

Predicting memory benefits in the production effect: the use and misuse of self-generated distinctive cues when making judgments of learning

Alan D. Castel · Matthew G. Rhodes · Michael C. Friedman

Published online: 23 August 2012
© Psychonomic Society, Inc. 2012

Abstract The production effect occurs when reading a word aloud leads to better memory for the item, relative to words that are read silently. In the present study, we assessed the degree to which judgments of learning (JOLs) are sensitive to the production effect, to determine whether people are aware of how distinctive cues can enhance memory. If the act of saying a word aloud is used as a cue for later memorability, then JOLs should be sensitive to production. Experiment 1 demonstrated that this was the case, as participants provided higher JOLs for produced items than for those read silently. This pattern of JOLs was also evident when participants silently mouthed words (Exp. 2). In Experiment 3, participants instead made a nonunique response as the production component (saying “yes” instead of the word itself). JOLs were still higher under production, although memory performance did not differ from that in a silent condition. The results suggest that the presence of both specific and nonspecific self-generated cues is used to make metacognitive judgments, likely due to the high accessibility of this information, but that participants are not precisely aware of how distinctiveness enhances encoding and retrieval. Such findings have implications for how distinctiveness is perceived by learners and for what cues would appropriately be incorporated when predicting future memory performance.

Keywords Metamemory · Memory · Production effect · Judgments of learning · Distinctiveness

A. D. Castel (✉) · M. C. Friedman
Department of Psychology, University of California, Los Angeles,
1285 Franz Hall, Box 951563, Los Angeles,
CA 90095-1563, USA
e-mail: castel@ucla.edu

M. G. Rhodes
Department of Psychology, Colorado State University,
Fort Collins, CO, USA

It is often important to accurately assess one’s own memory performance, a skill that has significant applied and theoretical implications for learning. To better understand how people monitor memory performance, researchers have examined predictions of memory performance and how the predictions then relate to actual memory performance. One way to assess this question is to ask people to make a judgment of learning (JOL) when studying information, indicating the likelihood that they will remember some bit of information on a future memory test. These predictions can then be compared with actual memory performance in order to assess the accuracy of the JOLs.

A host of work has suggested that JOLs are often at least moderately predictive of memory performance (e.g., Arbuckle & Cuddy, 1969; Nelson & Dunlosky 1991; Rhodes & Tauber, 2011b; Underwood, 1966), but it remains unclear how people incorporate any of the number of cues that are available when making metacognitive judgments. Indeed, several lines of research have demonstrated substantial discrepancies between predicted and actual memory performance that may reflect inappropriate weighting of the cues available to the rememberer (e.g., Benjamin, Bjork, & Schwartz, 1998; Kelley & Jacoby, 1996b; Koriat & Bjork, 2006; Koriat & Ma’ayan, 2005; Rhodes & Castel, 2008, 2009; Rhodes & Tauber, 2011a). For example, Benjamin et al. recorded participants’ latencies to answer general knowledge questions. Immediately after providing an answer, the participants predicted the likelihood that they would later remember that answer when given the opportunity for recall. The results showed that answers with the shortest retrieval latencies were given the highest JOLs; however, the opposite pattern was apparent for recall, as items with the longest latencies were the ones most likely to be recalled. Such discrepancies between predictions and memory performance provide important information regarding the cues that people use when making JOLs. For

example, people may misattribute ease of processing at study or retrieval to actual learning (Miele, Finn, & Molden, 2011), resulting in predictions of memory performance that deviate from the actual performance (e.g., Castel, McCabe, & Roediger, 2007; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Kornell, Rhodes, Castel, & Tauber, 2011; Matvey, Dunlosky, & Guttentag, 2001; Rhodes & Castel, 2008; Sungkhasettee, Friedman, & Castel, 2011; Zechmeister & Shaughnessy, 1980).

However, there are also instances when the perceived ease of processing leads to accurate JOLs (e.g., Arbuckle & Cuddy, 1969; Begg, Snider, Foley, & Goddard, 1989; Koriat & Ma'ayan, 2005). For example, participants' JOLs appear to be sensitive to the generation effect, which is the finding that information is better remembered when it is generated from cues rather than simply read (e.g., generating "fast" in response to the cue "rapid-f__" rather than simply reading "rapid-fast"; deWinstanley & Bjork, 2004; Matvey et al., 2001; Slamecka & Graf, 1978). For example, Begg, Vinski, Frankovich, and Holgate (1991) had participants generate some items at study and read others, and after each item, the participants made a prediction regarding their later memory performance. Begg et al. (1991) found that the participants deemed generated items to be more memorable than items that were read, and that this was consistent with actual memory performance, despite generation being associated with less ease of processing (see also deWinstanley & Bjork, 2004; Matvey et al., 2001; Mazzoni & Nelson, 1995).

It is unclear whether generation enhanced metacognitive awareness (e.g., by providing more diagnostic information about future memory performance) or whether people believe that any form of generation enhances memory. We examined this question in the present study by investigating the production effect. The *production effect* refers to the finding that simply pronouncing a word aloud, as compared to silently reading a word, leads to a sizable benefit in retention for the produced words (Conway & Gathercole, 1987; Hopkins & Edwards, 1972). These memorial benefits appear to accrue because producing words makes them more distinct (and more memorable) than silently read words (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; see also Forrin, MacLeod, & Ozubko, [in press](#); Hourihan & MacLeod, 2008). In a series of experiments, MacLeod et al. found robust production effects when participants said words aloud, and also when participants mouthed words silently, relative to reading words silently. However, the production effect was eliminated when participants made a nonunique response (i.e., saying "yes" for the produced words rather than saying the actual word), providing support for the notion that distinctiveness leads to memory enhancement for the produced words (see also Ozubko & MacLeod, 2010).

Are participants' memory predictions sensitive to the distinctiveness wrought by the production effect? More specifically, does production provide participants with access to more diagnostic memorial information, or are participants' JOLs sensitive to any form of production, regardless of whether it influences memory? A *distinctiveness hypothesis* would predict that production facilitates access to diagnostic memorial information. Thus, according to this hypothesis, participants should be sensitive to production when that production enhances memory performance by providing distinctive information. By extension, the distinctiveness hypothesis predicts that participants' JOLs should be insensitive to production when it does not influence memory performance, as when the production cues are not distinctive. An alternative hypothesis suggests that participants' judgments are not sensitive to distinctiveness, but to production per se. This *self-generation hypothesis* predicts that participants will deem produced items to be more memorable than items that are not produced, regardless of whether production actually facilitates memory performance.

We tested these competing hypotheses in three experiments. In Experiment 1, we sought to replicate the production effect found by MacLeod et al. (2010) using a free-recall task (cf. MacLeod, 2011) and also solicited JOLs for each item. In Experiment 2, we examined a more subtle form of production by having participants mouth words rather than read them aloud. On the basis of prior work on the generation effect and metamemory (e.g., Begg et al., 1991; deWinstanley & Bjork, 2004; Matvey et al., 2001), participants in Experiments 1 and 2 might deem produced words to be more memorable than words that were not produced (i.e., were read). Both of the accounts that we examined could accommodate such findings. That is, a memorial benefit following production could reflect access to diagnostic information (i.e., the *distinctiveness account*) or a simple rule that any produced item is more memorable than an item that was not produced (i.e., the *self-generation account*).

Thus, in Experiment 3, we sought to adjudicate between these accounts by investigating JOLs following a nonunique response. Specifically, MacLeod et al. (2010) observed that making a nonspecific response (saying "yes" for certain words), rather than saying the actual word, entirely eliminated the production effect. We adopted this procedure in Experiment 3 by requiring participants to make a nonunique response to half of the words that they studied (i.e., saying "yes" rather than producing the actual word). According to a distinctiveness account, participants should not regard produced words as more memorable under these conditions, as making a nonunique response does not generate any distinctive information. In contrast, the self-generation account predicts that participants would continue to regard produced

items as more memorable than nonproduced items, even when the response was nonunique. Such data would suggest that participants may not be basing JOLs on diagnostic cues regarding later recall, such as item-specific processing and distinctiveness, but rather on a more general, theory-based notion that producing *any* cue during encoding can enhance learning, even when this does not impact later recall.

Experiment 1

In Experiment 1, participants studied a list of words and, for each word, were given a cue to either say the word aloud or read the word silently (for a similar procedure, see MacLeod et al., 2010). Following the presentation of each word, participants made a JOL indicating the likelihood that they would be able to remember the item on a future test. Consistent with prior work (e.g., Conway & Gathercole, 1987; Gathercole & Conway, 1988; Hopkins & Edwards, 1972; MacDonald & MacLeod, 1998; MacLeod et al., 2010), we anticipated that participants would be more likely to recall produced relative to silent words. However, it was unclear whether JOLs would likewise be higher for produced than for silent words, as has been found for the generation effect (e.g., Begg et al., 1991).

Method

Participants A group of 20 undergraduate students at the University of California, Los Angeles, between the ages of 18 and 26 participated in the experiment in exchange for course credit.

Materials The materials consisted of a set of 40 common, six-letter nouns (e.g., *summer*, *office*, *dinner*) that had an everyday occurrence of approximately 35 per million, according to the Thorndike–Lorge count (Thorndike & Lorge, 1944), and that had been used in prior production effect experiments (see MacDonald & MacLeod, 1998, Appendix 1).

Procedure Participants were told that they would be shown a series of words to remember for a later memory test. They were instructed that if a word was presented in one font color (e.g., blue), they should read the word aloud, whereas if the word was presented in another (white), they should read the word silently. (The color of the cue to read the word was counterbalanced, such that half of the participants were told to read aloud words in blue and half to read aloud words in white.) Immediately after each word was presented, the participants were given up to 4 s to rate the likelihood that they would later be able to recall the word on a memory test,

on a scale from 0 % (*not likely at all*) to 100 % (*very likely*). Participants completed two practice trials and were asked whether they had any questions before they began the experiment.

The words were presented one at a time for 4 s in 44-point Arial font on a black background, after which the participants were prompted to make their JOLs. The experimenter recorded whether the participant accurately produced the word, as well as the JOL for each word. Immediately following the presentation of the list, participants engaged in a 30-s distractor task, consisting of counting backward from a prespecified three-digit number. Participants were then given 1 min to verbally free recall any of the words from the list, with the responses recorded by the experimenter. The procedure was then repeated with a second list of 20 words (with the words counterbalanced across lists and conditions). We included a second list to determine whether performance would change with a second study–test trial of new words. However, this manipulation did not interact reliably with any of the variables examined. Thus, in the interest of brevity, we collapsed data across the first and second lists for Experiment 1 and the subsequent experiments.

Results and discussion

Figure 1 shows the mean JOLs and mean recall performance from Experiment 1. Planned comparisons indicated that JOLs in the production condition ($M = 50.13$, $SE = 4.02$) were reliably greater than JOLs in the silent condition ($M = 35.04$, $SE = 3.52$), $F(1, 19) = 35.88$, $p = .001$, $\eta_p^2 = .65$. Likewise, recall was reliably better in the production condition ($M = 36.25$, $SE = 3.61$) than in the silent condition ($M = 21.25$, $SE = 2.59$), $F(1, 19) = 10.82$, $p = .004$, $\eta_p^2 = .36$. Thus, consistent with prior work (MacLeod et al., 2010), participants were more likely to recall items that had been

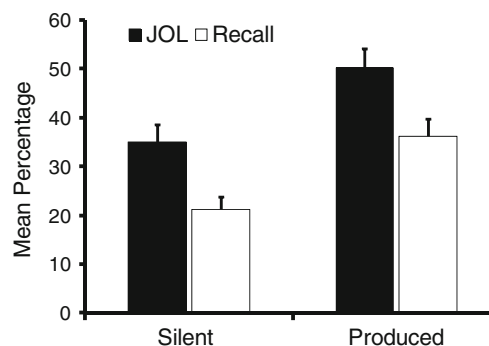


Fig. 1 Mean judgments of learning (JOLs) and mean percentages recalled as a function of silent reading and production in Experiment 1. Errors bars reflect the standard errors of the means

produced than items that had been read silently. As well, participants' JOLs were sensitive to this memorial advantage.

While our primary hypotheses were contingent on analyses of the magnitudes of JOLs and of the corresponding recall performance (i.e., absolute accuracy), for completeness we report gamma correlations between recall and the JOLs for each condition, in order to provide a measure of *relative accuracy* (i.e., the degree to which JOLs discriminated between items that were or were not later remembered). We first subjected each gamma correlation to one-sample *t* tests to determine whether the value reliably differed from zero. Collapsed across the production and silent conditions, the resulting gamma correlation ($G = .30$, $SE = .07$) reliably differed from zero, $t(19) = 4.13$, $p = .001$. The gamma correlation for the production condition ($G = .24$, $SE = .08$) differed from zero, $t(19) = 2.89$, $p = .009$, as did the gamma correlation for the silent condition ($G = .23$, $SE = .09$), $t(19) = 2.59$, $p = .018$. A follow-up test showed that the mean gamma correlations did not differ between the production and silent conditions, $F < 1$. Thus, participants were equally effective at assigning higher JOLs to items that they recalled, relative to items that were not recalled, regardless of whether the items had been produced aloud or read silently.

Experiment 2

Experiment 1 demonstrated that producing words enhanced memory, which is consistent with several previous studies documenting the production effect (MacLeod, 2011; MacLeod et al., 2010). Of greater interest, participants' JOLs were higher for produced items than for items read silently, suggesting that participants were sensitive to the enhancement in retention wrought by production. We attempted to replicate and extend these findings in Experiment 2 by using a more subtle manipulation of production. Specifically, when asked to produce a word, participants were instructed to mouth the word silently rather than say the word aloud (see Conway & Gathercole, 1987; MacLeod et al., 2010, for similar procedures). Prior work (e.g., MacLeod et al., 2010) had suggested that a production effect remains even when the production is silent rather than spoken aloud. If participants are sensitive to the distinctive act of producing a word, even when this production is silent, then JOLs should be higher for the mouthed ("silently produced") words than for the words that were read silently, much as in Experiment 1. In contrast, if participants' JOLs are sensitive to the sound produced by reading a word, then JOLs should be similar to those in the silent-reading condition, given that participants would no longer say a word aloud. With the exception of requiring participants to mouth

words rather than say them aloud, all other aspects of Experiment 2 were identical to those of Experiment 1.

Method

Participants A group of 32 undergraduate students at the University of California, Los Angeles, between the ages of 18 and 26 participated in the experiment in exchange for course credit. Two of the participants were excluded for failure to successfully mouth the words.

Procedure Experiment 2 was identical to Experiment 1, with one exception: Specifically, the participants were instructed to mouth words rather than say them aloud. Participants were given instructions on how to mouth the words, as well as two practice trials, and the experimenter observed each participant to ensure that he or she did mouth the target words.

Results and discussion

The mean recall and JOLs for the silent and produced conditions are presented in Fig. 2. Planned comparisons showed that JOLs in the production condition ($M = 49.32$, $SE = 2.94$) were reliably greater than JOLs in the silent condition ($M = 43.11$, $SE = 3.20$), $F(1, 31) = 21.30$, $p = .001$, $\eta_p^2 = .41$. However, recall was numerically, but not reliably, greater in the production condition ($M = 31.56$, $SE = 1.82$) than in the silent condition ($M = 27.34$, $SE = 2.08$), $F(1, 31) = 2.28$, $p = .141$, $\eta_p^2 = .07$. Thus, while recall only showed a trend in terms of benefits for produced words, the JOLs were significantly greater for produced than for silent items. We note that the lack of a significant production effect contrasts with prior studies that have shown a robust production effect when words were

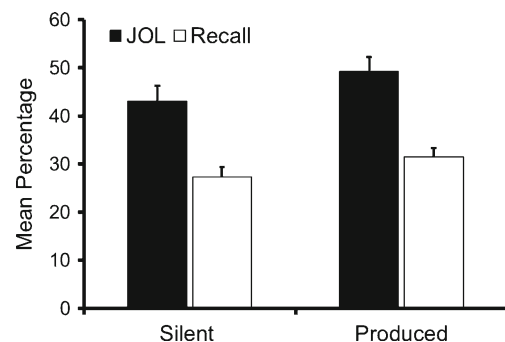


Fig. 2 Mean judgments of learning (JOLs) and mean percentages recalled as a function of silent reading and production (mouthing the words) in Experiment 2. Errors bars reflect the standard errors of the means

mouthed, although these studies typically used a recognition test (e.g., Conway & Gathercole, 1987; Forrin et al., *in press*; MacLeod et al., 2010). However, there was a trend toward a production effect, and critically for the present purposes, JOLs were significantly greater for the mouthed than for the silently read words, even when the memorial benefits of production were relatively small.

As in Experiment 1, we here report gamma correlations between recall and JOLs for each condition. Collapsed across the production and silent conditions, the resulting gamma correlation ($G = .39$, $SE = .03$) reliably differed from zero, $t(31) = 11.69$, $p < .001$. The gamma correlation for the production condition ($G = .35$, $SE = .05$) differed from zero, $t(31) = 6.39$, $p < .001$, as did the gamma correlation for the silent condition ($G = .44$, $SE = .05$), $t(31) = 8.34$, $p < .001$. A follow-up test showed that the mean gamma correlations did not differ between the production and silent conditions, $F(1, 31) = 1.19$, $p = .283$, $\eta_p^2 = .04$. Thus, relative accuracy did not differ for the mouthed and silently read items.

Experiment 3

Experiment 2 demonstrated that mouthing words led to higher JOLs than did silent reading. However, were participants' JOLs sensitive to activity that is unique to each word, or were they instead sensitive to production, because doing something, anything, leads to the belief that memory will be enhanced? As noted previously, the distinctiveness hypothesis predicts that JOLs are sensitive to the distinctive processing that is associated with the production of an item-specific cue, such that learners will be able to use this cue to generate more accurate JOLs. In contrast, the self-generation hypothesis suggests that JOLs are sensitive to any form of production. A corollary of the self-generation hypothesis is that participants' JOLs should favor produced items, regardless of whether memorial benefits are evident. The participants in Experiments 1 and 2 provided higher JOLs for produced items relative to items that were silently read, but those experiments also showed at least a trend for better memory for produced items, consistent with either account.

We attempted a stronger test of these accounts in Experiment 3 by having participants produce a nonunique response for each word. That is, instead of actually producing the word, as was done in Experiment 1, participants were instructed to say "yes" for half of the words, while silently reading the other words. Using a similar procedure, MacLeod et al. (2010) reported that this eliminated the production effect: Both words for which participants were instructed to say "yes" and words read silently led to similar levels of later memory performance. In the present

experiment, if participants' JOLs were sensitive to simply producing any sound, regardless of whether it benefited memory (i.e., the self-generation account), then JOLs should be higher for produced than for silent words. However, if participants were sensitive to the *item-specific* activities that enhance memory (i.e., the distinctiveness account), then JOLs should be equivalent between the conditions, given that producing a nonunique response should not have memorial benefits. Thus, the results from Experiment 3 would indicate whether JOLs are sensitive to distinctive item-specific processing or simply to producing any response during encoding, even if that response does not enhance memory performance.

Method

Participants A group of 20 undergraduate students at the University of California, Los Angeles, between the ages of 18 and 26 participated in the experiment in exchange for course credit.

Procedure The procedure was similar to that of Experiment 1, with the exception that for half of the words, participants were instructed to read the words silently, whereas for the other half they should say "yes" when presented with the word (for a similar procedure, see MacLeod et al., 2010).

Results and discussion

The mean recall and JOLs for the silent and produced conditions are presented in Fig. 3. Planned comparisons showed that JOLs in the production condition ($M = 54.25$, $SE = 2.43$) were reliably greater than JOLs in the silent condition ($M = 49.80$, $SE = 2.99$), $F(1, 19) = 14.09$, $p = .001$, $\eta_p^2 = .43$. However, recall was essentially equivalent

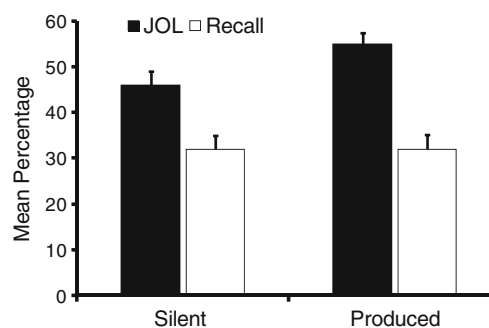


Fig. 3 Mean judgments of learning (JOLs) and mean percentages recalled as a function of silent reading and nonunique production (saying "yes") in Experiment 3. Errors bars reflect the standard errors of the means

between the production condition ($M = 33.50$, $SE = 3.10$) and the silent condition ($M = 33.00$, $SE = 2.91$), $F < 1$. Thus, it appears that participants' JOLs were sensitive to any type of production activity, even one that had no bearing on memory performance.

As in the prior experiments, we also examined measures of relative accuracy. Collapsed across the production and silent conditions, the resulting gamma correlation ($G = .30$, $SE = .06$) reliably differed from zero, $t(19) = 5.03$, $p < .001$. The gamma correlation for the production condition ($G = .31$, $SE = .06$) differed from zero, $t(19) = 4.96$, $p < .001$, as did the gamma correlation for the silent condition ($G = .32$, $SE = .09$), $t(19) = 3.53$, $p = .002$. A follow-up test showed that the mean gamma correlations did not differ between the production and silent conditions, $F < 1$. Thus, as in the prior experiments, relative accuracy did not differ on the basis of the study activity.

General discussion

In the present study, we assessed the degree to which judgments of learning (JOLs) were sensitive to the production effect in order to determine what cues guide (and potentially bias) metacognitive judgments. If the act of saying a word aloud is used as a cue for later memorability, then JOLs should be sensitive to the memory benefits afforded by production.

In Experiment 1, JOLs were greater for produced words, relative to silently read words, similar to the influence of production on recall. A similar finding was observed in Experiment 2, with JOLs being greater for mouthing words than for silent reading.

Such findings accord with two possible accounts of metacognition for produced items. That is, a distinctiveness account suggests that engaging in distinctive production provides privileged access to memory and enhances metacognition, leading participants to provide JOLs that are sensitive to the benefits of production. Conversely, a self-generation account suggests that participants deem any item associated with production to be more memorable than items not associated with production, likewise leading to elevated JOLs for produced items. Thus, in Experiment 3, we attempted to adjudicate among these accounts by having participants make a nonunique response as the production component (saying "yes" instead of the word itself). On the basis of prior work (MacLeod et al., 2010), we anticipated that such nonunique production would not lead to memory benefits for produced items. Indeed, memory was essentially equivalent for produced and silent items under these conditions. However, JOLs were still higher for items given a nonunique production, as participants predicted an "illusory" production effect that did not exist. Thus, we conclude

that while JOLs may be sensitive to production, they are not sensitive to the distinctiveness that gives rise to the production effect.

The memorial benefits of production may reflect more distinctive processing of produced words than of words read silently (Forrin et al., *in press*; MacLeod et al., 2010; Ozubko & MacLeod, 2010). Consistent with a distinctiveness account, Ozubko and MacLeod reported that production did not facilitate performance on a list discrimination task if participants were exposed to an intervening list that was also read aloud. If production simply enhanced the strength of an item, memorial benefits should be apparent, regardless of the nature of an intervening list. In addition, the production effect is strongest when participants themselves produce the words, as opposed to when another person produces them (MacLeod, 2011). Our experiments suggest that participants are largely unaware of the benefits of distinctiveness and may believe that *any* activity benefits memory, although it would also be important to test this assertion using other forms of production (e.g., whispering vs. silent reading; see Forrin et al., *in press*) and other types of memory tests, such as recognition (see also Ozubko, Gopie, & MacLeod, 2012). This is an important finding, given that there are many other instances in which distinctiveness can improve memory performance (see Hunt, 2006, for a review), as well as reduce memory illusions (e.g., Dodson & Schacter, 2001; Gallo, Perlmutter, Moore, & Schacter, 2008; Schacter & Wiseman, 2006). Because distinctiveness is such a potent cue to enhance memory and to reduce errors, it is critical for learners to understand the conditions in which distinctiveness can and cannot enhance memory. The present findings suggest that learners may have a general theory about the memorial benefits of production, but not a more precise awareness that production needs to be item-specific in order to enhance memory.

One reason that participants' JOLs were sensitive to production in all of the present experiments (even when the production effect did not exist, in Exp. 3) is that JOLs are highly sensitive to the availability of cues (see Koriat, 1997). Indeed, readily available cues, such as active versus passive processing in the generation effect (e.g., Begg et al., 1991) or direct cues to remember or forget items (e.g., Friedman & Castel, 2011), are often diagnostic of later recall. As these cues are highly accessible during learning, participants readily incorporate them into their JOLs, even when such cues are not indicative of later recall (e.g., font size; Rhodes & Castel, 2008). The present study, particularly Experiment 3, suggests that participants may rely on a more theory-based approach (see also Koriat & Bjork, 2006), regarding production of any kind as advantageous for memory. By extension, participants may deem any self-generated information produced at encoding as being more memorable than information that was not generated. This

may reflect a form of “adult-egocentrism” (Kelley & Jacoby, 1996a), in that self-generated subjective experiences elicited by to-be-remembered information may be regarded as a better basis for metacognitive judgments than are more analytical assessments of the precise nature of learning.

Overall, the present findings bear on the processes involved in making metacognitive judgments. Our data suggest that while participants may be aware of the memorial benefits of the production effect under some circumstances, they are not sensitive to the mechanisms that govern the effect. Instead, participants may deem any form of production to unilaterally enhance retention, and future research may shed light on whether this pattern of results emerges with any form of production or for other types of memory tests. While it appears that participants may adopt an overly broad notion that self-generation can enhance learning, such an approach does not adequately incorporate the influence of item-specific processing that is critical for the production effect. The present findings suggest that distinctiveness is not incorporated when people make JOLs for produced items. Consequently, future research should examine how participants learn to effectively monitor the benefits of production in a more complex learning task, such as second-language learning or learning from text (e.g., Ozubko, Hourihan, & MacLeod, *in press*), and how control operations, such as choosing to restudy information that was not initially produced, could also enhance learning. It will also be important to examine developmental changes in how distinctive cues are used appropriately, such as when children are learning and reading new vocabulary and when older adults try to remember important information (e.g., Castel, McGillivray & Friedman, 2012; Lin & MacLeod, 2012). In general, research of this type can provide insight regarding the cues that are used when making metacognitive judgments, and it could also facilitate enhanced learning and awareness of how certain types of processing can lead to memory benefits.

Author note Michael V. Roberts was instrumental in helping to design the experiments, as well as in collecting and analyzing portions of the data. Greg Hanson and Vu Nguyen helped with data collection. We sincerely thank Bob Bjork, John Dunlosky, and David McCabe for insightful comments during early stages of this research, and Colin M. MacLeod for sharing additional information about his research on this topic.

References

- Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, *81*, 126–131. doi:10.1037/h0027455
- Begg, I., Snider, A., Foley, F., & Goddard, R. (1989). The generation effect is no artifact: Generating makes words distinctive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 977–989. doi:10.1037/0278-7393.15.5.977
- Begg, I., Vinski, E., Frankovich, L., & Holgate, B. (1991). Generating makes words memorable, but so does effective reading. *Memory & Cognition*, *19*, 487–497. doi:10.3758/BF03199571
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metacognitive index. *Journal of Experimental Psychology: General*, *127*, 55–68. doi:10.1037/0096-3445.127.1.55
- Castel, A. D., McCabe, D. P., & III Roediger, H. L. (2007). Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bulletin & Review*, *14*, 107–111. doi:10.3758/BF03194036
- Castel, A. D., McGillivray, S., & Friedman, M. C. (2012). Metamemory and memory efficiency in older adults: Learning about the benefits of priority processing and value-directed remembering. In M. Naveh-Benjamin & N. Ohta (Eds.), *Memory and aging: Current issues and future directions* (pp. 245–270). New York: Psychology Press.
- Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *26*, 341–361. doi:10.1016/0749-596X(87)90118-5
- deWinstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition*, *32*, 945–955. doi:10.3758/BF03196872
- Dodson, C. S., & Schacter, D. L. (2001). “If I had said it I would have remembered it”: Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*, 155–161. doi:10.3758/BF03196152
- Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (*in press*). Widening the boundaries of the production effect. *Memory & Cognition*. doi:10.3758/s13421-012-0210-8
- Friedman, M. C., & Castel, A. D. (2011). Are we aware of our ability to forget? Metacognitive predictions of directed forgetting. *Memory & Cognition*, *39*, 1448–1456. doi:10.3758/s13421-011-0115-y
- Gallo, D. A., Perlmutter, D. H., Moore, C. D., & Schacter, D. L. (2008). Distinctive encoding reduces the Jacoby–Whitehouse illusion. *Memory & Cognition*, *36*, 461–466. doi:10.3758/MC.36.2.461
- Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & Cognition*, *16*, 110–119. doi:10.3758/BF03213478
- Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue utilized for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 22–34. doi:10.1037/0278-7393.29.1.22
- Hopkins, R. H., & Edwards, R. E. (1972). Pronunciation effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *11*, 534–537. doi:10.1016/S0022-5371(72)80036-7
- Hourihan, K. L., & MacLeod, C. M. (2008). Directed forgetting meets the production effect: Distinctive processing is resistant to intentional forgetting. *Canadian Journal of Experimental Psychology*, *62*, 242–246. doi:10.1037/1196-1961.62.4.242
- Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3–25). New York: Oxford University Press.
- Kelley, C. M., & Jacoby, L. L. (1996a). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, *35*, 157–175.
- Kelley, C. M., & Jacoby, L. L. (1996b). Memory attributions: Remembering, knowing, and feeling of knowing. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 287–307). Mahwah: Erlbaum.
- Koriat, A. (1997). Monitoring one’s own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of*

- Experimental Psychology. General*, 126, 349–370. doi:10.1037/0096-3445.126.4.349
- Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition*, 34, 959–972. doi:10.3758/BF03193244
- Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, 52, 478–492. doi:10.1016/j.jml.2005.01.001
- Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, 22, 787–794. doi:10.1177/0956797611407929
- Lin, O. Y. H., & MacLeod, C. M. (2012). Aging and the production effect: A test of the distinctiveness account. *Canadian Journal of Experimental Psychology*. doi:10.1037/a0028309
- MacDonald, P. A., & MacLeod, C. M. (1998). The influence of attention at encoding on direct and indirect remembering. *Acta Psychologica*, 98, 291–310. doi:10.1016/S0001-6918(97)00047-4
- MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review*, 18, 1197–1202. doi:10.3758/s13423-011-0168-8
- MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 671–685. doi:10.1037/a0018785
- Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, 29, 222–233. doi:10.3758/BF03194916
- Mazzoni, G., & Nelson, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1263–1274. doi:10.1037/0278-7393.21.5.1263
- Miele, D. B., Finn, B., & Molden, D. C. (2011). Does easily learned mean easily remembered? It depends on your beliefs about intelligence. *Psychological Science*, 22, 320–324. doi:10.1177/0956797610397954
- Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect. *Psychological Science*, 2, 267–270.
- Ozubko, J. D., Gopie, N., & MacLeod, C. M. (2012). Production benefits both recollection and familiarity. *Memory & Cognition*, 40, 326–338. doi:10.3758/s13421-011-0165-1
- Ozubko, J. D., Hourihan, K. L., & MacLeod, C. M. (in press). Production benefits learning: The production effect endures and benefits memory for text. *Memory*. doi:10.1080/09658211.2012.699070
- Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1543–1547. doi:10.1037/a0020604
- Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, 16, 550–554. doi:10.3758/PBR.16.3.550
- Rhodes, M. G., & Tauber, S. K. (2011a). Eliminating the delayed JOL effect: The influence of the veracity of retrieved information on metacognitive accuracy. *Memory*, 19, 853–870.
- Rhodes, M. G., & Tauber, S. K. (2011b). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, 137, 131–148. doi:10.1037/a0021705
- Schacter, D. L., & Wiseman, A. L. (2006). Reducing memory errors: The distinctiveness heuristic. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 89–107). New York: Oxford University Press.
- Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 592–604. doi:10.1037/0278-7393.4.6.592
- 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
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University, Teachers College.
- Underwood, B. J. (1966). Individual and group predictions of item difficulty for free learning. *Journal of Experimental Psychology*, 71, 673–679. doi:10.1037/h0023107
- Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, 15, 41–44.