

The Generation Effect in the Context of Lyrical Censorship

BRITTANY A. GOLDMAN

MATTHEW R. KELLEY*

Lake Forest College

When people are actively involved in generating information (e.g., solving a word fragment), they tend to remember that information better than when they process the information more passively (e.g., hearing a word). This phenomenon—the generation effect—has been applied to numerous settings including the field of education (e.g., teaching, learning, mathematics) and marketing (e.g., advertising). The current study reviewed the applied generation effect literature and then explored this effect within a new applied setting—lyrical censorship. Participants listened to and shadowed an original song which contained a mixture of partially or completely censored nouns. Participants were asked to repeat every word and generate the censored words throughout the song. Results showed an ironic effect of censorship: censored items were remembered significantly better than heard items. Results and implications are discussed.

One of the classic benchmark phenomena in the human memory literature is the generation effect (Hirshman & Bjork, 1988; Mulligan & Duke, 2002; Serra & Nairne, 1993; Slamecka & Graf, 1978). The generation effect refers to the finding that people remember information better when they are actively involved in generating that information such as the word fragment *f_otball*, than when they simply read the complete word *football*. Although using word fragments is a common manipulation in the generation effect literature, researchers have employed a variety of methods, including letter transposition, word stem completion, paired-associate generation, and picture fragment completion, to name but a few (Kinjo & Snodgrass, 2000; Mulligan, 2002; Mulligan & Duke, 2002). Moreover, researchers have demonstrated robust generation effects with a range of testing procedures including explicit free recall and recognition (Burns, 1990), explicit serial recall and reconstruction of order (Kelley & Nairne, 2001; Nairne & Kelley, 2004), and implicit memory tests (Mulligan, 2002; Nicolas, Ehrlich, & Facci, 1996).

On the Generality of the Generation Effect

In an attempt to assess the generality of the genera-

tion effect, many studies have explored the influence of generation across the lifespan, from childhood through adulthood and into old age (e.g., Nicholas et al., 1996; Pesta, Sanders, & Nemeč, 1996). For instance, McFarland, Duncan, and Bruno (1983) asked children aged 7, 9, 11, and 13 to generate or read words from semantic or phonetic categories for a subsequent test (free recall, recognition, or rhyme recognition). They reported that generation ability increased substantially from age 7 to 13, with older children able to generate both semantic and phonemic categories. Clearly, children must attain a certain level of cognitive maturity before they can fully benefit from the generation effect (McFarland et al., 1983).

The majority of research on the generation effect has involved college-aged participants and has shown robust memory enhancement for generated materials (e.g., Hirshman & Bjork, 1988). However, some studies have directly manipulated the ages of adults and have shown that the generation effect is present across the

* Faculty Sponsor

For convenience, please address correspondence to Matthew Kelley, Associate Professor of Psychology, Lake Forest College, 555 N. Sheridan Rd., Lake Forest, IL 60045, kelly@lakeforest.edu

lifespan, although its size is often reduced in old age (e.g., Tacconat et al., 2006; Troyer, Häfliger, Cadieux, & Craik, 2006). For example, Tacconat, Froger, Sacher, and Isingrini (2008) compared generation abilities for younger adults (20-36 yrs) and older adults (60-81 yrs). Participants were asked to generate or read lists of strong- or weak-associate pairs and then complete a cued-recall test. They showed that the magnitude of the generation effect did not differ between younger and older adults with strong-associate lists; however, younger adults exhibited a stronger generation effect for the weak-associate lists. Despite a reduction in magnitude, generation remains a useful mnemonic device throughout the lifespan.

Importantly, the act of generation also has been shown to benefit people with a variety of physical and cognitive impairments, including sufferers of multiple sclerosis (Basso, Lowery, Ghormley, Combs, & Johnson, 2006), traumatic brain injury (Lengenfelder, Chiarvaloti, & DeLuca, 2007), dementia of the Alzheimer's type (DAT; Barrett, Crucian, Schwartz, & Heilman, 2000), as well as other memory impairments (Tailby & Haslam, 2003). For instance, Multhaup and Balota (1997) examined healthy older adults (mean age 77), very mild DAT adults (mean age 76.7), and mild DAT adults (mean age 76.1). Using a counterbalanced within-subjects design, participants completed sentence fragments in an experimenter-participant task (half of the endings were completed by the experimenter; half by the participant) and a generate-read task (half required generation; half did not). The subsequent recognition test also employed a source judgment (experimenter- or participant-generated?; read or generated?). Multhaup and Balota reported significant generation effects for all three groups of participants (11% and 15% advantages for the experimenter-participant and generate-read tasks, respectively). Although reliable for each group, the magnitude of the generation effect declined with DAT severity. Furthermore, source monitoring accuracy declined as dementia increased across the groups.

Applications of the Generation Effect

Unlike many benchmark phenomena in the memory literature, the generation effect has clear applications to everyday life outside the laboratory. In particular, the mnemonic benefits of self-generation have been studied extensively in educational settings (e.g., Boyle & Weishaar, 1997), in the marketplace (e.g., Thompson & Barnett, 1981), and in other common, real-world situations (Kelley, Goldman, Briggs, & Chambers, 2009).

Learning. Research suggests that the generation effect can facilitate student learning (e.g., DeWinstanley & Bjork, 2004). That is, when attempting to learn new information, the act of generating the material has a

strong memory advantage over simply reading the material. For instance, Foos, Mora, and Tkacz (1994) compared student-generated questions and outlines with experimenter-generated questions and outlines. On an exam conducted 2 days after receiving/generating the outline and questions, the groups that generated their own study material scored significantly higher than the groups that received the experimenter-generated study material. Clearly, students can take advantage of this mnemonic benefit in an educational setting.

Mathematics. Along with learning and teaching, the generation effect has been shown to facilitate memory for mathematics problems. For instance, McNamara and Healy (2000) explored the influence of generation with simple and difficult multiplication problems. Participants were asked to generate half of the math problems (both simple and difficult) and read the remaining half. Although a generation effect was found in both conditions, a stronger generation effect was found with the simple math problems compared to the difficult math problems (for similar results with small versus large product-size problems, see Pesta et al., 1996).

Marketing and advertising. The generation effect has also been demonstrated successfully in marketing settings. For instance, Reardon and Moore (1996) created 19-min audio segments that contained 10 advertisements, three public service announcements, and three songs. The target advertisement was for an Optimax 35-mm camera, whose brand name was presented five times to each participant. The generation opportunity came at the end of the advertisement when a voice asked, "Hey, what was the name of that new camera again?"; this line was not included in the control advertisement. Following a short distractor task, participants were given a surprise cued recall test with 56 product names, including the target product. Participants were also asked to rate their confidence level of actually hearing the product name during the audio clip. Results indicated a significant generation effect—participants in the generation condition remembered the camera name significantly better than participants in the control condition. Hence, Reardon and Moore (1996, see also, Thompson & Barnett, 1981) were able to use the generation effect to help increase the recall of the brand name, which presumably might influence sales of that product.

The generation effect has also been demonstrated with print advertisements. For example, Sengupta and Gorn (2002) used print advertisements with two experimental conditions (element-absent and element-present) to test subsequent recall for Olympus cameras. The element-absent condition depicted a bare-chested male displaying the outline of a camera and its strap,

while the element-present condition depicted the bare-chested male with a real camera and strap. After viewing several print media pictures, participants were asked to complete a product and brand questionnaire. In addition, participants were asked to list everything they could remember regarding the specific image in question (i.e., the Olympus camera). The results indicated a generation effect—enhanced recall in the element-absent condition.

Lyrical censorship. Recently, Kelley, Goldman, Briggs, and Chambers (2009) explored the generation effect within the context of lyrical censorship. In their study, participants listened to an original song that contained 30 nonoffensive nouns—15 of which were either completely-censored or partially-censored and 15 of which were not. In the partially-censored condition, the word's initial phoneme was audible (e.g., the *t* sound in *tree*), while in a completely censored condition, the whole word was omitted. Participants were asked to shadow the lyrics as they were presented—repeating every *heard* word and generating every *censored* word—while an experimenter recorded the shadowing/generating accuracy. After song presentation, participants received a short distractor task where the participants gave a rating for their opinion of the song and then completed a forced-choice recognition test with a source component (e.g., if item was from song, was it heard or censored). An ironic effect of censorship was found for generating lyrics. If the participant successfully generated the censored lyric during presentation, his/her memory was enhanced by almost 20% over the heard items. Unfortunately, the generation task proved difficult for participants as they generated only 46% and 31% of the items in the partial and complete conditions, respectively. Moreover, source memory generally was poor. In other words, participants were poor at determining whether they actually heard a word or if that word was censored. Participants more often recalled that they heard the word when, in fact, it had been censored. Source memory was measured by calculating the sum of both the censored and not censored items and then dividing the sum by all of the censored and not censored items that were “recognized.” The current study is designed to replicate and extend the work of Kelley et al. (2009).

Present Experiment

The low generation rate reported by Kelley et al. (2009) could have been due to impoverished contextual information at the time of generation. That is, the necessary contextual information did not always precede the censored items, but often occurred within the next few lines of the song. To remedy this, participants in the current experiment were allowed to listen to

the entire song once—so they could understand the complete context—before they heard it again while completing the shadowing/generating task. Following a short distractor task, participants completed a similar forced-choice recognition test as used by Kelley et al. (2009) with one exception—the source recognition task was eliminated. Specifically, participants were forced to determine whether the word was from the song (either heard or censored) or whether it was a new item (distractor). By having participants hear the entire song before shadowing/generating, we expected that (a) participants would generate a greater proportion of items than on the previous experiment and (b) that they would remember the censored items better than the heard items.

Method

Participants

Twenty-four students from an introductory psychology course at a small liberal-arts college in the Midwest earned extra-credit for their participation in this task. Participants were tested individually in sessions lasting approximately 20 minutes. Stimuli were presented and controlled in a small testing room with a desktop computer.

Materials and Design

Thirty nonoffensive, high-frequency nouns were obtained from the Kucera and Francis (1967) norms with an average frequency of 241.72 (occurrences per million written words), a range of 107-787, and an average word length of 5.38 letters with a range of 4-7. Using each of the 30 target nouns, an original song consisting of 306 words was created about the start of a beautiful relationship between a man and a woman. A male vocalist recorded this song with the use of an acoustic guitar and a keyboard. Multiple versions of the song were created, with 15 of the nouns being intact (e.g., island, hand, friend) and 15 were censored, (e.g., market, voice, family). Censoring was accomplished by erasing the word (e.g., market) so no sound occurred at that point during the song. The censored words were either completely censored (eliminated from the song leaving a silent spot), or partially censored, leaving the word's initial phoneme audible (e.g., the *t* sound in *tree*).

In order to allow each noun to be presented equally often as a censored word and as an intact word, multiple versions of the song were employed. Thus, the second version counterbalanced the 15 censored nouns with the 15 intact nouns so that each word was presented in the same quantity across the participants. Overall, 5.5% of the words in the song were censored. Another 30 nouns were matched on all of the above

characteristics (i.e., frequency and word length) to pair together the censored/noncensored nouns and to serve as distractors in the recognition test following the completion of the experiment. The average frequency for the nouns in the song was 178.72 with a range of 101-470 and an average word length of 5.81 letters with a range of 4-7.

Kelley et al. (2009) provided an item analysis which reported the generation rate and recognition accuracy for each item when censored and not censored. This analysis showed the generation effect for 24 of 30 words. Additionally, there was no effect of generation rate. In other words, both easy and hard to generate items displayed the generation effect.

A 30 s clip of the theme song from the television show *Cheers* was used in a shadowing practice session. Participants were provided with a pair of headphones in order to listen to the lyrical song.

Procedure

Participants were given a brief set of instructions (see appendix) via the computer, prior to the start of the experiment. Participants received shadowing practice prior to hearing the experimental song. In other words, each participant was asked to repeat every word aloud. After completing the 30 s shadowing practice, participants listened to the entire version of the appropriate song for their condition (about 5 min long). Following the first song presentation, participants were asked to shadow the song lyrics while an experimenter recorded their shadowing/generating accuracy. During this phase, participants were instructed to fill in the censored words to the best of their ability. After the shadowing task, participants completed a short distractor task, rating the song on various qualities (e.g., pleasantness of the music). Finally, participants completed a 60-item, forced-choice recognition test, where they were asked if the noun was an old item (i.e., heard or censored from the song) or if the noun was a new item (distractor word that was not in or censored from the song).

Results and Discussion

Overall Recognition

Overall recognition performance was analyzed with a 2 (generation type: partial vs. complete) \times 3 (stimulus type: not censored, censored, distractor) mixed-factor ANOVA (see Figure 1). Only the main effect of stimulus type reached statistical significance, $F(2, 44) = 163.13, p < .001$ (see Table 1 for complete ANOVA information). A Newman-Keuls post-hoc analysis revealed that recognition accuracy was statistically equivalent for noncensored and censored words and performance in both conditions exceeded the false alarm rate for distractor items. The ANOVA did not reveal any other significant main effects or interactions. These results are consistent with Kelley et al. (2009)—namely, the overall data failed to produce a traditional generation effect. Kelley et al. (2009) suggested that, with the overall data, the generation effect might have been masked by the fact that not all censored words were generated. To control for this, a separate *conditional* analysis was run in which data inclusion was made conditional upon successful shadowing of noncensored items as well as successful generation and shadowing of censored items.

Conditional Recognition

After hearing the complete song once, participants successfully generated 51% ($SD = 10\%$) of the partially-censored items and 29% ($SD = 9\%$) of the completely-censored items. Recognition performance on the conditional data was analyzed with a 2 (generation type: partial vs. complete) \times 3 (stimulus type: not censored, censored, distractor) mixed-factor ANOVA (see Figure 1). The ANOVA revealed a significant main effect of stimulus type, $F(2, 44) = 231.47, p < .001$. A Newman-Keuls post-hoc analysis revealed a *traditional generation effect*—recognition accuracy for generated censored items was significantly higher than the noncensored items and both means far exceeded the false alarm rate. Once again, neither the main effect of censorship

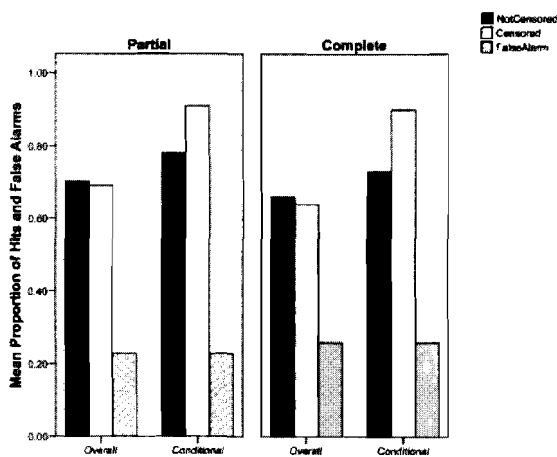
TABLE 1

ANOVA Results for Overall and Conditional Data

Overall	
Main Effect of Generation Type	$F(1, 22) = .10; p > .05$
Main Effect of Stimulus Type	$F(2, 44) = 163.13; p < .001$
Interaction	$F(2, 44) = 1.24; p > .05$
Conditional	
Main Effect of Generation Type	$F(1, 22) = .07; p > .05$
Main Effect of Stimulus Type	$F(2, 44) = 231.47; p < .001$
Interaction	$F(2, 44) = .82; p > .05$

FIGURE 1

Mean proportion of hits and false alarms as a function of stimulus type (not censored, censored, distractor), censorship type (partial, complete), and analysis type (overall, conditional).



type nor the interaction reached statistical significance. Indeed, a similar pattern of performance was apparent for both the partial (.91 > .78 > .23) and complete (.90 > .73 > .26) censorship conditions. Therefore, when censored words were generated successfully at encoding, those words were significantly more memorable than noncensored (heard) words.

The present investigation was designed to replicate and extend the findings of Kelley et al. (2009). One criticism of their original study was that generation rate was rather low (46% and 31% for the partial and complete censorship conditions, respectively). We posited that generation rate might be enhanced if participants heard the entire song because participants would understand the song's full context prior to completing the shadowing/generating task. Hearing the song twice only elicited a nominal improvement in generation rate for the partially-censored items (51% compared to 46% in previous study) and did not appear to influence generation rate for completely-censored items (29% vs. 31%). Clearly, complete censorship is a more effective technique than partial censorship. Yet, adult participants were still able to generate at least a portion of the omitted items. Despite these low generation rates, when participants successfully generated an item, they tended to remember it extremely well. Collapsed across the partial and complete conditions, generated censored items enjoyed a 15% mnemonic benefit over uncensored, heard items. Somewhat ironically, then, omitting song lyrics can actually enhance memory

for words that are censored. Censoring words did not appear to increase memory for the noncensored nouns. However, this could be tested by having a condition in which the entire song was heard and comparing performance for the heard items in a song without censorship versus the heard items in a song with censorship. To date, most generation effect literature shows the benefits are specific to the generated material.

One obvious limitation of the current study is that it used a sample of college students and censorship generally is intended for a younger audience. Because children are often impressionable, future research should focus on whether the generation effect with song lyrics occurs among youths. As described earlier, McFarland et al. (1983) showed that the generation effect in children was weak at age 7 but developed steadily until age 13, when the generation effect began to appear as for adults. Hence, one would expect even lower generation rates by young children. Moreover, many of the words that are typically censored from songs may not even be known to the children—they cannot generate what they do not know.

Another direction for future research might investigate the generation effect with regard to censorship of explicit versus nonexplicit words. Within the current study, the use of nonoffensive nouns was thought to be a tougher test of the lyrical generation effect. This is because offensive censored words contain a small, known set of items, which should make retrieval easier. If offensive words were used, memory performance might show a ceiling effect—regardless of whether the word was heard or generated. On the other hand, the nouns used in this study were derived from a huge set of all-high-frequency nouns, which in turn makes for a tougher test.

The present findings suggest that memory enhancement can occur when college-aged students were able to generate censored lyrics. This was apparent when the data was made conditional. In other words, during the encoding phase, participants showed a standard generation effect when they were able to successfully generate a censored word (i.e., listening and shadowing the song). Additionally, when participants recognized a word which was either heard or censored, they were more likely to report that word as "heard." Obviously, future research with different populations and types of words is needed before the generality of the lyrical generation effect can be assessed.

References

- Basso, M. R., Lowery, N., Ghormley, C., Combs, D., & Johnson, J. (2006). Self-generated learning in people with multiple sclerosis. *Journal of the International Neuropsychological Society, 12*, 640-648.
- Barrett, A. M., Crucian, G. P., Schwartz, R. L., & Heilman, K. M. (2000). Testing memory for self-generated items in dementia:

- Method makes a difference. *Neurology*, 54, 1258-1264.
- Boyle, J. R., & Weishaar, M. (1997). The effects of expert-generated versus student-generated cognitive organizers on the reading comprehension of students with learning disabilities. *Learning Disabilities Research & Practice*, 12, 228-235.
- Burns, D. J. (1990). The generation effect: A test between single- and multifactor theories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 1060-1067.
- DeWinstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition*, 32, 945-955.
- Foos, P. W., Mora, J. J., & Tkacz, S. (1994). Student study techniques and the generation effect. *Journal of Educational Psychology*, 86, 567-576.
- Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14, 484-494.
- Kelley, M. R., Goldman, B. A., Briggs, J. E., & Chambers, J. (2009). Ironic effects of censorship. In M. R. Kelley (Ed.), *Applied memory* (pp. 1-18). Hauppauge, NY: Nova Science.
- Kelley, M. R., & Nairne, J. S. (2001). Von Restorff revisited: Isolation, generation, and memory for order. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27, 54-66.
- Kinjo, H., & Snodgrass, J. G. (2000). Does the generation effect occur for pictures? *American Journal of Psychology*, 113, 95-121.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.
- Lengenfelder, J., Chiaravolloti, N. D., & DeLuca, J. (2007). The efficacy of the generation effect in improving new learning in a person with traumatic brain injury. *Rehabilitation Psychology*, 52, 290-296.
- McFarland, C. E., Duncan, E., & Bruno, J. M. (1983). Developmental aspects of the generation effect. *Journal of Experimental Child Psychology*, 36, 413-428.
- McNamara, D. S., & Healy, A. F. (2000). A procedural explanation of the generation effect for simple and difficult multiplication problems and answers. *Journal of Memory and Language*, 43, 652-679.
- Mulligan, N. W. (2002). The effects of generation on conceptual implicit memory. *Journal of Memory & Language*, 47, 327-342.
- Mulligan, N. W., & Duke, M. D. (2002). Positive and negative generation effects, hypermnnesia, and total recall time. *Memory & Cognition*, 30, 1044-1053.
- Multhaup, K. S., & Balota, D. A. (1997). Generation effects and source memory in healthy older adults and in adults with dementia of the Alzheimer type. *Neuropsychology*, 11, 382-391.
- Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language*, 50, 113-133.
- Nicolas, S., Ehrlich, M., & Facci, G. (1996). Implicit memory and aging: Generation effect in a word-stem completion test. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 15, 513-533.
- Pesta, B. J., Sanders, R. E., & Nemeck, R. J. (1996). Older adults' strategic superiority with mental multiplication: A generation effect assessment. *Experimental Aging Research*, 22, 155-169.
- Reardon, R., & Moore, D. J. (1996). The greater memorability of self-generated versus externally-presented product information. *Psychology & Marketing*, 13, 305-320.
- Sengupta, J., & Gorn, G. J. (2002). Absence makes the mind grow sharper: Effects of element omission on subsequent recall. *Journal of Marketing Research*, 39, 186-201.
- Serra, M., & Nairne, J. S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. *Memory & Cognition*, 21, 34-40.
- Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, 4, 592-604.
- Taconnat, L., Baudouin, A., Fay, S., Clarys, D., Vanneste, S., Tournelle, L., et al. (2006). Aging and implementation of encoding strategies in the generation of rhymes: The role of executive functions. *Neuropsychology*, 20, 658-665.
- Taconnat, L., Froger, C., Sacher, M., & Isingrini, M. (2008). Generation and associative encoding in young and old adults: The effect of the strength of association between cues and targets on a cued recall task. *Experimental Psychology*, 55, 23-30.
- Tailby, R., & Haslam, C. (2003). An investigation of errorless learning in memory-impaired patients: Improving the technique and clarifying theory. *Neuropsychologia*, 41, 1230-1240.
- Thompson, C. P., & Barnett, C. (1981). Memory for product names: The generation effect. *Bulletin of the Psychonomic Society*, 18, 241-243.
- Troyer, A. K., Häfliger, A., Cadieux, M., & Craik, F. I. M. (2006). Name and face learning in older adults: Effects of level of processing, self-generation, and intention to learn. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 61B, 67-74.