



The spacing effect stands up to big data

A. S. N. Kim^{1,2} · A. M. B. Wong-Kee-You¹ · M. Wiseheart^{1,3} · R. S. Rosenbaum^{1,2,4}

Published online: 8 January 2019
© The Psychonomic Society, Inc. 2019

Abstract

Many studies have shown that repetition of study material with temporal gaps between the repetitions (i.e., spaced in time) is more effective for long-term retention than are repetitions in immediate succession (i.e., massed; Greene, 1989). Although this *spacing effect* has proven to be robust in the laboratory (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006), its status in the real world is relatively understudied. Other research has demonstrated the benefit of memory retrieval on subsequent retrieval of the same information (Bjork, 1975, 1988; Roediger & Karpicke, 2006), referred to as the *testing effect*. However, it is not clear how spacing and retrieval can be optimally combined in order to enhance knowledge retention in a real-world setting. To investigate this question, we analyzed longitudinal data from 10,514 individuals, collected in the context of naturally occurring workplace training. To determine the impact of spaced retrieval on knowledge retention, these data were analyzed using a generalized linear mixed model with the following fixed factors: (1) spacing interval between repetitions of content training (retrieval practice), (2) retention interval, and (3) question format. Random factors included the specific content on which employees were trained, which was clustered by employee and, in turn, by company, resulting in a three-level hierarchy. The results showed a significant interaction between spacing interval and retention interval: the optimal amount of spacing between repeated retrieval events increased as the retention interval increased. These findings are in line with the results of laboratory studies, demonstrating the relevance and transferability of laboratory-based research to real-world contexts.

Keywords Episodic memory · Spacing effect · Distributed practice · Retrieval practice · Learning management system (LMS) · Training and development · Big data · Generalized linear mixed model (GLMM) · Real-world data

We strive to learn and remember information efficiently and effectively to achieve a variety of goals in everyday life, from a list of items to purchase at the grocery store and where we parked our car to the names of people we meet. We may attempt to do so through elaborate imagery (Bower, 1970; Paivio, 1969), associating newly learned material with prior knowledge (Brod, Werkle-Bergner, & Shing, 2013; Ryan, Moses, Barense, & Rosenbaum, 2013), or making the material relevant to ourselves (i.e.,

the self-reference effect; Carson, Murphy, Moscovitch, & Rosenbaum, 2016; Symons & Johnson, 1997). Even the simple and often spontaneous act of bringing this information to mind (retrieval practice) can be effective, especially when an interval is imposed between repeated presentations of the material (i.e., *spaced* repetition), as compared to when the repeated material is presented in immediate succession (i.e., *massed* repetition). The utility of this *spacing effect* as a learning technique has been evident in a vast number of laboratory experiments performed using a wide variety of methods in diverse populations (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006).

Extensive research has also demonstrated that the act of successful retrieval serves as a “memory modifier” to effectively increase the likelihood of successfully retrieving the same information at a future time (Bjork, 1975). The mere act of taking a test, or retrieving information from memory, improves subsequent retrieval, even more than passive repetition or restudy of the information; this is referred to as the *testing effect* (Bjork, 1975, 1988; Roediger & Karpicke,

✉ A. S. N. Kim
a_kim@yorku.ca

¹ Department of Psychology, York University, Toronto, Canada
² Rotman Research Institute, Baycrest Health Sciences, Toronto, Canada
³ LaMarsh Centre for Child and Youth Research, York University, Toronto, Canada
⁴ Vision: Science to Applications (VISTA) Program, York University, Toronto, Canada

2006). Although the effects of spacing and testing are both very well-supported by past studies (e.g., Gerbier & Toppino, 2015; Roediger & Karpicke, 2006), it is unclear how the two may be optimally combined to enhance knowledge retention in a real-world setting (although see the work by Storm, Bjork, & Storm, 2010). By analyzing a large, naturally occurring dataset of over 10,000 employees in the private sector (e.g., pharmaceutical and higher education companies and a large grocery retailer), the present study will help realize the potential of spaced retrieval as it extends to real-world learning in the workplace.

Attempts have been made to determine the efficacy of spaced repetition in the real world, primarily as it applies to classroom teaching (Kapler, Weston, & Wiseheart, 2015; Pashler et al., 2007; Sobel, Cepeda, & Kapler, 2011) and clinical interventions (Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Cermak et al., 1996; Green, Weston, Wiseheart, & Rosenbaum, 2014; Kim, Saberi, Wiseheart, & Rosenbaum, 2018), but its utility beyond these settings and in large, unconstrained population samples has yet to be determined. Workplace training could benefit from learning strategies that employ spaced repetition of information. Whether engaging in safety training or learning of new product information, workplace training often occurs in the context of varying levels of experience and motivation to learn, and changes in content and delivery of that content. Workplace training is also subject to unpredictable learning schedules in terms of time available to devote to learning in a given session or day, length of time between exposures to study material, and number of exposures to study material. This creates ideal conditions for examining the spacing effect and how to optimize it to promote learning in a heterogeneous, real-world context.

Here we analyzed longitudinal data that were collected in the context of workplace training. The data were collected via a learning management system (LMS) that delivers training content to employees across multiple sessions. An important feature embedded into the architecture of the LMS is multiple presentations of a given piece of information, with varying delays and number of intervening items between repetitions, resulting in training/learning sessions that are naturally spread (or spaced) apart and according to various schedules across individuals. In light of the extant literature, supported by numerous laboratory-based studies (for reviews see Cepeda et al., 2006; Gerbier & Toppino, 2015), we hypothesized that our large, real-world dataset would demonstrate a significant interaction between spacing and retention interval. More specifically, we hypothesized that the optimal spacing interval would increase as the retention interval increases. In this way, the present investigation was

hypothesis-driven based on past empirical work, unlike the more exploratory nature of most studies using big data.¹

Method

Axonify Inc.: The LMS

The data analyzed in the present study were collected using an LMS provided by Axonify Inc. As mentioned above, repeated and distributed retrieval practice of training content is integrated into the architecture of the LMS, allowing for investigation of spaced retrieval and its optimization using this large, real-world dataset. Through the LMS, client companies deliver training to employees across multiple sessions, with each session lasting approximately 3 min. In a typical session, employees answer two to three questions. When employees answer questions incorrectly, their learning is reinforced with corrective feedback. Thus, each training session was composed of questions and corrective feedback. The LMS offers the use of multiple question formats for content delivery, as described further below. The question formats used for training, along with the specific content delivered using the LMS (e.g., health and safety rules, product information), were selected at the discretion of the client companies.

Importantly, the LMS allows administrators to keep track of the performance of each individual employee, and thus identify the content for which individual employees need additional training. Training sessions are delivered on an individual basis, allowing content delivery to be personalized to individual employees so that training is focused on content that employees have yet to sufficiently acquire. Since the LMS platform enables the training process to be delivered to employees by way of computers, laptops, smartphones, point of sale systems, and security terminals, training can take place anywhere and at anytime. Typically, however, employees receive training in the workplace. Heterogeneity in the nature of the corresponding data comes with the real-world settings in which the data were collected. For example, some employees work on a part-time basis and consequently receive less training than full-time employees. Despite differences in the

¹ The use of big data in scientific investigations involves the “storing, retrieval, and analysis of large amounts of information” (Harlow & Oswald, 2016, p. 447), and as has been noted in the literature, the notion of what makes a dataset “big” is relative and ultimately depends on one’s perspective (Shiffrin, 2016; Wolfe, 2013). For example, past investigations of big data have included analyses conducted on 30,000 platelet transfusions (Guan et al., 2017) and 8,163,153 tweets (Thorstad & Wolff, 2018). In the present study, we conducted analyses on longitudinal data from 10,514 individuals, with approximately 220,000 data points. In the context of psychological and cognitive science research, where the number of participants included in a study typically falls well below a sample of 100 participants, we consider the present dataset to fall within the category of big data.

frequency of content delivery for a given employee and the nature of the content, all training delivered using the LMS implemented spaced retrieval and measured the impact it has on knowledge retention, which is central to the present investigation.

Data

We analyzed longitudinal data from 10,514 individuals (employees), collected in the context of workplace training from five client companies. Companies 1 and 5 are retailers, and, in both cases, the employees who received training are associates who work in their stores. Specifically, Company 1 is a high-end department store retailer, and employees received training on safety and store operations. Company 5, on the other hand, is a US grocery retailer, and employees received training on loss prevention, safety, and store operations. The data contributed by Companies 2 and 4 are from sales teams for two different pharmaceuticals companies, and the employees from both companies received training on product knowledge. Finally, Company 3 is a higher-education institution, and the corresponding data are from internal staff trained on training procedures. These individuals varied in age, gender, and occupation. Unfortunately, however, demographic data were not available for the present research, due to privacy policies and the nature of the data collection. Since we did not receive any personal or identifying information with the data, we were not required to receive informed consent from the employees whose data we analyzed or approval from our institutional ethics committees.

As mentioned above, the individual companies provided the content on which their employees were trained using the LMS and also selected the question format used throughout training (e.g., multiple choice, multiple answer, advanced multiple choice, matching, and/or fill in the blanks). Table 1 shows the breakdown (in proportions) of the question formats used by each company. The client companies also determined the number of unique questions (vs. occurrences of a question) delivered to an individual during training. Table 2 presents summary statistics for the number of unique questions

delivered to employees for each individual client company using the LMS described above. Table 3 presents a breakdown of the numbers of correct and incorrect responses on a first retrieval event (training section) as a function of question format and company. The Appendix lists sample questions for each of the question formats used by each company.

As we describe further below, our analyses only took into account individuals' performance on the first three occurrences of a question. Following the work of Cepeda et al. (2009), our analyses only considered data for which individuals provided an incorrect response to the first occurrence of a question, followed by a correct response for the second occurrence of a question (please refer to Fig. 1). We focused on items that were first answered incorrectly because we could not be certain whether those items that were first answered correctly were already known prior to the learning event in question. For this reason, we considered these data (i.e., items that were first answered correctly) to be potentially confounded, and their exclusion minimized the possibility that an individual employee was reviewing material that had already been learned, and possibly mastered, rather than learning new material. This type of confound does not apply in laboratory studies, since the type of information that is typically learned in laboratory studies is purposefully made to be random (e.g., pairs of words that are assembled to have little or no preexperimental relation). In the present study, however, the material was not random, and some of it is likely to have been previously known by the participant.

Because the focus of the present investigation was to assess the impact of spaced retrieval on learning new information, the abovementioned reasoning led us to exclude those items that were first answered correctly from our analyses. A correct response on the second occurrence of a question demonstrated that the participant had successfully learned the item. As we describe further below, responses to the third occurrence of a given question served as our response variable, which was dependent on a set of fixed and random factors. Aside from the criteria that we used to filter the data, we did not have any

Table 1 Proportions of question formats used by each client company

Company	MC	MA	AMC	Match	FIB	Sum
1	.83	.12	.02	.00	.02	1.00
2	.99	.00	.00	.01	.00	1.00
3	.72	.26	.00	.02	.00	1.00
4	.92	.02	.00	.06	.00	1.00
5	.53	.47	.00	.00	.00	1.00

MC = multiple choice, MA = multiple answer, AMC = advanced multiple choice, Match = matching, FIB = fill in the blanks

Table 2 Summary statistics for the number of unique questions presented to employees in each company

Company	Min	Max	Mean	Median	SD
1	1	58	8	5	8
2	1	12	4	4	3
3	1	119	13	6	18
4	1	11	3	2	3
5	1	18	7	7	4

Min = minimum, Max = maximum, SD = standard deviation

Table 3 Numbers of correct/incorrect responses on first retrieval event as a function of question format and client company

Company	Accuracy	MC	MA	AMC	Match	FIB	Total
1	Correct	421,074	10,609	23,071	1,244	3,601	459,599
	Incorrect	56,375	13,715	2,053	1,110	1,199	74,452
2	Correct	51,011	1,402	N/A	1,865	N/A	54,278
	Incorrect	13,695	1,009	N/A	1,207	N/A	15,911
3	Correct	23,199	3,561	N/A	23,199	N/A	49,959
	Incorrect	24,397	8,102	N/A	809	N/A	33,308
4	Correct	19,993	1,531	N/A	458	N/A	1,989
	Incorrect	4,959	1,270	N/A	329	N/A	6,558
5	Correct	15,274	2,371	N/A	N/A	N/A	17,645
	Incorrect	3,055	2,716	N/A	N/A	N/A	5,771

MC = multiple choice, MA = multiple answer, AMC = advanced multiple choice, Match = matching, FIB = fill in the blanks, N/A = not applicable

control over the amount of data we had to work with, and were thus unable to conduct a power analysis or otherwise to control sample size.

Model

In the present study, we investigated how to optimally distribute learning/training events by analyzing a large, real-world dataset. Our primary research interest centered on the impact of spacing and retention intervals on employees' knowledge retention, as well as on a potential interaction between these two variables. As a secondary research interest, we also investigated whether question format impacted employees' knowledge retention and whether this variable interacted with spacing and retention interval.

Both spacing and retention interval were used as fixed factors. Both of these fixed factors were continuous and quantified using number of days as a unit of measure. The

spacing interval corresponded to the time interval between the first and second occurrences of a question (O1 and O2, respectively; please refer to Fig. 1). The *retention interval* corresponded to the time interval between the second and third occurrences of a question (O2 and O3, respectively). Individuals' responses to a given question at O3 were taken as a measure of knowledge retention and dichotomously coded as being either correct (1) or incorrect (0). Since the duration of the spacing and retention intervals depended on the timing of O1 versus O2 and of O2 versus O3, respectively, we did not have any control over these timing-related variables. However, the inherent variability of these variables contributed to the scope of data accounted for by our model.

In addition to spacing and retention interval, question format was also treated as a fixed factor. This third fixed factor was categorical and consisted of five types of question format: multiple choice (MC), multiple answer (MA),

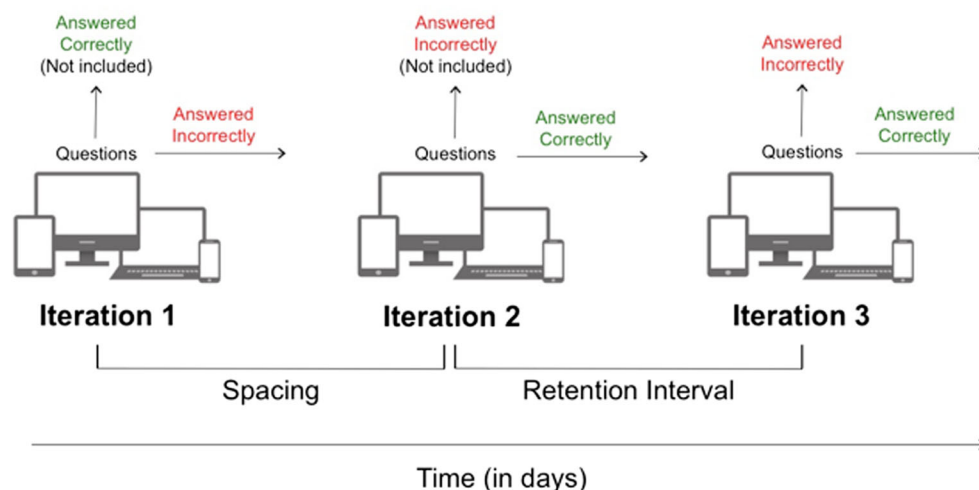


Fig. 1 The data considered in our analyses consisted of responses to questions to which individuals provided an incorrect response on its first occurrence, followed by a correct response on the second

occurrence of the question. Responses to the third occurrence of a given question served as our response variable, which was dependent on a set of fixed and random factors.

advanced multiple choice (AMC), fill in the blanks (FIB) and matching (Match). The MC and AMC formats required individuals to select one of multiple options provided to them as the most appropriate response to the given question. The difference between MC and AMC is that employees receive more detailed feedback for AMC than for MC questions; whereas employees receive a general explanation for incorrect responses to a MC question, for AMC questions they receive specific explanations outlining why each incorrect option is incorrect. For MA, multiple responses were provided and individuals were to select all (multiple) of the correct responses. For FIB and Match, individuals were tasked with filling in missing words and identifying appropriate pairings, respectively. Since question format was a categorical factor with five possible variables (MC, MA, AMC, FIB, and Match) and MC set as the reference, four binary variables were required to account for the effect of each variable. For example, to assess the effect of FIB, the binary variable for FIB would be set to one, MC serving as the reference, and the remaining three variables (MA, AMC, and Match) set to zero. In this way, a beta coefficient was derived for each of the four binary variables in the model.

The data in our analyses consisted of a three-level hierarchical structure; the questions (Level 1) clustered within a given individual (Level 2), who is, in turn, clustered within a specific company (Level 3). Responses of a given individual would therefore likely show some dependence, as would responses of individuals from a given company. To account for these dependencies, the following random factors were included in the model: the specific components of content that employees were trained on (e.g., question), which were clustered under employee (e.g., individual) and, in turn, under company, resulting in a three-level hierarchy.

A logistic generalized linear mixed model (GLMM) was fit to the employees' O3 response data. Spacing (or interval of time between O1 and O2), retention interval (or interval of time between O2 and O3) and question format were set as fixed predictor effects, whereas question, user and company were set as random effects. Our logistic GLMM model allowed us to examine the relation between a binary outcome variable (correct response = 1, incorrect response = 0) and our clustered predictor variables. The odds of a correct response was defined as the ratio between the probability of a correct response over the probability of an incorrect response. The logit transformation on the probabilities of a correct response was used to establish a linear relationship between our predictor variables and the odds of a correct response. Outliers were removed to meet the assumption of homogeneity of variance. All other model assumptions were met.

The odds of a correct response was defined as

$$\log \left(\frac{p_{ji}(x)}{1-p_{ji}(x)} \right) = \beta_o + \beta_{Spacing}X_{1ji} + \beta_{RI}X_{2ji} + \beta_{QF}X_{3ji} + \beta_{S:RI}X_{ji} + \beta_{RI:QF}X_{ji} + \gamma_j + \gamma_i + \gamma_k$$

where $p_{ji}(x)$ was the probability of a correct response, and $1 - p_{ji}(x)$ was the probability of an incorrect response. β_o refers to the intercept, $\beta_{Spacing}$ was the coefficient for spacing, β_{RI} was the coefficient for retention interval, β_{QF} was the coefficient for question format, $\beta_{S:RI}$ was the coefficient for the interaction between spacing interval and retention interval, $\beta_{RI:QF}$ was the coefficient for the interaction between retention interval and spacing interval, and γ_j , γ_i , and γ_k were the error terms for individual, company, and question, respectively.

Results

The results of our logistic GLMM, based on the penalized quasi-likelihood method, are presented in Table 4. Spacing, retention interval, and question format (specifically, AMC, MA, and Match) significantly affected the log-odds of a correct response to a given question within a training session (all z values significantly different from 0, $p < .005$). Second-order interactions were tested using the multivariate Wald test and revealed a significant interaction between spacing and retention intervals, as well as between retention interval and question format ($p < .05$).

Inverse log transformations of the predictor-variable estimates were conducted in order to interpret the results of our model as odds ratios. The transformed estimates can also be found in Table 4. Our results showed that when retention interval is set to zero and question format is set to multiple

Table 4 Summary of the results of the mixed model

Parameter	β	$\exp(\beta)$	SE	Statistic	p Value
Intercept	7.9468	2,826.6000	0.093	85.268	< .001
Spacing	-0.0216	0.9787	0.003	-7.978	< .001
RI	-0.0189	0.9812	0.001	-14.129	< .001
AMC	2.0938	8.1159	0.706	2.968	.003
FIB	-1.1260	0.3243	0.685	-1.644	.100
Match	-3.8089	0.0222	0.767	-4.966	< .001
MA	-1.1848	0.3058	0.182	-6.503	< .001
Spacing × RI	0.0003	1.0000	0.000	9.672	< .001
RI × AMC	0.0095	1.0100	0.008	1.249	.212
RI × FIB	-0.0230	0.9772	0.020	-1.173	.241
RI × Match	0.0237	1.0239	0.016	1.465	.143
RI × MA	-0.0317	0.9689	0.003	-9.128	< .001

RI = retention interval, AMC = advanced multiple choice, FIB = fill in the blanks, Match = matching, MA = multiple answer

choice (the selected reference for our model), increasing spacing by one day yields an odds of a correct response of .98 with respect to an individual answering correctly. In terms of odds ratios, any value less than 1 indicates that the probability of responding correctly was less than that of responding incorrectly. Given that past laboratory-based research has generally demonstrated better memory performance with increased spacing, one might have expected an odds value greater than 1 when the spacing interval was set to any value greater than 0. However, it is important to note the significance of the interaction between spacing and retention interval, as we discuss further below, which past work has shown to be an important contextual factor for detecting a spacing effect.

Our findings also show that when the spacing is set to zero, question format is set to multiple choice, and retention interval is increased by one day, the odds of an individual answering correctly is .99. Again, in terms of odds ratio, this finding suggests that the probability of responding correctly is less than that of responding incorrectly in the context of the specified parameters. To hone in on the impact of question format, retention interval was set to zero and spacing was held constant, whereas question format was set to AMC. This resulted in an increase in the odds ratio, whereas all the other levels of question format reduced the odds. Using these parameters for spacing and retention interval resulted in the following odds of answering correctly: 8.12 for AMC questions, 0.32 for FIB questions, 0.02 for matching questions, and 0.31 for MA questions.

As we mentioned above, we found a significant interaction between spacing and retention interval. Figure 2 shows the odds of answering correctly at O3 as spacing is increased and retention interval is fixed at various values. As we mentioned above, the spacing and retention interval variables were

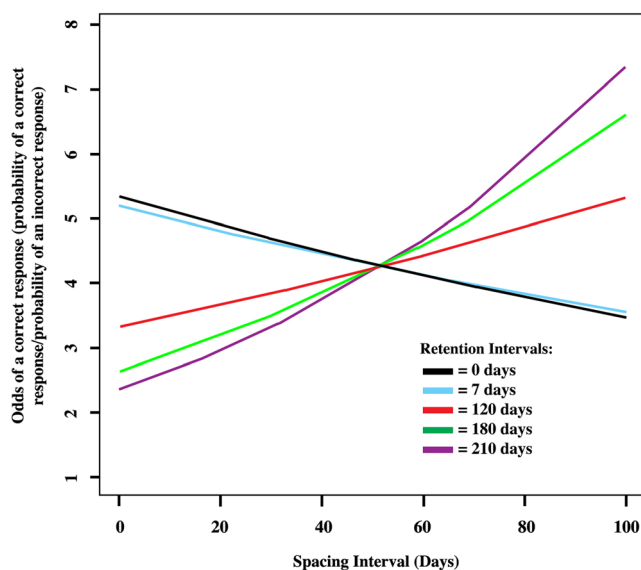


Fig. 2 Effects of spacing interval on the odds of answering correctly across retention intervals.

both continuous. Figure 2 shows that at shorter retention intervals (e.g., 0 and 7 days), the odds of answering correctly decreased as the spacing interval increased. In contrast, Fig. 2 shows that at longer retention intervals (e.g., 120 to 210 days), the odds of answering correctly increased as the spacing interval increased.

As was mentioned above, we also found a significant interaction between question format and retention interval. Figure 3 shows that when the spacing interval is held constant (e.g., at 30 days), the relation between retention interval and the odds of answering correctly differs across the various levels of question format. For example, when question format was set to multiple choice, the odds of answering correctly increased as retention interval increased. However, when the question format was set to AMC, the odds of answering correctly decreased as retention interval increased.

To unpack this interaction further, the odds of answering correctly at different retention intervals can be calculated for different question formats. For example, the following equation can be used to calculate the odds of answering correctly for the FIB question format, where RI refers to retention interval and R:FIB refers to the interaction between retention interval and the FIB question format. For demonstration purposes, the equation is set so that the retention interval is increased by one day and the spacing interval is held at 0.

$$\log \left(\frac{\text{odds}(Y = 1 \mid \text{Retention} = r + 1, \text{Question Format} = \text{FIB})}{\text{odds}(Y = 1 \mid \text{Retention} = r, \text{Question Format} = \text{FIB})} \right) = \beta_{RI} + \beta_{R:FIB} = -0.0189 - 0.0230 = -0.0419.$$

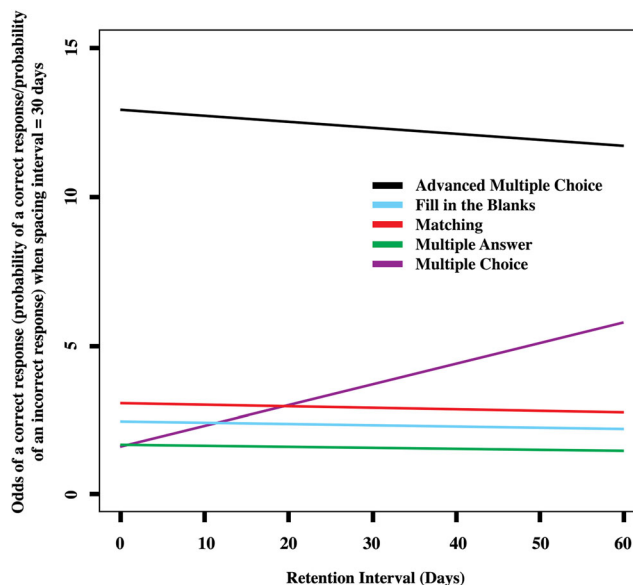


Fig. 3 Effects of retention interval, in days, on the odds of correctly answering across question formats when the spacing interval was set to 30 days.

The log-odds of answering correctly, derived using the formula above, can then be transformed into odds by using an inverse log transformation. Using this transformation, we found that the odds of answering correctly for the FIB question format and the specified parameters (retention interval increased by one day and spacing interval set to 0) was approximately 0.958. Table 5 presents the log-odds and odds of answering correctly for each level of question format, using the same parameters as those used above (retention interval increased by one day and spacing interval set to 0). Along these lines, the log-odds of answering correctly for each level of question format can also be calculated using different spacing-interval values. Again, using the example of the FIB question format, the log-odds of answering correctly can be calculated using the following formula:

$$\log\left(\frac{\text{odds}(Y = 1 \mid \text{Spacing} = s, \text{Retention} = r + 1, \text{Question Format} = \text{FIB})}{\text{odds}(Y = 1 \mid \text{Spacing} = s, \text{Retention} = r, \text{Question Format} = \text{FIB})}\right) = \beta_{\text{RI}} + \beta_{\text{R:FIB}} + \beta_{\text{S:R}} = -0.0419 + 0.0003 * s.$$

Using the inverse log transformation, the log-odds of answering correctly can be transformed into odds. Thus, the odds of answering correctly for FIB when the retention interval is increased by one day and spacing is not held at zero is approximately $0.9589e^{0.0003s}$. It then follows that if the spacing interval is set to 120 days, the odds of answering correctly for a FIB question is

$$0.9589e^{0.0003s} = 0.9589e^{0.0003*120} = 1.0270$$

Consequently, when the spacing interval is set to 120 days and the retention interval is set to one day, the odds of answering correctly for FIB is 1.0270. Table 6 presents both the log-odds and odds of answering correctly for each level of question format when the spacing interval is not held at 0 (e.g., spacing interval = 120 days, as in the example above).

Table 5 Log-odds and odds of correctly responding when retention interval was increased by one day and spacing was held at 0

Question Format	Log-Odds	Odds
AMC	- 0.0094	0.9907
FIB	- 0.042	0.9589
Match	0.0047	1.0047
MA	- 0.0507	0.9506

AMC = advanced multiple choice, FIB = fill in the blanks, Match = matching, MA = multiple answer

Discussion

In the present study, we modeled a large, real-world dataset to investigate the impact of spaced retrieval on knowledge retention beyond the lab. The data were collected via a LMS that delivered workplace training materials to employees in a naturally occurring, heterogeneous, spaced retrieval manner. Upon each occurrence of a given piece of information, employees were tested on their knowledge of the information. We organized these data to align with the simplest research design that can be used to investigate the impact of spaced retrieval on learning. Our analyses were conducted on individuals’ performance on the first three occurrences of a question. The first occurrence corresponded to an initial learning session (material was delivered to the employee for the first time using the LMS), and the second occurrence corresponded to a relearning session (the same material was revisited by the employee). Thus, the interval between these occurrences was operationally defined as the spacing interval. The third occurrence of a question corresponded to the final test. Accordingly, the interval between the second and third occurrences was operationally defined as the retention interval. In addition to significant main effects for all the variables included in the model (spacing interval, retention interval, questions format), the results of the present study also revealed a significant interaction between spacing interval and retention interval, as well as between retention interval and question format. These findings are in line with the results of laboratory studies, demonstrating the relevance and transferability of laboratory-based research to real-world contexts.

The interaction we found between spacing interval and retention interval was of prime interest and consistent with past laboratory studies demonstrating that the optimal amount of spacing between an initial learning event and a relearning session varies depending on the length of the retention interval (Cepeda et al., 2009; Cepeda et al., 2006; Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008). More specifically, the probability of retaining information in memory for a longer period of time (e.g., a month or longer) is higher if the spacing interval is also long (e.g., 11 days or longer) (Cepeda et al., 2009;

Table 6 Log-odds and odds of correctly responding when the retention interval was increased by one day and spacing was not held at 0

Question Format	Log-Odds	Odds
AMC	- 0.0094 + 0.0003 * s	$0.9907e^{(0.0003 * s)}$
FIB	- 0.042 + 0.0003 * s	$0.9589e^{(0.0003 * s)}$
Match	0.0047 + 0.0003 * s	$1.0047e^{(0.0003 * s)}$
MA	- 0.0507 + 0.0003 * s	$0.9506e^{(0.0003 * s)}$

s = spacing (in days), AMC = advanced multiple choice, FIB = fill in the blanks, Match = matching, MA = multiple answer

Glenberg & Lehmann, 1980; Küpper-Tetzel & Erdfelder, 2012; Küpper-Tetzel, Erdfelder, & Dickhäuser, 2013). In contrast, shorter spacing intervals (e.g., one day) have been found to be more beneficial for shorter retention intervals (e.g., one week). Thus, the optimal amount of spacing between an initial learning event and a relearning event increases as the retention interval increases. In fact, an influential review aimed at informing best practices in teaching and learning in educational settings concluded that the interval between two study occasions should be approximately 10% to 20% of the retention interval (Pashler et al., 2007). This coincides with the finding that although spacing out repeated study sessions generally benefits knowledge retention, excessive spacing results in decreased retention.

In workplace environments, corporations typically aim to onboard new employees as quickly as possible, and in general, aim for their employees to rapidly acquire the information they need to be successful in their positions, thereby benefiting the company. In this context, the interval between a given training event and the actual use of that knowledge on the job may be quite short (e.g., one week). Thus, on the basis of the results of the present study and the abovementioned laboratory research, a shorter spacing interval (e.g., one day) between repeated training events would be more beneficial than a longer spacing interval (e.g., one week). Additionally, however, corporations benefit and aim to support their employees in retaining information over much longer stretches of time even though employees may not make use of this information on a daily basis (e.g., emergency procedures). In these cases, the results of the present study and the abovementioned laboratory research suggest that longer spacing intervals between repeated training events would be most beneficial. Generally, optimizing spacing intervals between repeated training events requires consideration of how long the information is intended to be retained.

Interestingly, retention interval also seems to impact the optimal spacing schedule or the optimal amount of spacing between repetitions when there is more than one relearning event following the initial learning event and preceding the final test. An equal-interval schedule, for example, involves equally spaced out study episodes, whereas an expanding schedule consists of learning episodes that are spaced apart by incrementally increasing intervals. Mixed findings have been reported in the literature regarding whether an equal-interval schedule or expanding schedule is more beneficial for retention. For example, some studies do not show a difference between the two types of spacing schedules (e.g., Balota et al., 2006; Carpenter & DeLosh, 2005; Karpicke & Bauernschmidt, 2011; Kim et al., 2018), whereas other studies demonstrate a larger benefit from expanding over equal-interval spacing schedules in specific contexts (Gerber & Koenig, 2012; Karpicke & Roediger, 2007; Nakata, 2015). Interestingly, Karpicke and Roediger (2007) found an

interaction between retention interval and the benefit of different spacing schedules on memory performance in young adults—following a short (10-min) retention interval, the expanding schedule resulted in a larger benefit than did an equal-interval schedule. However, following a longer (two-day) retention interval, the equal-interval schedule resulted in a larger advantage than did the expanding schedule. Logan and Balota (2008) and Tsai (1927) have reported consistent findings. However, other studies that used long, multi-day study sessions and retention intervals have produced mixed findings. Moreover, Storm, Bjork, and Bjork (2012) and Balota et al. (2006) both found that an expanding retrieval practice schedule was most effective when the to-be-remembered materials were subject to rapid forgetting. Future research should investigate this further and would benefit from investigations using real-world data to assess ecological validity.

The results of our model also demonstrated an interaction between question format and retention interval, which may reflect differences in difficulty and retrieval cues available across the different types of question format ((Tulving & Osler, 1968; Tulving & Pearlstone, 1966)). For example, MC and MA questions included the correct answer within the questions, which serves as a strong retrieval cue, whereas FIB questions did not and required employees to retrieve the corresponding information from memory. The corresponding main effects of retention interval and question format on the odds of employees answering correctly on final tests was not surprising. Our model revealed a negative coefficient for retention interval, which is in line with the vast literature on forgetting curves that demonstrates that the probability of retaining information decreases exponentially from the time of the original learning event (Ebbinghaus, 1885/1964). In terms of question format, differences in the level of retrieval difficulty and retrieval cues may account for the observed differences that the various types of question format had on employees' odds of answering correctly on the final test.

As we described above, our analyses focused on items that were first answered incorrectly on the first retrieval attempt, because we could not be certain if those items that were answered correctly were already known prior to the learning event in question. For this reason, we considered these data (items that were first answered correctly) to be potentially confounded. Although it is beyond the focus of the present study, an intriguing question is whether the odds of providing a correct response at final recall varies as a function of a participants' retrieval success on the first and second occurrence of an item, across the four potential outcomes: correct first retrieval–correct second retrieval; correct first retrieval–incorrect second retrieval; incorrect

first retrieval–incorrect second retrieval; incorrect first retrieval–correct second retrieval.

In a relevant study by Kapler et al. (2015), participants engaged in spaced retrieval, with two instances of retrieval practice for each study item before the final test. The final test data were analyzed in two ways: (1) considering only items to which participants provided a correct response in the first and second retrieval events; and (2) considering all items regardless of whether participants provided a correct response in the first retrieval event. Importantly, participants could only progress in the study after providing a correct response on the second retrieval event, such that all instances of a second retrieval event for an item were coded as a correct response. The analyses conducted on this latter dataset showed that retrieval accuracy on the first retrieval event was approximately 55%. Both sets of data demonstrated a spacing effect, with final test performance being moderately higher and the effect size moderately larger when all data were included (regardless of participants' accuracy on the first retrieval event; factual-level condition: $p = .02$, $\eta^2 = .026$; higher-level condition: $p = .003$, $\eta^2 = .018$) than when the data were filtered to only include items that were correct on the first retrieval event (factual-level $p = .01$, $\eta^2 = .016$; higher-level $p = .06$, $\eta^2 = .009$). An important difference between the study by Kapler and colleagues and the present study is that participants in the former study did not know the material beforehand, whereas participants in the present study may very well have already been familiar with the content. Future research should investigate further the impact of consecutive retrieval successes on a final memory test.

Although working with a large dataset collected in a real-world context provides an opportunity to test the ecological validity of laboratory findings, some limitations are also intrinsic to the nature of these data. For example, in contrast to the controlled laboratory environment, the data analyzed in the present study were collected in a variety of uncontrolled environments. As mentioned above, the LMS allows the training, and thus data collection, to occur using a variety of electronic devices, including laptops and smartphones. Consequently, the training could have taken place anywhere and at any time. Moreover, we cannot be sure that employees answered the question on their own as opposed to working with a colleague, whether they understood the training task and how seriously they took it. Although one cannot be sure of the internal motivations and attitudes of participants in laboratory studies, participants in these studies are observed and in the company of an experimenter while performing the experimental task. In contrast, however, employees who were engaged in workplace training were not necessarily in the company of anyone, let alone an

authority figure. Thus, employees may have been less motivated when it came to focusing on and completing the training task to the best of their ability. Moreover, there was also a lack of control and/or measure of how much employees used the learned information between training sessions.

It is also important to note that learners received corrective feedback after answering questions incorrectly. Thus, we cannot be certain whether the act of spaced retrieval or the spacing of the feedback, or an interaction between the two, led to the present results. Additionally, each of the client companies trained their employees on different information, using different questions and question formats. Although we accounted for company and the specific components of content on which employees were trained (e.g., question) by setting them as random factors in our model, in addition to including question format as a variable, company and question still served as sources of variance. Another limitation is the lack of demographic information we have about the employees whose data we analyzed, particularly age. Without such information, we were unable to account for the impact of age on employees' learning trajectories, particularly with respect to spacing and retention intervals. Past work has shown that age is an important factor that affects learning and memory (e.g., Burke & Barnes, 2006; Old & Naveh-Benjamin, 2008), and had this information been made available, we could have added greater precision to the model and possibly gained a greater understanding of age interactions. Despite the abovementioned sources of variability, our model revealed significant effects of the variables tested on employees learning behavior, including an interaction between spacing interval and retention interval as supported by the results of laboratory research. The present investigation was guided by the extant literature on the spacing effect, which largely consists of small- to moderate-size studies (for reviews see Cepeda et al., 2006; Gerbier & Toppino, 2015). Thus, in contrast to the more exploratory nature of most studies using big data, the present investigation was hypothesis-driven based on past empirical work. Future research should continue to employ large real-world data to investigate the ecological validity of memory-related phenomena that are traditionally studied in the laboratory.

Author note We are grateful to Axonify Inc. for their continued research collaborations. This research was supported by grants from Natural Sciences and Engineering Research Council of Canada to R.S.R. and A.S.N.K., as well as a fellowship from Natural Sciences and Engineering Research Council of Canada to A.S.N.K. The authors declare no conflicts of interest.

Appendix

Table 7 Sample questions for each question format used by each company

Company	Type of Company; Nature of Training	Multiple Choice	Multiple Answer	Advanced Multiple Choice	Match	Fill in the Blanks
1	High-end department store retailer; safety and store operations	Lois is tidying up her department before closing time and she overhears co-workers Tammy and Denise talking about the customer accident that happened the other day in front of other customers. What should Lois do? (x) She should inform Tammy and Denise that the accident should not be talked about in front of other customers and explain why. () She should just ignore Tammy and Denise's conversation. () She should report Tammy and Denise to their manager.	Sharing product information and care instructions with customers before they leave is a great way to . . . (SELECT ALL THAT APPLY) (x) show you care about their business (x) show that you are an expert () ensure they don't make a return	Sample question not available.	N/A	Ensure the _____ and limbs are kept within _____ to prevent trip or struck-against hazards. (1) limbs (2) the platform edges () arms () legs () 30 inches away from the platform edges
2	Pharmaceutical sales; product knowledge	In the context of Medicare, readmissions have been generally defined as a patient being hospitalized within _____ days of an initial hospital stay. () 3 () 14 (x) 30 () 60	N/A	N/A	The order of tissues incised, from superficial to deep, when a midline abdominal incision is made into the body is: Skin=1 Subcutaneous Fat = 2 Fascia = 3 Peritoneum = 4	N/A
3	Higher education institution; training procedures	Veronica is currently in an active status; however, her grades are slipping. She would like to know what would happen to her at the conclusion of her term if she does not meet the minimum requirements outlined in her program. Select the best answer. (x) At the end of the term you will be evaluated. If you do not meet the minimum requirements as outlined in your program, your status will go from Active to Probation. () At the end of the term you will be evaluated. You must meet the	Which of the following are examples of question pitfalls? (Select all that apply.) (x) Early yes/no questions (x) Serial questions (x) Multiple-choice questions () Probing questions	N/A	Match the following Salesforce terms to their correct description. Activity = A planned task or event, optionally related to another type of record such as an account, contact, lead, opportunity, or case Task = Workflow and approval actions triggered by workflow rules or approval processes Inquiry = a lead in the following statuses: New Assign, Contact Attempt, Contact, Interview Complete and Remarketing	N/A

Table 7 (continued)

Company	Type of Company; Nature of Training	Multiple Choice	Multiple Answer	Advanced Multiple Choice	Match	Fill in the Blanks
		<p>minimum requirement of either the CGPA or Pace of Completion to stay in Active status.</p> <p>() At the end of the term you will be evaluated and if you do not meet the minimum requirements as outlined in your program, your status will go from Active to Final Probation.</p>				
4	Pharmaceutical sales; product knowledge	<p>In a sales call, you can do all of the following except:</p> <p>() Compare dosing regimens</p> <p>() Compare mechanisms of action</p> <p>(x) Compare the percentage of patients in remission at a particular time point</p> <p>() Compare indications</p>	<p>In managed care, each patient may be under different benefits designs such as: (Check all that apply)</p> <p>(x) Different formularies</p> <p>(x) Different "cost sharing" - co pays or co-insurance</p> <p>(x) Different SPP requirements</p> <p>(x) Different employer group carve outs</p>	N/A	<p>Match the spot on the product adoption/call continuum with the corresponding description.</p> <p>Awareness = Your customer indicates knowledge of the medication but has not prescribed the medication or had experience with patients using the drug.</p> <p>Evaluation = Your customer has had experience with patients on the medication and may have prescribed the medication for a very small number of patients.</p> <p>Trial = Your customer has prescribed the medication for some patients or in a specific patient population.</p> <p>Adoption = Your customer prescribes the medication for a large number of patients in a population that is crucial to his practice</p> <p>Advocacy = Your customer not only regularly prescribes the medication for the appropriate patient types but is also willing to influence and encourage colleagues to prescribe the medication.</p>	N/A
5	U.S. Grocery Retailer; loss prevention, safety, and store operations	<p>Who should I contact if a customer requests nutritional information on a product?</p> <p>() Company website</p> <p>() Grocery Clerk</p> <p>() Receiver</p>	<p>How does internal theft impact honest associates?</p> <p>(x) It reduces the opportunities for more job promotions.</p> <p>(x) It reduces the opportunities for better wages.</p>	N/A	N/A	N/A

Table 7 (continued)

Company Type of Company; Nature of Training	Multiple Choice	Multiple Answer	Advanced Multiple Choice	Match	Fill in the Blanks
	(x) Store Team Leader	(x) It reduces the job security of all associates. () It doesn't impact honest associates in any way.			

A sample advanced multiple choice question was not available for Company 1. N/A = not applicable for the designated company.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger, H. L., III. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging, 21*, 19–31. doi:<https://doi.org/10.1037/0882-7974.21.1.19>
- Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 123–144). Hillsdale, NJ: Erlbaum.
- Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues. Vol. 1: Memory in everyday life* (pp. 396–401). London, UK: Wiley
- Bower, G. H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior, 9*, 529–533. doi:[https://doi.org/10.1016/S0022-5371\(70\)80096-2](https://doi.org/10.1016/S0022-5371(70)80096-2)
- Brod, G., Werkle-Bergner, M., & Shing, Y. L. (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience, 7*, 139.
- Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience, 7*, 30–40. doi:<https://doi.org/10.1038/nrn1809>
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology, 19*, 619–636. doi:<https://doi.org/10.1002/acp.1101>
- Carson, N., Murphy, K. J., Moscovitch, M., & Rosenbaum, R. S. (2016). Older adults show a self-reference effect for narrative information. *Memory, 24*, 1157–1172. doi:<https://doi.org/10.1080/09658211.2015.1080277>
- Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology, 56*, 236–246. doi:<https://doi.org/10.1027/1618-3169.56.4.236>
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380. doi:<https://doi.org/10.1037/0033-2909.132.3.354>
- Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*, 1095–1102. doi:<https://doi.org/10.1111/j.1467-9280.2008.02209.x>
- Cermak, L. S., Verfaellie, M., Lanzoni, S., Mather, M., & Chase, K. A. (1996). Effect of spaced repetitions on amnesia patients' recall and recognition performance. *Neuropsychology, 10*(2), 219–227. <https://doi.org/10.1037/0894-4105.10.2.219>
- Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. New York, NY: Dover. (Original work published 1885)
- Gerbier, E., & Koenig, O. (2012). Influence of multiple-day temporal distribution of repetitions on memory: A comparison of uniform, expanding, and contracting schedules. *Quarterly Journal of Experimental Psychology, 65*, 514–525. doi:<https://doi.org/10.1080/17470218.2011.600806>
- Gerbier, E., & Toppino, T. C. (2015). The effect of distributed practice: Neuroscience, cognition and education. *Trends in Neuroscience and Education, 4*, 49–59. doi:<https://doi.org/10.1016/j.tine.2015.01.001>
- Glenberg, A. M., & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory & Cognition, 8*, 528–538. doi:<https://doi.org/10.3758/BF03213772>

- Green, J., Weston, T., Wiseheart, M., & Rosenbaum, R. (2014). Long-term spacing effect benefits in developmental amnesia: Case experiments in rehabilitation. *Neuropsychology*, *28*, 685–694. doi:<https://doi.org/10.1037/neu0000070>
- Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 371. doi:<https://doi.org/10.1037/0278-7393.15.3.371>
- Guan, L., Tian, X., Gombar, S., Zemek, A. J., Krishnan, G., Scott, R., . . . Pham, T. D. (2017). Big data modeling to predict platelet usage and minimize wastage in a tertiary care system. *Proceedings of the National Academy of Sciences*, *114*, 11368–11373. doi:<https://doi.org/10.1073/pnas.1714097114>
- Harlow, L. L., & Oswald, F. L. (2016). Big data in psychology: Introduction to the special issue. *Psychological Methods*, *21*, 447–457.
- Kapler, I. V., Weston, T., & Wiseheart, M. (2015). Spacing in a simulated undergraduate classroom: Long-term benefits for factual and higher-level learning. *Learning and Instruction*, *36*, 38–45. doi:<https://doi.org/10.1016/j.learninstruc.2014.11.001>
- Karpicke, J. D., & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1250. doi:<https://doi.org/10.1037/a0023436>
- Karpicke, J. D., & Roediger, H. L., III. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 704. doi:<https://doi.org/10.1037/0278-7393.33.4.704>
- Kim, A. S. N., Saberi, F. M., Wiseheart, M., & Rosenbaum, R. S. (2018). Ameliorating episodic memory deficits in a young adult with developmental (congenital) amnesia. *Journal of the International Neuropsychological Society*, *24*, 1003–1012. doi:<https://doi.org/10.1017/S1355617718000589>
- Küpper-Tetzel, C. E., & Erdfelder, E. (2012). Encoding, maintenance, and retrieval processes in the lag effect: A multinomial processing tree analysis. *Memory*, *20*, 37–47. doi:<https://doi.org/10.1080/09658211.2011.631550>
- Küpper-Tetzel, C. E., Erdfelder, E., & Dickhäuser, O. (2013). The lag effect in secondary school classrooms: Enhancing students' memory for vocabulary. *Instructional Science*. Advance online publication. doi:<https://doi.org/10.1007/s11251-013-9285-2>
- Logan, J. M., & Balota, D. A. (2008). Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, and Cognition*, *15*, 257–280. doi:<https://doi.org/10.1080/13825580701322171>
- Nakata, T. (2015). Effects of expanding and equal spacing on second language vocabulary learning: Does gradually increasing spacing increase vocabulary learning? *Studies in Second Language Acquisition*, *37*, 677–711. doi:<https://doi.org/10.1017/S0272263114000825>
- Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*, 104–118. doi:<https://doi.org/10.1037/0882-7974.23.1.104>
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*, 241–263. doi:<https://doi.org/10.1037/h0027272>
- Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). Organizing instruction and study to improve student learning (NCER 2007-2004). Washington, D.C.: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
- Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210. doi:<https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- Ryan, J. D., Moses, S. N., Barense, M., & Rosenbaum, R. S. (2013). Intact learning of new relations in amnesia as achieved through unitization. *Journal of Neuroscience*, *33*, 9601–9613. doi:<https://doi.org/10.1523/JNEUROSCI.0169-13.2013>
- Shiffrin, R. M. (2016). Drawing causal inference from Big Data. *Proceedings of the National Academy of Sciences*, *113*, 7308–9. doi:<https://doi.org/10.1073/pnas.1608845113>
- Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, *25*, 763–767. doi:<https://doi.org/10.1002/acp.1747>
- Storm, B. C., Bjork, R. A., & Storm, J. C. (2010). Optimizing retrieval as a learning event: When and why expanding retrieval practice enhances long-term retention. *Memory & Cognition*, *38*, 244–253. doi:<https://doi.org/10.3758/MC.38.2.244>
- Storm, B. C., Bjork, E. L., & Bjork, R. A. (2012). On the durability of retrieval-induced forgetting. *Journal of Cognitive Psychology*, *24*, 617–629. doi:<https://doi.org/10.1080/20445911.2012.674030>
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*, 371–394. doi:<https://doi.org/10.1037/0033-2909.121.3.371>
- Thorstad, R., & Wolff, P. (2018). A big data analysis of the relationship between future thinking and decision-making. *Proceedings of the National Academy of Sciences*, *115*, E1740–E1748. doi:<https://doi.org/10.1073/pnas.1706589115>
- Tsai, L. S. (1927). The relation of retention to the distribution of relearning. *Journal of Experimental Psychology*, *10*, 30–39. doi:<https://doi.org/10.1037/h0071614>
- Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, *77*, 593–601. doi:<https://doi.org/10.1037/h0026069>
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*, 381–391. doi:[https://doi.org/10.1016/S0022-5371\(66\)80048-8](https://doi.org/10.1016/S0022-5371(66)80048-8)
- Wolfe, P. J. (2013). Making sense of big data. *Proceedings of the National Academy of Sciences*, *110*, 18031–18032. doi:<https://doi.org/10.1073/pnas.1317797110>