Journal of Memory and Language 109 (2019) 104050

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Wait for it…performance anticipation reduces recognition memory☆

Noah D. Forrin*, Brandon C.W. Ralph, Navi K. Dhaliwal, Daniel Smilek, Colin M. MacLeod

University of Waterloo, Canada

ARTICLE INFO

Keywords:
Episodic memory
Performance anticipation
Production effect

ABSTRACT

Conventional wisdom suggests that there is an encoding decrement prior to performing in front of others. We hypothesized that this pre-performance memory deficit—the next-in-line effect (Brenner, 1973)—should also occur in the context of mixed-list memory experiments where one of the conditions requires performance. As the testing ground for this prediction, we used the production effect (i.e., enhanced memory for words that are read aloud vs. silently). Specifically, we examined whether performance anticipation imposes a memorial cost on silent items studied in a mixed list (among "performed" aloud items) relative to a pure-silent list. Experiment 1 established this mixed-list cost in recognition (replicating Bodner, Taikh, & Fawcett, 2014). In Experiments 2 and 3, providing foreknowledge of the task to be performed on upcoming study items—thereby allowing participants to see when they would have to read aloud—led to diminished memory for silent items that were studied immediately before aloud items. In Experiment 4, in the absence of an experimenter, the pre-performance cost to silent items was non-significant (with Bayesian evidence supporting the null), consistent with the notion that the presence of the experimenter (a social factor) contributed to performance anticipation. Taken together, these results imply that performance anticipation drives the mixed-list cost of production shown by the silent items (and may explain costs observed in other memory research). Performance anticipation may reduce memory for pre-performance information by diverting attention away from that information.

Introduction

People often report that their memory is poor for information presented prior to their own public performance. For example, a researcher presenting at a conference may remember relatively few details from talks that immediately preceded their own. Similarly, co-workers taking turns introducing themselves may struggle to remember their colleagues' names—especially those names said shortly before their own introduction. Indeed, across numerous professional and social settings, the anticipation of one's own public performance seems to impose a cost on memory for preceding information.

The next-in-line effect

In Brenner's (1973) seminal investigation of this phenomenon, which he referred to as *the next-in-line effect*, he had 22 participants sit in pairs around a table. One member of each pair was a "reader" and the other member was a "listener." The 22 readers took turns reading words aloud to the group; listeners only had to listen. In this way, the readers would presumably experience performance anticipation because they were aware of the position in which they would have to read and thus when their turn was drawing near. The listeners, conversely, would never experience performance anticipation. After words had been read around the table twice, both readers and listeners recalled as many words as they could.

Brenner's (1973) work produced two important observations. First, readers had better recall for words that they themselves read aloud relative to the words that they heard others read, which Brenner attributed to advantages from additional motor, visual, and auditory cues to recall. Second, Brenner observed a *next-in-line-effect*, whereby readers had diminished recall for words read by other readers shortly before their own turn (starting three words prior). The listener participants, conversely, had consistent recall performance prior to their partner reading aloud. Brenner argued that readers were anticipating their performance as their own turn drew near, and that this competing demand on their attention imposed a cost on memory for the passively heard words. The next-in-line effect has been replicated in free recall (e.g., Bond & Kirkpatrick, 1982; Bond & Omar, 1990; Brown & Oxman,

☆ This research was supported by Discovery Grant A7459 from the Natural Sciences and Engineering Research Council of Canada (NSERC). We thank Safa Bajwa, Kaitlyn Crook, Tenaaz Irani, Seth Mahon, James Siklos-Whillans, and Pearl Zaki for their assistance in collecting the data. All programs used in this research, as well as the raw data files, are publicly available on the Open Science Framework (https://osf.io/z689c).

* Corresponding author at: Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada.
E-mail addresses: nforrin@uwaterloo.ca, cmacleod@uwaterloo.ca (N.D. Forrin).

https://doi.org/10.1016/j.jml.2019.104050
Received 1 August 2018; Received in revised form 14 August 2019; Accepted 17 August 2019
Available online 22 August 2019
0749-596X/

1978; Innes, 1982; Walker & Orr, 1976), in cued-recall (Bond, 1985), and in recognition (Bond, Pitre, & van Leeuwen, 1991).

The production effect

The substantial memory boost that Brenner (1973) found for the words that participants read aloud dovetails with our own more recent research on *the production effect* (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; see MacLeod & Bodner, 2017, for a brief review). The production effect is the finding that memory is better for words that are produced (e.g., read aloud) relative to those that are not produced (e.g., read silently). For example, individuals have better memory for words that they have read aloud compared to words that they have heard others read which, in turn, are better remembered than words that they have read silently (Forrin & MacLeod, 2018; MacLeod, 2011). The explanation for the production effect that has garnered the most empirical support is the distinctiveness account (Conway & Gathercole, 1987; MacLeod et al., 2010; see MacLeod & Bodner, 2017, for a review). Under this account, words that are read aloud at study benefit from *distinctive processing* (see Hunt, 2006, 2013) because they undergo additional encoding that results in them standing out against a "baseline" established by the control, silent reading condition. For example, in a typical mixed-list production experiment, reading aloud involves at least three distinct processes—articulation, audition, and self-reference—compared to reading silently (Forrin & MacLeod, 2018). Similar to Brenner (1973) claim that reading aloud gives participants additional cues to recall, MacLeod et al. (2010) argued that participants use a distinctiveness heuristic (Dodson & Schacter, 2001; Israel & Schacter, 1997) on a memory test to probe their memories for distinctive information confirming that a word was studied (e.g., "I remember saying it out loud, so I must have studied it").

In accord with the distinctiveness account, the production effect is substantially larger in a within-subject (i.e., mixed-list) design than in a between-subjects (pure-list) design (Fawcett, 2013). Although the distinctiveness heuristic could aid recognition of words read aloud in either a mixed or a pure study list, only words read aloud in a mixed list benefit from distinctive processing at study (for a discussion, see Fawcett, 2013). Moreover, Forrin, Groot, and MacLeod (2016) found a benefit of reading aloud in a mixed list compared to reading aloud in a pure-aloud list, which is also consistent with distinctiveness enhancing memory for aloud items at the time of encoding (see also Ozubko & MacLeod, 2010; Ozubko, Major, & MacLeod, 2014).

It is important to consider, however, that in a mixed list containing both aloud and silent words, the production effect reflects not only a benefit of reading aloud but also a *cost* of reading silently (Bodner, Taikh, & Fawcett, 2014; Forrin et al., 2016). Silent words are more poorly remembered—in terms of both recognition and free recall—when they are studied in a mixed list relative to a pure list consisting of only silent words.[1] It is this observation that led us to consider the possibility that the cost to reading silently in a mixed list could be driven, at least in part, by participants' *anticipation* of having to read aloud. That is, the next-line-effect that Brenner (1973) demonstrated over 45 years ago might also explain the recently observed cost (in recognition and free recall) in mixed-list production research: While reading a silent item, participants anticipate the next aloud item, causing that silent item to be less well encoded.

Can performance anticipation explain the mixed-list cost of production?

Here we investigated a novel hypothesis inspired by Brenner's (1973) next-in-line research: that *performance anticipation* contributes to the cost in the production effect. Because of the random ordering of conditions in a mixed study list, participants cannot predict when they will have to read aloud, which may result in persistent performance anticipation that particularly influences silent trials. That is, when participants are studying silent items, they may tend to be distracted by the impending prospect of having to read aloud, and this in turn could undermine their encoding of silent items. Although this memory decrement ought to be evident in both recall and recognition tests, we focused on recognition here because the cost of reading silently is not as well understood in recognition as it is in recall (see Forrin & MacLeod, 2016).

In the context of the next-in-line effect, Bond and Kirkpatrick (1982) found evidence that prolonged performance anticipation can impose a "blanket" cost on items studied passively. They induced prolonged performance anticipation by modifying Brenner's (1973) procedure so that the order in which participants read aloud was randomized (thereby preventing them from predicting when they would be called upon to read aloud). In contrast to the standard next-in-line effect, which is limited to a few items prior to performance, randomizing the performance order extended the effect across all non-performance words—ostensibly because performance anticipation was likewise extended. Importantly, Bond and Kirkpatrick further showed that performance anticipation is required to obtain a next-in-line effect: When participants did not expect to read a word aloud—but were still called on to do so—the effect was eliminated.

Bond (1985) subsequently obtained converging evidence that performance anticipation imposes its cost at encoding rather than at retrieval. Bond and Omar (1990) revealed that the next-in-line effect was significantly larger in socially anxious (vs. non-anxious) individuals, presumably because socially anxious individuals are more likely to have their attention diverted by the prospect of performance—in this case, of reading aloud. And Bond et al. (1991) found that instructing participants to deeply encode pre-performance study items eliminated the next-in-line effect. This collection of results led Bond et al. (1991) to conclude that "anticipation diverts cognitive resources from memorization" (p. 436), resulting in the shallow encoding of pre-performance study items.

Bond and colleagues' findings therefore served as a basis for our hypothesis that performance anticipation could be driving the cost for silent items in the mixed-list production effect. Whereas participants may perpetually anticipate reading aloud while studying the silent items in a mixed list, performance anticipation would be absent while studying the items in a pure-silent list.

The present research

Across four experiments, we evaluated this performance anticipation hypothesis for the cost of production by examining whether memory was worse for silent items preceding items that participants expected to read aloud (relatively high anticipation) compared to silent items preceding items that they expected to read silently (relatively low anticipation). In accord with Bond and Kirkpatrick's (1982) view that performance anticipation diverts attention from the passively encoded pre-performance study items, we propose that anticipating reading aloud in a production experiment diverts participants' attention when reading silently. For example, the prospect of reading aloud may give rise to self-presentation concerns (given that a research assistant is present during the experiment; see Baumeister, 1982, for a review). Even in the absence of conscious thoughts regarding reading aloud, performance anticipation may increase anxiety to a level that is detrimental to encoding (see e.g., Moran, 2016; Vogel & Schwabe, 2016, for reviews).

The present experiments also allowed us to evaluate the viability of a related explanation for the mixed-list cost of the production effect in recognition: the lazy reading account (Begg & Snider, 1987). This account, which derives from work on the generation effect (the finding that words generated from a cue are remembered better than words

[1] When free recall is tested, only the mixed-list cost to reading silently, not the benefit to reading aloud, emerges reliably, and the between-subjects production effect is consistently nonsignificant (Forrin & MacLeod, 2016; Jones & Pyc, 2014; Jonker et al., 2014; Lambert, Bodner, & Taikh, 2016).

that are simply read aloud; Slamecka & Graf, 1978), states that people devote less processing to the items that they see as relatively less important or less cognitively demanding in a mixed study list.

MacLeod et al. (2010, Experiment 7) tested the lazy reading account by having participants generate all of the words at study, thereby forcing elaborative encoding of all items and preventing lazy reading of the silent items. Contrary to the lazy reading account, the production effect remained significant in that experiment. Forrin, Jonker, and MacLeod (2014, Experiment 1) took this a step further, demonstrating that the production effect for generated words was statistically equivalent in size to that for read words (see also Taikh & Bodner, 2016), and, in a second experiment, that imaging the referents of words at study also did not diminish the size of the production effect relative to imaging the orthography of the words. Conversely—consistent with lazy reading—Fawcett and Ozubko (2016) found that participants reported paying less attention to silent items than to aloud items in a mixed study list, and Sousa, Carriere, and Smilek (2013) found that participants reported mind-wandering more frequently when reading passages silently vs. aloud. Thus, although experimental manipulations that targeted lazy reading clearly did not support the lazy reading account, participants' self-reported lack of attentiveness to silent items suggests that the account warrants further research.

Importantly, our performance anticipation account constitutes a *specific* explanation for why participants may pay less attention to silent items (i.e., lazily read) in a mixed-list production experiment: Attention is diverted from silent items in anticipation of reading aloud. There are, of course, other possible reasons why participants might pay less attention to silent items; for example, they may judge silent items to be less important than aloud items. The present research does not assess the viability of these other explanations for lazy reading. Rather, we test the performance anticipation account in particular using the following logic: If performance anticipation imposes a cost on silent items in a mixed list, then memory for silent items should be particularly poor when those items immediately precede other items that participants anticipate reading aloud. This hypothesized pattern of results would not rule out the possibility that other factors (e.g., judgments of low importance or cognitive demands) further diminish attention—and memory—for silent items. That said, our hypothesized empirical pattern would suggest that those other factors do not *fully* account for why participants lazily read silent items in a mixed-list production experiment. Uniquely, the performance anticipation account predicts a gradient of cost for the silent items.

In sum, our goal was to answer the question "Does the cost imposed by performance anticipation manifest itself in laboratory production experiments just as it does in "next-in-line" experiments? It should, if the same mechanism underlies both phenomena. Although we focused here on explaining the cost of production, our results have broader implications for memory research. Performance anticipation may account for costs observed for other widely-researched encoding techniques such as the generation effect (e.g., Slamecka & Graf, 1978) and the enactment effect (e.g., Engelkamp & Dehn, 2000), where anticipated performance in one condition may undermine memory for the unperformed material in another condition.

Experiment 1: A within-subject vs. between-subjects design

Before investigating whether performance anticipation can account for the mixed-list cost in the production effect, we thought it prudent to replicate this cost using the same within-subject vs. between-subjects design used by Bodner et al. (2014) (see also Forrin et al., 2016). In this design type, one group of participants studies a mixed list (of randomly ordered 'aloud' and 'silent' items) and two other groups study pure lists of either all aloud or all silent items. A mixed-list *benefit* occurs when recognition is better for aloud items studied in a mixed list vs. a pure list. A mixed-list *cost* occurs when recognition is worse for silent items studied in a mixed list vs. a pure list. We expected to obtain both a benefit and a cost in terms of hits, replicating Bodner et al. and Forrin et al.[2] The mixed-list benefit would be consistent with distinctiveness enhancing memory. The mixed-list cost could be explained by multiple factors, including performance anticipation (which we see as a specific form of lazy reading in which participants shallowly process silent items because they are distracted by the prospect of reading aloud).

Method

Participants

A total of 150 undergraduate students from the University of Waterloo participated for course credit. The first 50 participants were assigned to the within-subject (i.e., mixed list) group. The next 100 participants were randomly assigned to one of the between-subjects groups (i.e., pure-aloud list or pure-silent list). A sensitivity analysis using the statistical software G*Power 3.0 (Erdfelder, Faul, & Buchner, 1996) showed that this sample size had adequate statistical power (0.80) to detect an effect size as small as $d = .057$ at a significance criterion of $\alpha = .05$.

Stimuli

We used the same word pool as in Forrin et al. (2016), consisting of 240 words obtained from the MRC Psycholinguistic Database (http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm). All words had frequencies of greater than 30 per million (Thorndike & Lorge, 1944) and were 5–10 letters long. For each participant, 80 words were randomly assigned to the study phase and a further 80 were randomly assigned to be distractors during the test phase.

Apparatus

Stimuli were presented and responses were collected using E-Prime 2.0 software (Psychology Software Tools Inc., Pittsburgh, PA). Instructions and stimuli were displayed on a 17″ LCD monitor.

Procedure

Participants took part in the experiment individually in a laboratory room with an experimenter present. In the within-subject version of the experiment, font color (blue or white) indicated whether each study item should be read aloud or silently. Both font color and study condition were counterbalanced; in each case, participants were instructed in advance to try to remember all of the words regardless of their color. The study list consisted of 80 words—40 blue and 40 white, in random order—presented centrally in 16 pt. Courier New lowercase bold font against a black background. Each word was presented for 3 s followed by a 500-ms inter-stimulus interval. The study lists in the between-subjects versions were identical with the exception that the 80 words were either all in blue (aloud) or all in white (silent) print.

The study phase was immediately followed by a 160-item old/new recognition test in which the 80 study items were randomly intermixed with 80 new distractor items. Test items were presented one at a time in yellow font against a black background. Participants made keypress responses to label each item as old ('m') or new ('c'), with each word remaining on the screen until the participant responded.

It is worth noting the extent of a research assistant's interactions

[2] Bodner et al. (2014) reported only a significant cost in terms of d'. In terms of hits, however, a re-analysis of their raw data revealed both a statistically significant cost and a statistically significant benefit (as reported in Forrin et al., 2016). Forrin et al. (see also Fawcett, Quinlan, & Taylor, 2012) argued that hits are the only reliable measure for assessing cost and benefit in recognition because the within-subject design yields only a single false alarm (FA) rate for both aloud and silent items. When Forrin et al. dissociated within-subject FAs, they found evidence of a cost in hits, but not in FAs (i.e., participants had comparable FAs for silent items that were studied in a mixed list vs. in a pure list). Those results suggest that the cost in recognition is not driven by response bias; instead, it reflects lower sensitivity to recognizing silent items when they are studied in a mixed (vs. pure) list. We therefore focused in the present investigation on examining the cost in hits.

with a participant: Those interactions constituted a social component that could have increased performance anticipation and, consequently, resulted in a cost to memory (as has been found in research on the next-in-line effect; see Bond & Omar, 1990). Importantly, how research assistants interacted with participants was identical to prior production effect experiments that we have conducted, including experiments that have yielded costs in mixed-list production (e.g., Forrin et al., 2016).

Five undergraduate research assistants conducted the experiment. Each research assistant met a participant in the waiting room at the scheduled time and led that participant to the laboratory room. Research assistants were trained to be professional, friendly, and consistent in their interactions with participants. Once in the laboratory room, the research assistant requested that the participant's phone be silenced. After signing the consent form, the participant was asked to read the study instructions on the computer monitor and to summarize them orally. In the rare event that a participant misunderstood the instructions, the RA corrected them. For the duration of the experiment, the research assistant sat at a table approximately 10–15 feet behind the participant (depending on which laboratory room was used), and outside the participant's line of sight. Research assistants followed the same protocol in subsequent experiments.

Thus, although the experiment itself was non-social (participants read words aloud to themselves rather than to the research assistant), the presence of the research assistant in the room constituted a social factor—an audience of sorts—that may have brought about performance anticipation. This possibility is addressed in Experiment 4.

Results

False alarms (FAs) were significantly lower for participants who studied a pure-aloud list ($M = .15$, $SE = .02$) than for those who studied a pure-silent list ($M = .23$, $SE = .02$), $t(98) = 3.19$, $p = .002$, $d = 0.64$, with the mixed-list FAs intermediate ($M = .18$, $SE = .02$). FAs could not be meaningfully compared across the two experimental designs because separate FA rates for 'aloud' and 'silent' study items cannot be obtained from a within-subject design (which also undermines comparisons of *d'* across experimental designs; see Forrin et al., 2016, for a discussion). We therefore focused on comparisons of hits.

Table 1 shows the mean hits for each condition. Unsurprisingly, the within-subject (mixed-list) production effect in hits was robust, $t(49) = 9.61$, $p < .001$, $d = 1.28$. The between-subjects (pure-list) production effect, while more modest in size, was also statistically significant, $t(98) = 2.12$, $p = .04$, $d = 0.42$, consistent with recent work showing a small but reliable between-subjects production effect (e.g., Forrin et al., 2016; see Fawcett, 2013, for a meta-analysis) and mirrored the between-subjects FAs. To compare the magnitude of within-subject and between-subjects production effects, we used Erlebacher's (1977) analysis, a modified ANOVA used to compute the interaction between Design Type (within-subject vs. between-subjects) and the independent variable manipulated in both designs—in this case, Item Type (aloud vs. silent). Erlebacher's analysis revealed a significant Design Type × Item Type interaction, $F(1, 98) = 15.49$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .04$, signifying that the production effect was significantly larger within-subject than between-subjects.

Of main interest, hits were 8 percentage points lower for words read silently in a mixed list compared to those in a pure silent list, $t(98) = 2.94$, $p = .004$, $d = 0.59$ (i.e., a significant cost). Hits were also 5 percentage points higher for words read aloud in a mixed list relative to those in a pure aloud list, although this benefit was nonsignificant, $t(98) = 1.69$, $p = .10$, $d = 0.34$. This nonsignificant benefit had a comparable effect size to those obtained in Forrin et al. (2016), which suggests that the present experiment was underpowered to detect this (modest) effect size.

The overall pattern of hits in Experiment 1 closely resembled those obtained by Bodner et al. (2014) and by Forrin et al. (2016). We found a robust within-subject production effect and a significantly smaller (though still significant) between-subjects production effect. The larger within-subject production effect was primarily attributable to a significant cost for reading silently in a mixed list. Having successfully replicated this cost—the basis for our present investigation—we turned our attention to testing whether this cost could be explained by performance anticipation diminishing memory for silent items in a mixed list.

Table 1
Experiment 1: Mean hits (with SEs) for each group and item type.

	Aloud Items	Silent Items
Mixed-list Group	.75 (.02)	.56 (.02)
Pure-list Groups	.70 (.02)	.64 (.02)

Experiment 2: Alternating between five aloud and five silent items

In Experiment 2, we explored whether the cost in the mixed-list production effect observed in Experiment 1—and in prior research (see Bodner et al., 2014; Forrin et al., 2016)—could be explained by performance anticipation for aloud items weakening the encoding of silent items. Performance anticipation may be continuous in a mixed-list design (because participants cannot foresee when aloud items will appear in a randomized study list), resulting in a "blanket" cost in memory for silent items (Bond & Kirkpatrick, 1982). If this explanation is viable, then it follows that informing participants when aloud items will appear in a mixed list ought to heighten the performance anticipation that participants experience while studying silent items that immediately precede aloud items, resulting in relatively poor memory for those silent items.

We made the occurrence of 'aloud' items predictable in Experiment 2 by having the study list alternate between blocks of five consecutive aloud items and blocks of five consecutive silent items, and by informing participants of this item order in the study instructions. As participants progress through a block of five silent items, performance anticipation ought to increase as a block of aloud items draws nearer. Conversely, participants should not experience performance anticipation while progressing through a block of five aloud items because they are already engaged in the performance of reading aloud. We therefore hypothesized that hits would decline over the five positions of each block of silent study items (as in the results from the next-in-line literature; e.g., Brenner, 1973), while remaining constant over the five positions of each block of aloud study items. This empirical pattern would be consistent with performance anticipation imposing a cost on silent items in a mixed-list production experiment.

If performance anticipation indeed diminishes memory for silent items that participants expect to be followed shortly by aloud items (in Experiment 2) and for silent items in a randomly mixed list (in Experiment 1), then a similar cost to memory should be observed in both cases. We therefore hypothesized that the mean hit rate for silent items studied in position 5 of each silent block (Experiment 2) would be comparable to that for silent items studied in a mixed list (Experiment 1). Relatedly, we hypothesized that the mean hit rate for silent items studied in position 5 of each silent block (Experiment 2) would be significantly *lower* than that for silent items studied in a *pure* list (Experiment 1), which could not be influenced by performance anticipation. That is, we predicted that there would be a cost to silent items immediately followed by aloud items, just as there is a cost to silent items studied in a mixed list (that participants realize *could* be immediately followed by aloud items).[3] Along these lines, we also

[3] Although Bodner et al. (2014) did not find evidence of a cost to production in a within-subject blocked design (in which participants studied two large blocks, one of aloud items and one of silent items, with the order counterbalanced), participants in their experiment were, importantly, *not aware* that they would be studying a second block of words. Thus, their participants who studied a 'silent block' first did not know that they would subsequently study an 'aloud block,' so performance anticipation could not have occurred in their study.

predicted that the mean hit rate for silent items studied in position 1 of each silent block (Experiment 2) would be comparable to that for silent items studied in a pure list (Experiment 1) because performance anticipation would be low/absent in both cases.

Method

Participants

Fifty undergraduate students from the University of Waterloo participated for course credit; this sample size was equivalent to each of the three conditions in Experiment 1.

Stimuli and apparatus

These were the same as in Experiment 1.

Procedure

The procedure was the same as for the mixed-list condition of Experiment 1, except that the 80-item study list alternated between blocks of five aloud items and blocks of five silent items, with 8 blocks of each type. The pairing of condition and font color was randomly counterbalanced, as was the initial condition of the list (i.e., starting with an aloud block vs. a silent block). Instructions informed participants that the list would alternate in this manner; the experimenter also asked participants to summarize the study instructions before beginning (as was the case in Experiment 1), thereby ensuring that participants understood the structure of the study list. The recognition test was identical to that of Experiment 1, featuring all 80 study items and 80 new distractor items in a purely random sequence.

To stay consistent with Experiment 1, the study list did not indicate the relative position of each aloud item or silent item within each block of five items, nor was there any indicator of when a block of five items started or ended. Thus, in terms of our performance anticipation hypothesis, participants ought to have been aware of when their performance was drawing near, even if they did not keep track of exactly when the next aloud item would occur.

Results

Fig. 1 displays hit rates (and SEs) for aloud and silent items in each block position. The mean FA rate was .18 (*SE* = .02). Once again, our analyses focused on hits because the single FA rate yielded by this within-subject design was not informative for testing our account that performance anticipation reduces memory for silent items in a mixed list.

Under our performance anticipation account, hits for silent items should, on average, decrease from position 1 to position 5 of the silent blocks but should remain consistently high across the five positions of the aloud blocks. Fig. 1 displays precisely this pattern of means. We tested this hypothesis statistically by conducting an Item Type (aloud vs. silent) by Block Position (1–5) repeated measures ANOVA. Unsurprisingly, the Item Type main effect was significant, $F(1, 49) = 129.59$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .73$, indicating a robust production effect. The Block Position main effect was nonsignificant, $F(4, 196) = 1.17$, $MSE = 0.02$, $p = .33$, $\eta_p^2 = .02$. Most important, and consistent with our hypothesis, the Item Type × Block Position interaction was significant, $F(4, 196) = 4.53$, $MSE = 0.02$, $p = .002$, $\eta_p^2 = .08$, as was the linear contrast of this interaction, $F(1, 49) = 13.41$, $MSE = 0.02$, $p = .001$, $\eta^2 = .21$.

To unpack the significant interaction, we conducted separate ANOVAs for aloud items and for silent items. For aloud items, the ANOVA was nonsignificant, $F(4, 196) = 1.45$, $MSE = 0.02$, $p = .22$, $\eta_p^2 = .03$, signifying that position within study block did not differentially affect memory for aloud items. Conversely, for silent items, the ANOVA was significant, $F(4, 196) = 4.13$, $MSE = 0.02$, $p = .003$, $\eta_p^2 = .08$; indeed, the linear contrast was also significant, $F(1, 49) = 13.67$, $MSE = 0.02$, $p = .001$, $\eta_p^2 = .22$, indicating that mean

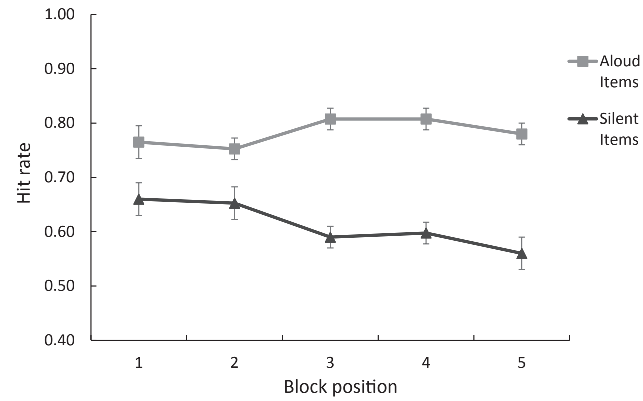

Fig. 1. Experiment 2: Mean hit rate (with SE) for each block position and item type. Aloud items: Block position 1 (*M* = .77, *SE* = .03), position 2 (*M* = .75, *SE* = .02), position 3 (*M* = .81, *SE* = .02), position 4 (*M* = .81, *SE* = .02), position 5 (*M* = .78, *SE* = .02). Silent items: Block position 1 (*M* = .66, *SE* = .03), position 2 (*M* = .65, *SE* = .03), position 3 (*M* = .59, *SE* = .02), position 4 (*M* = .60, *SE* = .02), position 5 (*M* = .56, *SE* = .03).

hits for silent items decreased linearly across block positions, consistent with the prediction that memory should decrease as performance anticipation increases. (The quadratic contrast was nonsignificant; $p = .90$.) Post-hoc *t*-tests (with a Bonferroni correction) revealed that mean hits were significantly fewer for silent items studied in position 5 than for those studied in position 1, $t(49) = 3.35$, $p = .02$, $d = 0.48$. The difference between position 5 and position 2 was marginal, $t(49) = 2.81$, $p = .07$, $d = 0.32$. All other differences between block positions were nonsignificant (*p*s > .20).

Comparing means across Experiments 1 and 2 also yields support for our performance anticipation account (cf. Tables 1 and 2). As hypothesized, mean hits for silent items studied in a *mixed* list in Experiment 1 (*M* = .56, *SE* = .02) were comparable to (indeed, numerically equal to) mean hits for silent items studied in block position 5 in Experiment 2 (*M* = .56, *SE* = .03), a nonsignificant difference, $t(98) = 0.14$, $p = .89$, $d = 0.03$. We submit that recognition of silent items was relatively low in both cases because performance anticipation was relatively high in both cases. Of main interest, mean hits were significantly lower for silent items studied in block position 5 in Experiment 2 than for silent items studied in a pure list in Experiment 1, $t(98) = 2.34$, $p = .02$, $d = 0.47$, a result that dovetails with the mixed-list cost observed in Experiment 1. Relatedly, mean hits were nonsignificantly different when comparing silent items studied in a *pure* list in Experiment 1 (*M* = .64, *SE* = .02) to silent items studied in block position 1 in Experiment 2 (*M* = .66, *SE* = .03), $t(98) = 0.58$, $p = .56$, $d = 0.12$; we contend that recognition was relatively high in both of these cases because performance anticipation was relatively low (or

Table 2
Experiments 1, 3, and 4: Mean hits (with SEs) for each item type as a function of the next item type.

Experiment	Aloud Next	Silent Next	Overall
Experiment 1			
Aloud Items	.74 (.02)	.75 (.02)	.75 (.02)
Silent Items	.56 (.02)	.54 (.03)	.55 (.02)
Experiment 3			
Aloud Items	.77 (.01)	.77 (.01)	.77 (.01)
Silent Items	.53 (.02)	.58 (.02)	.55 (.01)
Experiment 4			
Aloud Items	.70 (.02)	.69 (.02)	.70 (.02)
Silent Items	.58 (.02)	.59 (.02)	.58 (.02)

Note. These means do not include recognition performance for the final study item because there was no item following it.

absent, in the case of the pure-silent list). Thus, the pattern of mean hits across Experiments 1 and 2 is entirely consistent with our account that performance anticipation reduces memory for silent items.[4]

Discussion

In Experiment 2, we found that when a mixed list followed a predictable pattern—alternating between blocks of five aloud items and five silent items—memory for the silent items decreased linearly within each silent block. This result is consistent with our explanation that performance anticipation disrupts the encoding of silent items, which may therefore also account for the mixed-list cost in the production effect that we replicated in Experiment 1. In Experiment 1—as in most mixed-list production experiments—anticipation of reading aloud loomed continuously throughout the study phase and may consequently have imposed a "blanket" cost on items read silently. In Experiment 2, however, anticipation of reading aloud increased over the course of silent items in a study block, consistent with the gradual decrease in recognition. This pattern of results is less compatible with the typical formulation of the lazy reading account (inspired by Begg & Snider, 1987; see MacLeod et al., 2010, for a discussion in the context of the production effect), which purports that silent items generally tend to be shallowly processed in a mixed list and which does not predict any position effects. If individuals do tend to (indiscriminately) pay less attention to silent items, then one would expect a consistently widespread cost imposed on silent items in Experiment 2; instead, we found that the cost was absent at the beginning of each study block (relative to Experiment 1) and gradually increased, a pattern entirely consistent with our performance anticipation account.

There is another possible explanation for the pattern of results in Experiment 2: Individuals could be more likely to mind-wander (i.e., have task-unrelated thoughts) while reading silently than while reading aloud (see Sousa et al., 2013; cf. Franklin, Mooneyham, Baird, & Schooler, 2014). Essentially, silent reading imposes relatively low attentional/working memory demands, which "frees up" more of those resources for mind-wandering (see Smallwood & Schooler's, 2006, *executive resource hypothesis*). In Experiment 2, then, perhaps memory for silent items decreased within silent blocks because individuals were increasingly likely to initiate task-unrelated mind-wandering over a block of passive trials. However, if task-unrelated mind-wandering reduces memory for silent items in production effect research, then memory ought to be particularly poor for silent items studied in a pure list—which consists of one long sequence of passive processing (and thus presents a prolonged opportunity for mind-wandering). To the contrary, silent words are better remembered when they are studied in a pure list vs. a mixed list (e.g., Experiment 1). It therefore seems improbable that task-unrelated mind-wandering accounts for the mixed-list cost of reading silently, or for the pattern of results observed in Experiment 2.[5]

[4] We further analyzed the nonsignificant differences between experiments using Wagenmakers (2007) Bayesian approximation procedure. Posterior odds were calculated from the sum of squares output by the ANOVA and were converted to pBIC (see Masson, 2011), which quantifies support for the null hypothesis on a scale from 0 to 1. For the difference between the silent items in a mixed list (Experiment 1) and the silent items in block position 5 (Experiment 2), the Bayesian approximation procedure yielded "positive" support for a null effect, $p\text{BIC} = .91$, according to Raftery (1995) labeling system ["weak" (.50–.75), "positive" (.75–.95), "strong" (.95–.99), "very strong" (> .99)]. For the difference between silent items in a mixed list (Experiment 1) and silent items in block position 1 (Experiment 2), the Bayesian approximation procedure also yielded "positive" support for a null effect, $p\text{BIC} = .89$.

[5] A reviewer raised the possibility that a primacy effect (e.g., Murdock, 1962) could explain the results of Experiment 2—that participants disproportionately rehearsed silent items at the beginning of each block of silent items. We are not aware, however, of any theory that would predict a primacy effect only for the silent blocks and not for the aloud blocks. Moreover, in Experiment 3, using a different paradigm in which primacy is not possible, we will show further evidence of a pre-performance cost in memory for silent items.

Experiment 3: Revealing upcoming conditions in a mixed list

The pattern of results in Experiment 2 is consistent with our account that performance anticipation reduces memory for silent items in a mixed study list, but the design that we used—consisting of alternating blocks of five aloud items and five silent items—differed from the *randomly* mixed list design previously used to demonstrate a cost in the production effect literature (Experiment 1; Bodner et al., 2014; Forrin et al., 2016). We have contended that this previously observed cost is driven by essentially continuous performance anticipation (i.e., throughout a mixed list, it is always possible that an aloud item is "next-in-line" in the study list).

Our objective in Experiment 3 was to assess whether performance anticipation affects memory within the framework of a randomly mixed study list. We did so by taking an approach similar to that in Experiment 2: revealing to participants the condition (aloud/silent) of upcoming study items. Previewing what is "next in line" should influence the amount of performance anticipation experienced by participants; specifically, while studying a silent item, participants ought to experience more performance anticipation when they see that an aloud item (vs. a silent item) is next in line. Of course, while studying an aloud item, participants are already in the act of performing so we do not expect that the condition of the next item would have much, if any, influence on performance anticipation.

Consequently, we hypothesized that when participants are aware of the condition that is coming next, they should have poorer memory for silent items followed by aloud items (higher performance anticipation) than for silent items followed by silent items (lower performance anticipation). Of course, we expected memory for aloud items to be better overall than memory for silent items (the typically robust mixed-list production effect) and not to be significantly influenced by whether the next item was aloud or silent. This pattern of results would support our account that performance anticipation imposes a cost on memory for silent items in a randomly mixed list.

Method

Participants

We decided to run a robust sample in case our manipulation yielded a small effect. The G*Power 3.0 software indicated that 90 participants would yield adequate statistical power (.80) to detect a small two-tailed effect ($d = .03$) at an $\alpha = .05$ significance level. We therefore ran 100 University of Waterloo students who were reimbursed with course credit.

Stimuli

The same word pool was used as in Experiments 1 and 2. A left-to-right series of blue and white horizontal lines, each 3 cm long and two mm thick, indicated the condition of each study word (i.e., 'aloud' or 'silent'). We counterbalanced font color and study condition (and ran fifty participants for each pairing) to ensure that the color indicating the task for the upcoming study item did not influence our results.

Each horizontal line was positioned 1 cm below the middle of the screen (akin to an underline for each word in the study list), and these lines were evenly spread across the screen (separated by 2 cm). One underline was positioned in the center of the screen below the current study item, which was the only item displayed on the trials. Four underlines on the right side of the screen represented the position in line of the four upcoming study items, and could appear in blue or in white. For example, if the underline that was immediately to the right of the current study item was blue, then the next study item would be a blue (e.g., 'aloud') item. Likewise, four underlines on the left side of the screen corresponded to the four previous study items. After each study word was presented, the underlines slid one position to the left in one fluid motion/animation (resembling a conveyor belt) and a new central word was presented. The leftmost underline "slid off" the screen and a new underline emerged from the right side of the screen. Of course, when the first study item was presented, there were no underlines on the left side of the screen; and when the last study item was presented, there were no underlines on the right side of the screen.

Apparatus

This was identical to Experiment 1.

Procedure

The procedure was mostly identical to the mixed-list condition of Experiment 1. The single significant exception was the presence of the blue and white underlines that indicated the condition of each study word and shifted to the left by one position after each word was studied. Participants were not instructed to pay attention to the underlines (and could have readily completed the experiment without doing so). The study instruction screen showed two examples that illustrated the orientation of the study items and the underlines. The instructions described the underlines as follows:

> "The lines indicate the colours of the upcoming words in the study list, as well as the colours of the previous words. In the above example, the lines to the right of "horse" indicate that the next word will be blue, the word after that will be white, etc. The lines to the left of "horse" indicate that the previous word was blue, the word before that was white, etc. After you study each word, the lines will all move one position to the left (like a conveyor belt) and the next word will appear in the middle of the screen. Please try to remember all of the words, regardless of their color."

As in Experiments 1 and 2, each study word was presented for 3 s and then disappeared leaving only the lines present. After another 200 ms, the underlines moved one "position" to the left in a fluid animation. The timing of this animation varied somewhat (the animation was programmed by presenting several screens for 5 ms each, but some lag occurred due to the monitor's refresh rate). Consequently, the animation took between 249 ms and 400 ms on any given trial ($M = 281$ ms, $SD = 29$ ms). After the underlines moved to their new positions, there was a 200-ms delay before the next study item appeared. (These short delays were inserted to make it easier to follow what was happening.) In total, then, there was approximately a 681-ms interval between successive study items (slightly longer than the 500-ms interval in Experiments 1 and 2).

The study phase progressed in this manner until all 80 words had been presented. A 160-item old/new recognition test followed that was identical to those in the prior experiments. The instructions for the recognition test were also unchanged (i.e., there was no further mention of the underlines).

Results

The influence of the next study condition on memory

Table 2 shows the hits for aloud and silent study items as a function of whether they were immediately followed by an aloud item or a silent item (i.e., the condition that was "next in line"; Brenner, 1973). The mean FA rate was .20 ($SE = .01$). To test our performance anticipation account—that memory for silent items (but not for aloud items) would be poorer when they were followed by an aloud item—we conducted a 2 (Current Item: aloud vs. silent) × 2 (Next Item: aloud vs. silent) repeated measures ANOVA. There was a significant main effect of Current Item, $F(1, 99) = 234.28$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .70$, indicative of a robust production effect. The main effect of Next Item was also significant, $F(1, 99) = 6.35$, $MSE = 0.01$, $p = .01$, $\eta_p^2 = .06$, signifying that hits were lower overall when an aloud (vs. silent) item was next. Importantly, the Next Item main effect was driven by a significant Current Item × Next Item interaction, $F(1, 99) = 6.34$, $MSE = 0.01$, $p = .01$, $\eta_p^2 = .06$. Consistent with our hypothesis, silent hits were significantly lower when an aloud item was next relative to when a silent item was next, $t(99) = 3.38$, $p = .001$, $d = 0.28$, whereas aloud hits were unaffected by the condition of the next item, $t(99) = 0.12$, $p = .90$, $d = 0.01$.

The influence of the previous study condition on memory

Next, we explored whether the condition of the previous study item (aloud vs. silent) affected memory for silent items. Table 3 shows the hits for silent items as a function of the condition of both the previous study item and the next study item.

In line with our performance anticipation account, we had hypothesized that recognition of silent items would be influenced by the *next* condition (with lower hits when an aloud item came next). We did not expect the *previous* condition to modulated the size of that effect. To test these predictions, we conducted a Previous Item (aloud vs. silent) × Next Item (aloud vs. silent) ANOVA on silent items. Of main interest, there was a significant Next Item main effect, $F(1, 99) = 12.90$, $MSE = 0.02$, $p = .001$, $\eta_p^2 = .12$, signifying that participants had worse memory for silent items when an aloud (vs. silent) item was next. In contrast, the main effect of Previous Item was non-significant, $F(1, 99) = 2.87$, $MSE = 0.03$, $p = .09$, $\eta_p^2 = .03$; participants tended to have slightly (but non-significantly) better memory for silent items that were preceded by an aloud item (vs. a silent item). Importantly, the Next Item × Previous Item interaction was non-significant ($F < 1$), indicating worse memory for silent items followed by an aloud item (vs. a silent item), regardless of the condition of the previous study item. Whether the previous item was aloud, $t(99) = 2.20$, $p = .03$, $d = 0.25$, or silent, $t(99) = 2.78$, $p = .007$, $d = 0.26$, participants had significantly worse memory for silent items followed by an aloud item (vs. a silent item). Also in accord with our hypothesis, memory for aloud items was quite consistent regardless of the condition of the neighboring study items. For aloud items, a Next Item (aloud vs. silent) × Previous Item (aloud vs. silent) ANOVA revealed non-significant main effects of Next Item and Previous Item ($Fs < 1$) as well as a non-significant Next Item × Previous Item interaction, $F(1, 99) = 2.76$, $MSE = 0.02$, $p = .10$, $\eta_p^2 = .03$.

In sum, memory for silent items was significantly influenced by the study condition of the next item (consistent with our performance anticipation account), but not by the study condition of the previous item.

Discussion

Providing participants with advance information about upcoming study conditions (i.e., read aloud vs. silently) influenced memory entirely in accord with our performance anticipation account: Memory for silent items was worse when an aloud item was next (relatively high

Table 3
Experiments 1 and 3: Mean hits (with SEs) for silent items as a function of the previous item type and the next item type.

Experiment	Aloud Next	Silent Next
Experiment 1		
Aloud Previous	.58 (.03)	.56 (.03)
Silent Previous	.56 (.03)	.54 (.03)
Experiment 3		
Aloud Previous	.54 (.02)	.59 (.02)
Silent Previous	.51 (.02)	.57 (.02)

Note. These means include recognition performance only for study items for which there was both a previous item and a next item (i.e., the first item and final item are not included).

performance anticipation) relative to when a silent item was next (relatively low performance anticipation).[6]

Conversely, memory performance for aloud items was not affected by the condition of the study item that followed (or preceded) them. This would seem to rule out the alternate explanation that the anticipation of task-switching resulted in a memory decrement. Rather, our results suggest that it is specifically the anticipation of switching from reading silently to reading aloud that reduces memory, a consequence of performance anticipation.

The condition of the previous study item, in contrast, did not significantly influence memory. Importantly, the negative effect that an upcoming aloud item had on memory for silent items was apparent regardless of whether the *previous* study item was aloud or silent. This result provides further evidence against the alternate explanation that the "next-in-line" effect obtained in Experiment 2 emerged due to a string of prior silent items that prompted task-unrelated mind-wandering. Rather, we contend that, within a mind-wandering framework, anticipation of an upcoming aloud item while reading silently can be seen as having prompted *task-related* mind-wandering (i.e., mind-wandering regarding a component of the experiment—in this case reading aloud—that interferes with ongoing performance; see, e.g., Smallwood et al., 2004). We expand on this point in the General Discussion.

The present results are consistent with our proposal that performance anticipation diverts attention resources from silent items in a standard mixed-list production experiment. While prior discussions of the lazy reading account in the context of production (see, e.g., MacLeod et al., 2010) have presumed that participants generally tend to process silent items shallowly (i.e., lazily read), the present findings suggest that a context-specific factor brings about this shallow processing: performance anticipation. We contend that in a mixed-list design, performance anticipation is ordinarily relatively constant across silent items (given that participants are unaware of when they will have to read aloud), which results in a "blanket" cost in memory (cf. Bond & Kirkpatrick, 1982).

Combined analyses of Experiment 1 (mixed-list group) and Experiment 3

In accord with our performance anticipation account of the mixed-list cost in production, the added study design feature in Experiment 3—underlines indicating the order of study conditions—resulted in worse memory for silent items that preceded an aloud item (vs. a silent item). If our account is valid, then the mixed-list group in Experiment 1 (a typical mixed-list production experiment) should not yield this study order effect: Performance anticipation ought to be relatively consistent across all silent study items (i.e., a "blanket cost") because participants were unaware of when they would have to read aloud. The presence of a study order effect in Experiment 1 would therefore undermine our performance anticipation account, and would instead support an alternate explanation: that reading aloud retroactively interferes with encoding silent item (i.e., a "retroactive amnesia" effect; see Tulving, 1969).

Table 2 shows the hits for aloud and silent items in Experiment 1 (mixed-list group) and in Experiment 3. Inspection of the means confirms that the study order effect observed in Experiment 3 (i.e., worse memory for silent items followed by aloud items) did not occur in Experiment 1. This observation is supported by an Experiment (1 vs. 3) × Current Item (aloud vs. silent) × Next Item (aloud vs. silent) ANOVA, which yielded a significant three-way interaction, $F(1, 148) = 5.16$, $MSE = 0.01$, $p = .02$, $\eta_p^2 = .03$. Critically, the Current Item × Next Item interaction that was significant in Experiment 3 (as reported above), was non-significant in Experiment 1, $F(1, 49) = 1.02$, $MSE = 0.01$, $p = .32$, $\eta_p^2 = .02$. In Experiment 1, neither silent hits, $t(49) = 0.37$, $p = .71$, $d = 0.04$, nor aloud hits, $t(49) = 0.87$, $p = .39$, $d = 0.11$, were significantly influenced by the condition of the next study item.

It is also worth noting that in this three-way ANOVA the main effect of Experiment was non-significant ($F < 1$), as was the Experiment × Current Item interaction, $F(1, 148) = 1.36$, $MSE = 0.02$, $p = .25$, $\eta_p^2 = .01$. All other main effects and interactions were also non-significant ($ps > .05$), except for the Current Item main effect of production ($p < .001$). Thus, the presence of the underlines denoting study condition in Experiment 3 did not alter the overall mixed-list production effect from that seen in Experiment 1. In fact, overall hits were remarkably similar across the two experiments (see Table 2). The difference in aloud hits was non-significant, $t(148) = 1.20$, $p = .23$, $d = 0.20$, as was the difference in silent hits, $t(148) = 0.14$, $p = .89$, $d = 0.03$. [FAs were also non-significantly different across Experiment 1 ($M = .18$, $SE = .02$) and Experiment 3 ($M = .20$, $SE = .01$), $t(148) = 0.91$, $p = .37$, $d = 0.16$.]

Last, we explored whether the condition of the previous study item might also influence hits. When silent item hit rate (see Table 3) was the dependent measure, an Experiment (1 vs. 3) × Previous Item (aloud vs. silent) × Next Item (aloud vs. silent) ANOVA revealed non-significant main effects of Experiment, Previous Item, and Next Item [$F < 1$; $F(1, 148) = 3.07$, $MSE = .03$, $p = .08$, $\eta_p^2 = .02$; and $F(1, 148) = 1.93$, $MSE = .02$, $p = .17$, $\eta^2 = .01$, respectively]. The only statistically significant interaction (all others were $p > .70$) was the Experiment × Next Item interaction, $F(1, 148) = 6.79$, $MSE = .02$, $p = .01$, $\eta_p^2 = .04$, again confirming that the condition of the next item influenced memory only for silent items in Experiment 3 (when participants could foresee the upcoming study conditions). Unsurprisingly, when aloud hit rate was the dependent measure, the ANOVA yielded uniformly non-significant main effects and interactions (all ps $> .15$). Put simply, in both experiments, aloud hits were influenced neither by the previous condition nor by the next condition.

To summarize the cross-experiment analyses: In keeping with our performance anticipation account, the condition of the next study item influenced memory for silent items (i.e., worse memory when an aloud item was next) only when participants could see the upcoming study condition (Experiment 3). When that information was not shown to participants (Experiment 1), the condition of the next study item did not significantly influence memory for silent (or aloud) items. (The condition of the previous study item did not significantly influence memory in either experiment, or when the data from both experiments were combined.) Notably, the overall patterns of means across these two experiments were strikingly similar (see Table 2), which supports our view that the results of Experiment 3 (suggesting that performance anticipation reduces memory) can reasonably account for the "blanket" cost that broadly affected silent items in a mixed list in Experiment 1.

The fact that memory for silent items was uniformly poor in Experiment 1—regardless of the condition of the next study item—is inconsistent with the notion that the process of reading aloud retroactively interfered with the encoding (or retrieval) of prior silent study items. This result aligns with Bond and Kirkpatrick's (1982) prior work showing that "retroactive amnesia" (Tulving, 1969) does not account for the next-in-line effect.

Experiment 4: An absent experimenter

Research on the next-in-line effect (Brenner, 1973) suggests that performance anticipation is engendered by self-presentation concerns and other social factors, which draw attention away from encoding ongoing information. Consistent with this view, Bond (1985) found the

[6] A similar comparison can be made in Experiment 2 by comparing hits for silent position 5 ($M = .560$, $SE = .03$)—the silent items that immediately preceded aloud items—to the average hit rate across the first four silent positions ($M = .625$, $SE = .02$). This difference in means is statistically significant, $t(49) = 2.68$, $p < .01$, which affirms that Experiments 2 and 3 yielded consistent results.

next-in-line effect to be larger for socially anxious (vs. non-anxious) individuals. In mixed-list production research (and other basic laboratory memory studies), the presence of the experimenter constitutes an inherent social factor. If the presence of the experimenter—who participants may perceive as an expert evaluating their performance—contributes to performance anticipation, then having the experimenter absent for the duration of the experiment could eliminate the cost to silent items that precede aloud items (or yield a trivially small effect size). We explored this possibility by conducting a replication of Experiment 3 in which the experimenter was absent.

Method

Participants

One hundred undergraduate students from the University of Waterloo participated for course credit; this sample size was equivalent to that of Experiment 3.

Stimuli and apparatus

These were the same as in Experiment 3.

Procedure

The procedure was identical to that of Experiment 3, except that the experimenter left the room immediately after going over the study instructions with the participant (which followed the same protocol as in Experiments 1–3). The experimenter was absent for the duration of the experiment (both the study phase and recognition test).

Results

Table 2 shows the hits for aloud and silent study items as a function of whether they were immediately followed by an aloud item or a silent item. The mean FA rate was .21 (SE = .01).

A 2 (Current Item: aloud vs. silent) × 2 (Next Item: aloud vs. silent) repeated measures ANOVA revealed a significant main effect of Current Item, $F(1, 99) = 85.43$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .46$, signifying a robust production effect. However, the main effect of Next Item was nonsignificant ($F < 1$), and so was the Current Item × Next Item interaction, $F(1, 99) = 2.05$, $MSE = 0.01$, $p = .16$, $\eta_p^2 = .02$. The hit rate for silent items followed by aloud items was only 1 percentage point lower than the hit rate for silent items followed by silent items, a non-significant difference, $t(99) = 0.87$, $p = .38$, $d = .08$. Moreover, the Bayesian approximation procedure yielded “positive” support for a null effect, pBIC = .87. Thus, there was no evidence of a performance anticipation decrement on memory. Aloud hits were also unaffected by the condition of the next item, $t(99) = 1.21$, $p = .23$, $d = .10$. Combined analyses for Experiments 3 and 4 are reported in Appendix A.

Discussion

Experiment 4 replicated the method of Experiment 3—a mixed-list production experiment in which participants could foresee the upcoming study conditions—with only one difference: The experimenter was absent. In contrast to Experiment 3, in which hits were lower for silent items that were followed by aloud items relative to silent items that were followed by silent items (p = .001; d = .28), this difference was non-significant in Experiment 4 (p = .38; d = .08); the effect size was miniscule and a Bayesian analysis supported the null (pBIC = .87). This pattern of results is consistent with our performance anticipation account, insofar as the experimenter constitutes a social presence that, in keeping with next-in-line research (Brenner, 1973), may engender performance anticipation in participants. In particular, when reading an upcoming study item aloud, participants’ attention may be diverted from studying silent items by the impending prospect of being evaluated by the experimenter. Removing the experimenter may have eliminated such evaluative concerns, thereby mitigating performance anticipation.

That said, the fact that hit rates for aloud items were 7 percentage points lower when the experimenter was absent (Experiment 4; M = .70) vs. present (Experiment 3; M = .77) raises an alternate explanation: Participants may have been less likely to comply with reading words aloud when the experimenter was absent. A decrease in the frequency of performance (i.e., reading aloud) could also account for a decrease in performance anticipation (and the in concomitant cost to silent items that precede aloud items). In sum, participants’ recognition of silent items that preceded aloud items may have been worse in Experiment 3 because the experimenter constituted a social/evaluative presence who engendered performance anticipation. Alternately, participants could have performed less often in Experiment 4, which might also explain reduced performance anticipation. In either case, the memory deficit for silent items is reduced because performance anticipation is reduced.

General discussion

The prospect of one's upcoming performance (e.g., giving a talk at a conference) seems to diminish memory for preceding information (i.e., the prior talk). Prior research (Bond & Kirkpatrick, 1982; Bond et al., 1991; Bond, 1985; Brenner, 1973) has demonstrated this pre-performance memory deficit using a task in which the members of a large group of participants took turns reading words aloud: Individuals had poorer memory for words that they heard shortly before those that they read aloud—a “next-in-line” effect (Brenner, 1973). We propose that the negative effect of performance anticipation on memory could drive the cost that has been observed in the control conditions in within-subject memory research (McDaniel & Bugg, 2008), including in particular the cost in recognition for reading silently in mixed lists obtained in recent production effect research (Bodner et al., 2014; Forrin et al., 2016).

In Experiment 1, we replicated the mixed-list cost of production: Individuals had poorer recognition for silent items studied in a mixed list relative to those in a pure-silent list. In Experiment 2, we assessed whether performance anticipation reduced memory for silent items. We did this by having a mixed list follow a predictable pattern (alternating between blocks of five aloud items and five silent items), such that participants could anticipate the aloud items. Participants' recognition of silent items decreased over the five block positions as the next aloud block drew nearer, consistent with performance anticipation being detrimental to memory. In Experiment 3, we also gave participants foreknowledge of when aloud items would occur—using underlines to denote upcoming (and previous) study conditions—while still adhering to a *randomly* mixed list design (in which the cost of production occurs). Memory for silent items was worse when an aloud item (vs. a silent item) was next in the list, again suggesting that performance anticipation reduced memory for silent items. In contrast, the condition of the previous item did not affect memory for the silent items (and memory for the aloud items was entirely unaffected by the condition of prior or subsequent items). Importantly, there was no pre-performance memory deficit in Experiment 1, which rules out the alternate possibility that reading aloud retroactively reduces memory for prior silent items. Removing the experimenter (Experiment 4) also virtually eliminated (d = .08) the pre-performance memory deficit, which suggests that an audience (even a lone experimenter) may be required for performance anticipation—and for the corresponding cost to pre-performance memory—to occur.

Performance anticipation impairs memory

Taken together, the present experiments yield novel evidence that performance anticipation reduces memory for silent items in the context of a straightforward laboratory memory experiment. These results fit nicely with Brenner’s (1973) next-in-line effect. Whereas Brenner found that anticipation of reading aloud to an audience of peers imposed a cost on pre-performance memory, our results suggest that a multiple-person audience is not necessary: The anticipation of reading aloud (in the presence of an experimenter) is sufficient to yield this cost.

This pre-performance cost in memory may account for the cost apparent for items read silently in mixed-list production research (including our Experiment 1). The unpredictable occurrence of aloud items in a mixed list may cause participants to suffer persistent performance anticipation while studying silent items, thereby imposing a "blanket cost" on memory for silent items (see Bond & Kirkpatrick, 1982, for an analogous finding in the next-in-line literature). The results of Experiment 3 are entirely consistent with this possibility: Previewing upcoming study conditions resulted in worse memory for silent items followed by aloud items (relatively high anticipation) than for silent items followed by silent items (relatively low anticipation). Furthermore, aside from this study order effect, overall memory performance was comparable across Experiments 1 and 3, which supports our claim that the results of Experiment 3 are applicable to memory performance in a typical mixed-list production experiment.

To be clear, we do not claim that performance anticipation is the *only* factor responsible for the memorial cost of reading silently in a mixed list. Prior research (Forrin & MacLeod, 2016; Jonker, Levene, & MacLeod, 2014) has demonstrated another factor that contributes to the robust cost in free recall: Participants have inferior order memory for silent items that are studied in a mixed list (vs. a pure list), consistent with the item-order account (Nairne, Riegler, & Serra, 1991). While performance anticipation may also account for the cost in free recall, we suspect that it is a larger contributor to the cost in recognition because order information is less useful for guiding memory on a recognition test (see McDaniel & Bugg, 2008), in which test words typically appear randomly, in an order determined not by the participant but by the researcher. It is also worth noting that the observed memory reduction due to performance anticipation can only explain the *cost* to silent items in a mixed-list production experiment. The within-subject production effect has also been found to reflect a *benefit* to reading aloud (Forrin et al., 2016), which has been attributed to the distinctiveness conferred on aloud study items (see MacLeod & Bodner, 2017, for a review).

Performance anticipation as a type of 'lazy' reading

In the introduction, we posited that performance anticipation represents one factor that could lead participants to pay less attention to silent items in a mixed-list production experiment (and may thereby account for the mixed-list cost of production). That is, performance anticipation may constitute a type of context-specific "lazy reading" (Begg & Snider, 1987): Participants lazily read (i.e., shallowly process) silent items when they anticipate reading aloud. Although the present results are consistent with our performance anticipation account, it may well be the case that additional factors lead to the shallow processing of silent items (e.g., participants deem silent items to be less important or demanding than aloud items). Nevertheless, the present results suggest that these other potential factors are not sufficient in explaining lazy reading in a mixed-list production experiment: If participants uniformly paid less attention to silent (vs. aloud) items, then we ought to have observed uniformly poor memory for silent items (regardless of whether participants could foresee the occurrence of aloud items). Instead, we found that this foreknowledge led to worse memory for silent items only when they preceded aloud items and not when they preceded other silent items, which suggests that performance anticipation diverts attention from silent items.

Notably, the present results are compatible with Forrin et al.'s (2014) finding that the mixed-list production effect is equivalently large for generated words and for read words. Although having participants generate words silently in that experiment ensured that those words were not completely ignored, participants likely still paid less attention to words that they generated silently due to the prospect of an upcoming performance—generating aloud. Thus, the prospect of generating aloud may have diverted attention from generating silently in that experiment—perhaps to a similar extent as the prospect of reading aloud diverts attention from reading silently. In sum, a novel theoretical contribution of this research is that it elucidates a specific condition under which shallow processing occurs: due to the diversion of attention when participants anticipate a future performance.

Why does performance anticipation impair memory?

Brenner (1973) explained the next-in-line effect in terms of a participant having "incompatible demands on his attention" (p. 322). Bond et al. (1991) similarly concluded that "anticipation diverts cognitive resources from memorization" (p. 436). In the case of a mixed-list production experiment, we posit that attention is diverted from reading silently by the prospect of reading aloud, resulting in shallow processing of silent items. In contrast, while studying a pure-silent list, individuals cannot have their attention diverted by the prospect of reading aloud.

Several findings from the mind-wandering literature (see Smallwood & Schooler, 2006, for a review) are consistent with this "attention diversion" explanation. First, mind-wandering worsens performance on a wide variety of tasks, including reading (e.g., Franklin, Smallwood, & Schooler, 2011; Feng, D'Mello, & Graesser, 2013; Forrin, Risko, & Smilek, 2017; Schooler, Reichle, & Halpern, 2004; Smallwood, McSpadden, & Schooler, 2008; Smilek, Carriere, & Cheyne, 2010). Second, mind-wandering is more frequent when task demands are low (Baird, Smallwood, & Schooler, 2011; Smallwood, Nind, & O'Connor, 2009), as would be the case in silent reading. (Correspondingly, individuals would be less likely to mind-wander while reading aloud.) And third, individuals frequently mind-wander about current concerns (Klinger, 1978, 1999). In experiments requiring an overt response, a particularly salient current concern may well be their impending performance. In a mixed-list production experiment, individuals may have brief performance-related thoughts while reading silently, such as "get ready," "speak clearly," or "don't mess up." It might be possible in future to test this hypothesis using thought probes (e.g., Antrobus, 1968; Giambra, 1995; Schooler et al., 2004).

Mind-wandering that pertains to one's performance of the current task has been called *task-related interference* (e.g., Smallwood et al., 2004; Smallwood, Riby, Heim, & Davies, 2006). Smallwood et al. (2004) found that individuals mind-wander retrospectively about their task errors, which can worsen ongoing task performance. They concluded that "a pre-occupation with one's task performance […] seems to be best conceptualised as a strategic attempt to deploy attentional resources in response to a perception of environmental demands which exceed one's ability to perform the task" (p. 657). In accord with this view, we propose that individuals also mind-wander *prospectively* about upcoming task demands (e.g., a performance such as reading aloud), at the expense of ongoing task performance (e.g., encoding silent items). This *prospective task-related interference* should increase as the relatively more demanding task component draws nearer, thereby increasing the cost to the less demanding component. This explanation fits with the pattern of decreasing recognition of silent items as aloud items draw nearer (Experiment 2) as well as with the poorer memory for silent items that precede aloud (vs. silent) items (Experiment 3).

It is notable that research on prospective mind-wandering has, thus far, focused on *task-unrelated* thoughts (e.g., D'Argembeau, Renaud, & Van der Linden, 2011; Smallwood et al., 2009). For example, Baird et al. (2011; see also Stawarczyk, Cassol, & D'Argembeau, 2013) found that the content of prospective mind-wandering often involves the planning of personally-relevant, goal-directed behavior (in accord with Klinger's current concern theory, 1978, 1999). Likewise, we propose that prospective *task-related* interference may be particularly likely to occur when an upcoming component of the task evokes concerns (e.g., performing). That said, an individual's attention may be drawn to an upcoming performance even in the absence of performance-related concerns. Even a non-concerning performance may be perceived as more demanding (due to being active) than passive study (e.g., silent reading). And given that self-referential information (e.g., one's name) tends to divert attention from an ongoing task (e.g., Moray, 1959), the self-referential nature of an upcoming performance may be similarly distracting. In short, we suggest the intriguing possibility that the active, self-referential nature of performance—two components that underlie the benefit of production (Forrin & MacLeod, 2018)—may also drive its cost by diverting attention from silent items. Indeed, attention diversion may even occur in the absence of any conscious thoughts about one's impending performance: Rather, one's attention may spontaneously shift in preparation for performance.

There is another possible account for why performance anticipation reduces memory. The prospect not just of speaking but of speaking in front of the experimenter may have increased participants' anxiety to a degree that was detrimental to memory. Indeed, public speaking phobia is a pervasive, clinically-documented phenomenon (Paul, 1966) that is evident both in self-reports (Trexler & Karst, 1972) and in physiological measures (Knight & Borden, 1979). A large body of research has demonstrated that increased anxiety/arousal is detrimental to encoding (see, e.g., Moran, 2016, for a review). Although small audiences tend to engender less performance anxiety than larger audiences (Latané & Harkins, 1976), an expert audience tends to increase performance anxiety (Geen & Gange, 1977), and participants may perceive the experimenter as an expert who is scrutinizing their performance (increased scrutiny has also been found to reduce memory; see Kimble & Zehr, 1982; Lord, Saenz, & Godfrey, 1987). This could explain why, when the experimenter was absent in Experiment 4, the memory deficit to silent items that preceded aloud items was no longer observed (d = .08).

Participants' memory for aloud items may similarly be impaired by performance anxiety, but this would not influence the mixed-list *cost* of production, which is assessed in terms of the difference in memory between silent items studied in a mixed list and those studied in a pure list (Bodner et al., 2014). Individuals would clearly not experience performance anxiety while studying silent items in a pure list. So, in a mixed list, even if both aloud and silent items were negatively affected by performance anxiety, this would not alter the magnitude of the cost. And we would still attribute the better memory for aloud items to an advantage conferred by distinctiveness (e.g., MacLeod et al., 2010).

In sum, we have proposed two mechanisms by which performance anticipation may diminish memory for ongoing information: (1) diverting attention (i.e., prospective task-related interference), and (2) eliciting anxiety. That said, these two mechanisms could well be connected: Attention diversion could be the active ingredient—the mechanism that anxiety invokes—thereby underpinning the anxiety account. Indeed, Mrazek et al. (2011) found that mind-wandering mediates the relation between anxiety and math performance, so it could similarly mediate a relation between anxiety and memory performance.

To what extent does the performance component affect anticipation?

It is important to clarify that the performance itself does not reduce memory for pre-performance information. If it did, we would have observed memory deficits for silent items that were followed by aloud items in Experiment 1 (a standard mixed-list production experiment). Rather, the factor that harms pre-performance memory is *anticipation* of performance. The expectation of a future event (i.e., anticipation) can elicit an emotional response (Castelfranchi & Miceli, 2011). In our research, we have surmised that the expectation of reading aloud evokes anxiety that may impede memorization. But not all upcoming performances induce anxiety. For example, the expectation of an enjoyable future event could lead to positive emotions (which may enhance memory; for a review see Isen, 1987). Nevertheless, the expectation of a positive event could still reduce memory via *attention diversion*.

We therefore propose that a future event reduces memory for ongoing current events when the individual appraises the future event as more important—and thus more deserving of attention—than the current events. Several factors could result in an event attaining priority status, including—but certainly not limited to—a personal performance. Upcoming events that are demanding, personally relevant, distinctive (and so on) may also be designated "high priority," as would be true for events that participants infer are important within the context of the experiment.

Conclusion

Expanding Brenner (1973) finding that performance anticipation reduces memory for ongoing information in social settings, the present results suggest that performance anticipation is also detrimental in the context of mixed-list memory experiments (at least those in which the experimenter is present), and may readily account for the mixed-list costs observed for production (e.g., Bodner et al., 2014). Performance anticipation may also contribute to memorial costs that have been reported for other encoding techniques (see McDaniel & Bugg, 2008). Even when both conditions in a mixed list involve overt performance (e.g., the generation effect when the read condition is aloud; Slamecka & Graf, 1978) participants may deem one of the conditions to involve a more demanding performance and hence to have a higher priority (in the case of generation, perhaps due to the evaluative component) which would in turn result in heightened anticipation for that condition.

Anticipating an upcoming performance may impose a particularly large cost on memory by evoking concerns that divert attention from the ongoing information and by eliciting anxiety that degrades encoding. Certainly, however, there are other "high priority" events that individuals anticipate that do not involve performing, yet that still divert attention. Anticipation is therefore a crucial factor for researchers to consider when using a within-subject design: Individuals' performance on a trial involving one task may well be influenced by their anticipation of a trial involving a forthcoming different task.

Appendix A. Combined Experiment 3 and 4 analyses

Overview

Prior to reporting combined analyses of Experiments 3 and 4, we briefly outline the background of Experiment 4, our aim being to explain why these combined analyses should be considered as supplemental and exploratory. In fact, reviewers of our initial submission suggested that we conduct an experiment in which the experimenter was absent to assess whether the social presence of an experimenter was required to observe a cost in memory to silent items that preceded aloud items. We first considered conducting an experiment that would test the hypothesis that the absence (vs. presence) of an experimenter would result in a significantly smaller pre-performance cost to silent items. The appropriate way to test this hypothesis would be to manipulate the presence vs. absence of the experimenter between-subjects, with random assignment to the experimenter/no experimenter conditions, focusing on the critical interaction. (It would not have been appropriate to use participants' data from Experiment 3, since those data were collected prior to forming this hypothesis.)

We therefore conducted an *a priori* power analysis to determine an appropriate sample size for such an experiment. Given the small effect size (d = .28) that we had obtained in Experiment 3, we presumed that the hypothesized interaction effect would be small (even assuming a null effect when the experimenter was absent, which would be the best-case scenario for detecting an interaction effect). G*Power 3.0 revealed that 194 participants were required in each condition to have adequate statistical power (.80) to detect a small interaction effect (η_p^2 = .02) at a significance criterion of α = .05. Considering our resource limitations, we decided that carrying out an experiment with samples of this size was not possible.

Instead, we concluded that our best available option was to run a fairly large sample (n = 100; like Experiment 3) with the experimenter absent. This became Experiment 4. Although consequently not designed to test for the aforementioned interaction effect, Experiment 4 still enabled us to obtain an effect size estimate for the cost in recognition to silent items when the experimenter was absent. Moreover, we could also use Bayesian statistics to assess the strength of the evidence supporting a null next-in-line cost when the experimenter was absent.

Keeping in mind the issues just raised, we decided to include this (supplemental) combined analysis section for readers who might be interested in these results. To summarize, these analyses were intended to be exploratory given that (i) Experiment 4 was not intended to test the hypothesis that the presence (vs. absence) of the experimenter modulates the size of the next-in-line cost to silent items, and (ii) we did not have adequate statistical power to reliably detect this interaction effect in a direct comparison of Experiment 3 and 4.

Analyses

The false alarm rates in Experiment 3 (M = .20) and Experiment 4 (M = .21) were virtually identical, $t(198) = 0.43$, $p = .67$, $d = .06$. An Experiment (3 vs. 4) × Current Item (aloud vs. silent) × Next Item (aloud vs. silent) ANOVA revealed a non-significant main effect of Experiment, $F(1, 198) = 1.37$, $MSE = 0.07$, $p = .24$, $\eta_p^2 = .01$, signifying that overall hit rate for silent items was similar in the two experiments. The main effect of Current Item was significant, $F(1, 198) = 312.47$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .61$, reflecting a robust overall production effect. The main effect of Next Item was marginal, $F(1, 198) = 2.84$, $MSE = 0.01$, $p = .09$, $\eta_p^2 = .01$, due to hits being slightly lower overall for study items that were followed by aloud items. Notably, the Experiment × Current Item interaction was significant, $F(1, 198) = 34.23$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .15$, reflecting a significantly larger production effect in Experiment 3 than in Experiment 4. However, the Experiment × Current Item × Next Item interaction was non-significant ($F < 1$).

Given that the three-way interaction was non-significant, the following results should be interpreted cautiously. Because we were specifically interested in whether the *silent* items would be affected by removing the experimenter, we ran a follow-up Experiment × Next Item ANOVA that included hit rates for silent items only. Of main interest, the Experiment × Next Item interaction was marginally significant, $F(1, 198) = 2.97$, $MSE = 0.01$, $p = .086$, $\eta_p^2 = .01$; the effect of the next study condition on silent items was larger in Experiment 3 (experimenter present) than in Experiment 4 (experimenter absent). Hits for silent items that were followed by aloud items were significantly lower in Experiment 3 (M = .53) than in Experiment 4, (M = .58), $t(198) = 2.13$, $p = .034$, $d = .30$.

For aloud items, the Experiment × Next Item ANOVA revealed a non-significant main effect of Next Item and a non-significant Experiment × Next Item interaction (Fs < 1). Hits for aloud items were not influenced by the next study condition in either experiment. There was, however, a significant main effect of Experiment, $F(1, 198) = 14.79$, $MSE = 0.04$, $p < .001$, $\eta_p^2 = .07$, reflecting more aloud hits in Experiment 3 than in Experiment 4.

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2019.104050.

References

Antrobus, J. S. (1968). Information theory and stimulus-independent thought. *British Journal of Psychology, 59*, 423–430.

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition, 20*, 1604–1611.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*, 3–26.

Begg, I., & Snider, A. (1987). The generation effect: Evidence for generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 553–563.

Bodner, G. E., Taikh, A., & Fawcett, J. M. (2014). Assessing the costs and benefits of production in recognition. *Psychonomic Bulletin & Review, 21*, 149–154.

Bond, C. F., Jr. (1985). The next-in-line effect: Encoding or retrieval deficit? *Journal of Personality and Social Psychology, 48*, 853–862.

Bond, C. F., Jr., & Kirkpatrick, C. K. (1982). Distraction, amnesia, and the next-in-line effect. *Journal of Experimental Social Psychology, 18*, 307–323.

Bond, C. F., Jr., & Omar, A. S. (1990). Social anxiety, state dependence, and the next-in-line effect. *Journal of Experimental Social Psychology, 26*, 185–198.

Bond, C. F., Jr., Pitre, U., & van Leeuwen, M. D. (1991). Encoding operations and the next-in-line effect. *Personality and Social Psychology Bulletin, 17*, 435–441.

Brenner, M. (1973). The next-in-line effect. *Journal of Verbal Learning and Verbal Behavior, 12*, 320–323.

Brown, A. S., & Oxman, M. (1978). Learning through participation: Effects of involvement and anticipation of involvement. *American Journal of Psychology, 91*, 461–472.

Castelfranchi, C., & Miceli, M. (2011). Anticipation and emotion. In P. Petta, C. Pelachaud, & R. Cowie (Eds.). *Emotion-oriented systems* (pp. 483–500). Heidelberg: Springer.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language, 26*, 341–361.

D'Argembeau, A., Renaud, O., & Van der Linden, M. (2011). Frequency, characteristics and functions of future-oriented thoughts in daily life. *Applied Cognitive Psychology, 25*, 96–103.

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.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments & Computers, 28*, 1–11.

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 671.

Fawcett, J. M., & Ozubko, J. D. (2016). Familiarity, but not recollection, supports the between-subject production effect in recognition memory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 70*, 99–115.

Fawcett, J. M., Quinlan, C. K., & Taylor, T. L. (2012). Interplay of the production and picture superiority effects: A signal detection analysis. *Memory, 20*, 655–666.

Fawcett, J. M. (2013). The production effect benefits performance in between-subject designs: A meta-analysis. *Acta Psychologica, 142*, 1–5.

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review, 20*, 586–592.

Forrin, N. D., Groot, B., & MacLeod, C. M. (2016). The d-Prime directive: Assessing costs and benefits in recognition by dissociating mixed-list false alarm rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1090–1111.

Forrin, N. D., Jonker, T. R., & MacLeod, C. M. (2014). Production improves memory equivalently following elaborative vs non-elaborative processing. *Memory, 22*, 470–480.

Forrin, N. D., & MacLeod, C. M. (2016). Order information is used to guide recall of long lists: Further evidence for the item-order account. *Canadian Journal of Experimental Psychology, 70*, 125–138.

Forrin, N. D., & MacLeod, C. M. (2018). This time it's personal: The memory benefit of hearing oneself. *Memory, 26*, 574–579.

Forrin, N. D., Risko, E. F., & Smilek, D. (2017). On the relation between reading difficulty and mind-wandering: A section-length account. *Psychological Research, 83*, 485–497.

Franklin, M. S., Mooneyham, B. W., Baird, B., & Schooler, J. W. (2014). Thinking one thing, saying another: The behavioral correlates of mind-wandering while reading aloud. *Psychonomic Bulletin & Review, 21*, 205–210.

Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic bulletin & review, 18*, 992–997.

Geen, R. G., & Gange, J. J. (1977). Drive theory of social facilitation: Twelve years of theory and research. *Psychological Bulletin, 84*, 1267–1288.

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition, 4*, 1–21.

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt, & J. B. Worthen (Eds.). *Distinctiveness and memory* (pp. 1–25). New York: Oxford University Press.

Hunt, R. R. (2013). Precision in memory through distinctive processing. *Current Directions in Psychological Science, 22*, 10–15.

Innes, J. M. (1982). The next-in-line effect and the recall of structured and unstructured material. *British Journal of Social Psychology, 21*, 1–5.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. *Advances in Experimental Social Psychology, 20*, 203–253.

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review, 4*, 577–581.

Jones, A. C., & Pyc, M. A. (2014). The production effect: Costs and benefits in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 300–305.

Jonker, T. R., Levene, M., & MacLeod, C. M. (2014). Testing the item-order account of design effects using the production effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 441–448.

Kimble, C. E., & Zehr, H. D. (1982). Self-consciousness, information load, self-presentation, and memory in a social situation. *The Journal of Social Psychology, 118*, 39–46.
Klinger, E. (1978). Modes of normal conscious flow. In K. S. Pope, & J. L. Singer (Eds.). *The stream of consciousness* (pp. 225–258). New York: Plenum.
Klinger, E. (1999). Thought flow: Properties and mechanisms underlying shifts in content. In J. A. Singer, & P. Salovey (Eds.). *At play in the fields of consciousness* (pp. 29–50). Mahwah, NJ: Lawrence Erlbaum Associates.
Knight, M. L., & Borden, R. J. (1979). Autonomic and affective reactions of high and low socially-anxious individuals awaiting public performance. *Psychophysiology, 16*, 209–213.
Lambert, A. M., Bodner, G. E., & Taikh, A. (2016). The production effect in long-list recall: In no particular order? *Canadian Journal of Experimental Psychology, 70*, 165–176.
Latané, B., & Harkins, S. (1976). Cross-modality matches suggest anticipated stage fright is a multiplicative power function of audience size and status. *Perception & Psychophysics, 20*, 482–488.
Lord, C. G., Saenz, D. S., & Godfrey, D. K. (1987). Effects of perceived scrutiny on participant memory for social interactions. *Journal of Experimental Social Psychology, 23*, 498–517.
MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review, 18*, 1197–1202.
MacLeod, C. M., & Bodner, G. E. (2017). The production effect in memory. *Current Directions in Psychological Science, 26*, 390–395.
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.
Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*, 679–690.
McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review, 15*, 237–255.
Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin, 142*, 831–864.
Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56–60.
Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology, 47*, 1243–1248.
Murdock, B. B., Jr (1962). The serial position effect of free recall. *Journal of experimental psychology, 64*, 482–488.
Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 702–709.
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.
Ozubko, J. D., Major, J., & MacLeod, C. M. (2014). Remembered study mode: Support for the distinctiveness account of the production effect. *Memory, 22*, 509–524.
Paul, G. L. (1966). Insight vs. desensitization in psychotherapy. An experiment in anxiety reduction. *The American Journal of the Medical Sciences, 252*, 500.
Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.). *Sociological methodology 1995* (pp. 111–196). Cambridge, MA: Blackwell.
Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.). *Thinking and seeing: Visual metacognition in adults and children* (pp. 203–226). Cambridge, MA: MIT Press.
Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory, 4*, 592–604.
Smallwood, J., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic sustained attention to response task. *Consciousness and Cognition, 15*, 218–231.
Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and cognition, 13*, 657–690.
Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36*, 1144–1150.
Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition, 18*, 118–125.
Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*, 946–958.
Smilek, D., Carriere, J. S., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science, 21*, 786–789.
Sousa, T. L. V., Carriere, J. S., & Smilek, D. (2013). The way we encounter reading material influences how frequently we mind wander. *Frontiers in Psychology, 4*, 892.
Stawarczyk, D., Cassol, H., & D'Argembeau, A. (2013). Phenomenology of future-oriented mind-wandering episodes. *Frontiers in Psychology, 4*, 425.
Taikh, A., & Bodner, G. E. (2016). Evaluating the basis of the between-group production effect in recognition. *Canadian Journal of Experimental Psychology, 70*, 186–194.
Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Teacher's College, Columbia University.
Trexler, L. D., & Karst, T. O. (1972). Rational-emotive therapy, placebo, and no-treatment effects on public-speaking anxiety. *Journal of Abnormal Psychology, 79*, 60–67.
Tulving, E. (1969). Retrograde amnesia in free recall. *Science, 164*, 88–90.
Vogel, S., & Schwabe, L. (2016). Learning and memory under stress: Implications for the classroom. *npj Science of Learning, 1*, 16011. https://doi.org/10.1038/npjscilearn.2016.11.
Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14*, 779–804.
Walker, B. S., & Orr, F. E. (1976). Anxiety and the next-in-line effect. *Journal of Educational Psychology, 68*, 775.