow down and struggle through them even in a somewhat superficial way improved content comprehension in our experiment. When developing student materials, especially in active learning contexts such as ours, educators should consider incorporating planned difficulties. Our study showed the use of a difficult-to-read font to be the more effective than planned pedagogical mistakes.

The self-deception effect also held true in our business context. There were significant differences in distributions regarding self-assessed comprehension. These differences were counterintuitive. The mean score for students in the control group whom had the largest percentage of students with feelings of mastery was in fact the lowest of the three experimental groups. The other two experimental groups had higher mean test scores, but lower percentages in these groups felt a sense of material mastery. The students with the font-modified and pedagogically altered materials seemed to have difficulty developing a true self-confidence in their abilities when in fact their self-confidence should be the greatest since they have worked through challenges others did not encounter. Ease and clarity in workshop materials did give students confidence of mastery but did not equate to assimilation of concepts. Just the opposite was true: a workshop with challenges (a difficult-to-read font) was associated with higher levels of performance later.

Students using materials with the altered font showed higher levels of comprehension than those using materials with the clear font. This alteration was significant and generated a disfluency effect, which increased student focus. Students who had to struggle a bit, albeit in a superficial way, showed increased learning and retention. Our results are the first from a university business school environment (operations management core class sections) and consistent with experiments performed primarily in liberal arts domains (education and psychology). In these domains, there is strong support for the belief that this type of alteration can lead to improved retention and performance. When students struggle with material, rather than rapidly and smoothly reading through it, a cue is triggered that they may not have mastery over it (Diemand-Yauman et al. (2011)). Alter et al. (2007) found that students were less confident in their ability to answer reasoning type questions

when text was printed in a disfluent font. Alter et al. (2007) also found that students with disfluent font materials answered more questions correctly than students with materials in a fluent font. Struggling students seem to engage in deeper reasoning strategies to compensate for even a superficial leaning obstacle. The introduction of clear and obvious hindrances in materials slows students down and forces them to try harder to master concepts. Subtle obstacles, such as our implementation of planned mistakes, may be missed and disregarded by students rendering these modifications less effective.

The material with planned pedagogical mistakes did not provide a statistically significant difference in scores. The mean score for students using these materials was 0.69 versus a mean score of 0.67 for students using the base set of materials, better but not statistically significant in our experiment of 155 students. We believe the reason for this result was that our two planned pedagogical mistakes were not obvious enough for our undergraduate student subjects to observe. It is likely the majority of students simply did not notice our planned incorrect mean task time calculation and erroneous mean class grade. No additional difficulty was introduced with these materials. As Alter et al. (2007) and Dreisbach and Fischer (2011) have observed, difficult-to-read fonts seem to introduce a mental trigger or signal that tells us we must cognitively engage more deeply to understand. Kahneman (2011) refers to this as activation of “System 2” or “effortful mental” processing. It is surprising that such a superficial adjustment can produce this effect, but clearly, many researchers have performed experiments with difficult-to-read fonts and observed results showing improved comprehension. As educators, we look for creative ways to more deeply engage students in the learning process.

The search for enhanced active learning methods continues and the introduction of complexities into materials seems to have merit. The trend toward creating more active learning environments in business school curriculum necessitates deep and varied exploration of student knowledge gained. In our experiment incorporating a lighter, difficult-to-read font into materials was an alteration associated with improved learning. Student judgment in assessing what they learned was inaccurate and in fact exactly the opposite of actual performance. It should be emphasized to students that increased effort, even in a superficial way, may lead to improved comprehension and recall. Similar to a number of studies conducted in education and liberal arts domains, our best performing participant groups were the ones that had the lowest levels of confidence in material mastery. These results demonstrate that incorporating obvious challenges in core operations management content can increase student focus and retention.

Appendix A

Experimental Proportions

Material set\class section	Section 1	Section 2	Section 3	Total
Control	0.2903	0.2941	0.3810	0.3161
Pedagogical mistakes	0.4032	0.3333	0.2857	0.3484
Unusual fonts	0.3065	0.3725	0.3333	0.3355

Chi-Square Test Class Section 1

Material set\class section 1	Sample size	Hypothesized proportions
Control	18	0.3333
Pedagogical mistakes	25	0.3333
Unusual fonts	19	0.3333

$$\chi^2 = \sum_{i=1}^3 \frac{(f_i - e_i)^2}{e_i} = 1.3871$$

$$p - \text{value} = P(\chi^2 > 1.3871) = 0.4998^1$$

¹Not statistically significant, we cannot conclude the sampled proportions for class section 1 are different from the hypothesized proportions.

Chi-Square Test Class Section 2

Material set\class section 2	Sample size	Hypothesized proportions
Control	15	0.3333
Pedagogical mistakes	17	0.3333
Unusual fonts	19	0.3333

$$\chi^2 = \sum_{i=1}^3 \frac{(f_i - e_i)^2}{e_i} = 0.4706$$

$$p - \text{value} = P(\chi^2 > 0.4706) = 0.7903^2$$

²Not statistically significant, we cannot conclude the sampled proportions for class section 2 are different from the hypothesized proportions.

Chi-Square Test Class Section 3

Material set\class section 3	Sample size	Hypothesized proportions
Control	16	0.3333
Pedagogical mistakes	12	0.3333
Unusual fonts	14	0.3333

$$\chi^2 = \sum_{i=1}^3 \frac{(f_i - e_i)^2}{e_i} = 0.5714$$

$$p - \text{value} = P(\chi^2 > 0.5714) = 0.7515^3$$

³Not statistically significant, we cannot conclude the sampled proportions for class section 3 are different from the hypothesized proportions.

Appendix B Assessment Questions

Question	Bloom level
Which of the following is a distribution in which all of the probabilities are equally likely?	Low (knowledge/comprehension)
A distribution that shows outcomes and their associated probabilities or likelihoods is called what?	Low (knowledge/comprehension)
Which of the following is a popular distribution in which 3 numbers are specified? These 3 numbers completely describe the distribution.	Low (knowledge/comprehension)
When rolling a single die, what is the probability of an even number outcome (a 2, 4, or 6)?	Medium (application/analysis)
Which of the following Monte Carlo simulation results is most likely?	Medium (application/analysis)
based on a long history of buying a particular part from various suppliers, ABC Inc. believes costs can be anywhere from \$8 to \$12 per unit. The table (Table 3) shows the probabilities of the part costs, based on the historical information from purchasing. A Monte Carlo simulation is run for 100 trials. Which of the following results is not possible?	Medium (application/analysis)
Figure 3 is a picture of a beta distribution. Like the normal distribution, the beta distribution is a continuous distribution, but may be skewed to either side, which changes the likelihood of outcomes to be greater on one side versus another. Which of the following statements is true regarding the beta distribution shown in the figure?	High (synthesis/evaluation)
Which statement best describes why the MC simulation mean result may differ from the statistically calculated mean?	High (synthesis/evaluation)
What is the value or use of Monte Carlo Simulation?	High (synthesis/evaluation)
We are not sure how much money project ABC will cost. A team of business experts believes the cost could be as low as \$100,000, as high as \$200,000, but most likely \$180,000. A good distribution to model the possible project cost would be which distribution? Which of the following is the mean cost for the ABC project?	Medium (application/analysis)
Considering Figure 1, a Monte Carlo simulation was run with 700 trials, what is the probability of loss?	Medium (application/analysis)
Considering Figure 2, a Monte Carlo simulation was run with 700 trials, what is the expected or mean <i>revenue</i> ?	Medium (application/analysis)
Considering Figure 2, a Monte Carlo simulation was run with 700 trials, what is the expected or mean <i>profit</i> ?	Medium (application/analysis)

References

- Alter A (2013) The benefits of cognitive disfluency. *Current Directions Psych. Sci.* 22(6):437–442.
- Alter A, Oppenheimer D, Epley N, Eyre R (2007) Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *J. Experiment. Psych.* 136(4):569–574.
- Auster E, Wylie K (2006) Creating active learning in the classroom: A systematic approach. *J. Management Ed.* 30(2):333–353.
- Barak M, Lipson A, Lerman S (2006) Wireless laptops as means for promoting active learning in large lecture halls. *J. Res. Tech. Ed.* 38(3):245–263.
- Bjork RA (1994) Memory and metamemory considerations in the training of human beings. Metcalfe J, Shimamura AP, eds. *Metacognition: Knowing About Knowing* (MIT Press, Cambridge, MA), 185–205.
- Bloom BS (1956) *Taxonomy of Educational Objectives: The Classification of Educational Goals Book 1: Cognitive Domain* (Longman, New York).
- Bonwell CC, Sutherland TE (1996) The active learning continuum: Choosing activities to engage students in the classroom. *New Directions for Teaching Learn.* 67(3):3–16.
- Botvinick NM, Braver TS, Carter CS, Barch DM, Cohen JD (2001) Cognitive monitoring and cognitive control. *Psych. Rev.* 108:624–652.
- Carey B (2011) Come on, I thought I knew that. *The New York Times* (April 18) Health, D5.
- Castel AD, McCabe DP, Roediger HL (2007) Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bull. Rev.* 14(1):107–111.
- Coy S, Adams J (2012) Walking the talk: Continuous improvement of a quality management field exercise. *Decision Sci. J. Innovative Ed.* 10(2):223–244.
- Diemand-Yauman C, Oppenheimer D, Vaughan E (2011) Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. *Cognition* 118(1):111–115.
- Dreisbach G, Fischer R (2011) If it's hard to read...try harder! Processing fluency as signal for effort adjustments. *Psych. Res.* 75:376–383.
- Frederick S (2005) Cognitive reflection and decision making. *J. Econom. Perspect.* 19:25–42.
- Gill MJ, Swann WB Jr, Silvera DH (1998) On the genesis of confidence. *J. Personality Soc. Psych.* 75(5):1101–1114.
- Gorman MF (2008) Evaluating the integration of technology into business statistics. *INFORMS Trans. Ed.* 9(1):10–19.
- Hansen JD (2006) Using problem-based learning in accounting. *J. Ed. Bus.* 81(4):221–224.
- Hembroke H, Gay G (2003) The laptop and the lecture: The effects of multitasking in learning environments. *J. Comput. Higher Ed.* 15(1):46–64.
- Hernandez I, Preston JL (2013) Disfluency disrupts the confirmation bias. *J. Experiment. Soc. Psych.* 49(1):178–182.

- Kahneman D (2011) *Thinking, Fast and Slow* (Farrar, Straus and Giroux, New York).
- Kelley CM, Lindsay DS (1993) Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *J. Memory and Language* 32(1): 1–24.
- Kornell N, Rhodes M, Castel A, Tauber S (2011) The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psych. Sci.* 22(6):787–794.
- Lee KH (2007) Chocolate chip cookies as a teaching aid. *Amer. Statist.* 61(4):351–355.
- Leon L, Seal KC, Przasnyski ZH (2006) Captivate your students' minds: Developing interactive tutorials to support the teaching of spreadsheet modeling skills. *INFORMS Trans. Ed.* 7(1): 70–87.
- Lieberman MD, Gaunt R, Gilbert DT, Trope Y (2002) Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. *Adv. Experiment. Soc. Psych.* 34: 199–249.
- Oppenheimer D (2008) The secret life of fluency. *Trends in Cognitive Sci.* 12(6):237–241.
- Raelin J, Coghlan D (2006) Developing managers as learners and researchers: Using action learning and action research. *J. Management Ed.* 30(5):29–41.
- Rasch D, Kubinger KD, Moder K (2011) The two-sample *t* test: Pre-testing its assumptions does not pay off. *Statist. Papers* 52.1:219–231.
- Rhodes MG, Castel AD (2008) Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *J. Experiment. Psych. General* 137:615–625.
- Rienzo T, Han B (2011) Does ERP hands-on experience help students learning business process concepts?. *Decision Sci. J. Innovative Ed.* 9(2):177–207.
- Sungkhasettee VW, Friedman MC, Castel AD (2011) Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psych. Bull. Rev.* 18(5):973–978.
- Ueltschy LC (2001) An exploratory study of integrating interactive technology into the marketing curriculum. *J. Marketing Ed.* 23(1):63–72.
- Welch BL (1947) The generalization of “student’s” problem when several different population variances are involved. *Biometrika* 34(1):28–35.
- Weltman D, Whiteside M (2010) Comparing the effectiveness of traditional and active learning methods in business statistics: Convergence to the mean. *J. Statist. Ed.* 18(1).
- Yue C, Castel AD, Bjork RA (2012) When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory and Cognition* 41:229–241.