BRIEF REPORT

Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning

Lynne M. Reder[1,2] · Xiaonan L. Liu[1,2] · Alexander Keinath[1,2] · Vencislav Popov[1,3]

Abstract Despite vast efforts to better understand human learning, some principles have been overlooked; specifically, that less familiar stimuli are more difficult to combine to create new knowledge and that this is because less familiar stimuli consume more working memory resources. Participants previously unfamiliar with Chinese characters were trained to discriminate visually similar characters during a visual search task over the course of a month, during which half of the characters appeared much more frequently. Ability to form associations involving these characters was tested via cued recall for novel associations consisting of two Chinese characters and an English word. Each week performance improved on the cued-recall task. Crucially, however, even though all Chinese *character pairs* were novel each week, those pairs consisting of more familiar characters were more easily learned. Performance on a working-memory task was better for more familiar stimuli, consistent with the claim that familiar stimuli consume fewer working memory resources. These findings have implications for optimal instruction, including second language learning.

✉ Lynne M. Reder
reder@cmu.edu

[1] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[2] Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[3] Department of Cognitive Science and Psychology, New Bulgarian University, Sofia, Bulgaria

Keywords Episodic memory · Human memory and learning · Working memory · Encoding effects

In the quest to ever improve humans' ability to learn, researchers have explored many factors that impact learning (Ebbinghaus, 1964; Hintzman, 1974; Nelson & Shiffrin, 2013; Reder et al., 2000; Tulving & Kroll, 1995; Tulving, Markowitsch, Craik, Habib, & Houle, 1996). For example, rehearsal of information (e.g., Ericsson, Krampe, & Teschromer, 1993), distributed practice (Glenberg, 1979; Greene, 1989; Ross & Landauer, 1978), encoding variability (Bird, Nicholson, & Ringer, 1978; Postman & Knecht, 1983), and attempting to retrieve the studied information all improve memory performance (e.g., Karpicke & Roediger, 2008). Likewise, we know that it is easier to encode information that can be "chunked" into meaningful units (Miller, 1956; Simon, 1974). A simple demonstration is that people have more difficulty recalling a string of letters such as BIC IAJ FKI RSU SAF than another string of letters, such as FBI CIA JFK IRS USA, even though the order of letters in the two strings is almost identical. Likewise, a chess master can remember the position of pieces shown on a chessboard displayed only briefly while novice chess players can recall very few pieces; however, when the chess pieces are randomly placed on the chessboard instead of in a configuration that would occur in chess, chess masters are no better than novices at recalling the positions of the pieces (Chase & Simon, 1973).

While it is well known that it is easier to encode and remember knowledge when it can be "chunked" into high-level units, it is not known whether the chunk's *familiarity* also affects its encoding. Diana and Reder (2006) found that low-frequency words had an encoding disadvantage, and Reder, Paynter, Diana, Ngiam, and Dickison (2007) proposed that the encoding of stimuli consumes working memory resources in proportion to

their familiarity. Given that working memory is implicated in the formation of new memories (Blumenfeld & Ranganath, 2006), we propose that it should be more difficult to encode and create new knowledge structures when their elements are less familiar.

We find those assumptions plausible because they can explain a number of puzzling findings. For example, Reder et al. (2007) demonstrated that a simple computational model, in which high-frequency words consume less of the available working-memory resources, reproduced the result that recognition memory for low-frequency words in a list benefit when the other words in the list are high-frequency, and high-frequency words on a list are hurt when the other words on the list are low-frequency.

The idea that the familiarity of stimuli affects memory for them is not new (Glanzer & Adams, 1985). Short-term memory span is greater for words than for nonwords (Multhaup, Balota, & Cowan, 1996) and differential exposure to non-words produces the same mirror effects found in recognition for words (Reder, Angstadt, Cary, Erickson, & Ayers, 2002), as well as previously unfamiliar Chinese characters (Nelson & Shiffrin, 2013). Cued recall is better for high-frequency cues and targets (Criss, Aue, & Smith, 2011), although the high-frequency advantage is not always found at encoding (Madan, Glaholt, & Caplan, 2010). Immediate serial recall and nonword repetition are better for nonword syllables that occur frequently in polysyllabic words, compared to those that do not (Nimmo & Roodenrys, 2002), demonstrating that the familiarity of the constituents, not just that of the total structure, is important.

There is also evidence consistent with the idea that familiarity aids binding. It is easier to recall episodic encoding details for familiar proverbs compared to completely novel proverbs (Poppenk, Kohler, & Moscovitch, 2010), and the advantage of context reinstatement for face recognition occurs only for well-known faces (Reder et al., 2012). Finally, the familiarity of the category to which an item belongs predicts source memory accuracy for that item but the effect disappears when attention is divided during item encoding (DeWitt, Knight, Hicks, & Ball, 2012), supporting the view that more familiar stimuli are easier to encode by providing more attentional resources to enable their binding to the episodic context.

While a relatively simple model can explain many of the results reviewed above, most of these experiments involved *quasiexperimental* designs. Therefore, the cause of these memory effects might be due to another variable associated with familiarity but not familiarity, per se. Furthermore, these studies addressed whether memory for a single stimulus differs as a function of its familiarity and not whether the familiarity of the constituent elements would affect memory for the complete structure.

The goal of this study is to examine whether forming new knowledge structures is affected by the *familiarity* of the elements that must be combined to form the new information. The idea of a relationship between the chunks' strength, working memory, and formation of higher level knowledge

structures has not been experimentally tested before. Here, we experimentally manipulate the familiarity of stimuli that were previously unfamiliar to our participants. We argue that (a) the ease of forming complex knowledge structures is affected by the familiarity of the chunks to be combined and (b) that working memory resource depletion is inversely related to the familiarity of the information to be processed. A strong case for this proposal requires evidence for both aspects of our thesis. We expect people to more successfully learn combinations of familiar elements than of unfamiliar elements. We will also test whether more familiar stimuli consume less working memory resources.

Method

Participants

Twenty U.S. college students with no prior experience learning Chinese participated in this experiment. Our sample size was determined using pilot data with an a priori power analysis by GPower (Faul, Erdfelder, Buchner, & Lang, 2009) with power (1-β) set at 0.80 and α = 0.05. The recommended effect sizes used for this assessment are as follows: small ($\eta_p^2 = .01$), medium ($\eta_p^2 = .06$), and large ($\eta_p^2 = .14$; Cohen, 1992). Because each participant had 12 training sessions and the ratio between high-frequency and low-frequency was 20:1, we expected a large effect size. This analysis indicated that we needed at least 15 participants to detect a large effect.

Materials

We familiarized participants with 64 Chinese characters over 4 weeks, with some characters given 20 times the exposure given to other characters. The characters were grouped into 16 sets of four characters (64 in all) based on their visual similarity. By using visually similar distractors, participants had to encode the entire character rather than a subset of features. For each participant, 8 of the 16 sets were randomly selected to be in the high-frequency condition. The differential familiarity of the characters was achieved exclusively by differential exposure during the visual search task.

Procedure

Participants performed four different tasks over the course of 4 to 8 weeks.

Visual search task Participants performed a visual search task for 3 hour-long sessions per week for a total of 4 weeks, and each session consisted of several hundred trials. Figure 1 (top panel) illustrates a single visual search training
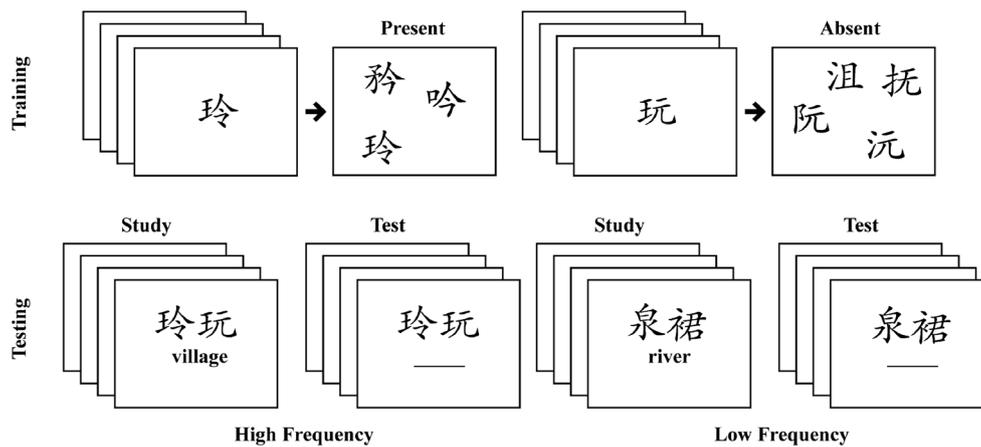
Fig. 1 Top panel shows example trials of the visual search task. The bottom panel shows study and test trials. Study Lists and Tests were segregated by the frequency of the constituent Chinese characters

trial. Each trial showed a target character followed by a display of three to five characters, and participants had to indicate whether the target character was present in the display (it was, 50 % of the time). The visual search display contained exactly three out of the four characters from a target's similarity set along with zero to two additional characters from different character sets of the same frequency class (i.e., high or low frequency).

Weekly paired-associate/cued-recall test Every week we tested our hypothesis that ease of learning depends on the familiarity of the constituents to be associated by presenting for study new combinations of two characters from the same frequency condition, paired with a new English word, and then testing the participant's ability to recall the English word when cued with the associated Chinese character pair. See Fig. 1 (bottom panel) for an illustration. By differentially familiarizing previously unfamiliar Chinese characters over a period of weeks, we were able to track improvement in learning *novel combinations* of these stimuli by week as well as by frequency condition. There were 10 to 11 pairs on each list and three lists of each frequency. Each character on a list was used in two different pair combinations so that participants were forced to attend to both characters of the pair when learning to associate the character pairs with an English word. A given character pair was shown for 3 seconds with its word before the next association appeared. A cued-recall test was given after all pairs had been shown with their respective words. Character pairs were tested one at a time in a different random order from the study list, with the participant attempting to recall the word just studied with it.

This paired-associate/cued-recall test was administered before the first visual search session for each week, starting with the second week. After 4 weeks of visual search training and three cued-recall tests as described above, participants were dismissed for 2 to 4 weeks before returning for a final test.

Final paired-associate/cued-recall test After this delay period, participants returned to learn a new set of arbitrary character pairs assigned to different English words. Participants learned a list of 10 to 11 character pairs, all either high- or all low-frequency characters, and then were immediately tested to recall the English word. Each character pair was studied for 2 seconds. After testing memory for all six lists (three high-frequency lists, three low-frequency lists), the procedure repeated with the same lists, for a total of five rounds of cued recall. This task focused on whether participants would still find it easier, after a longer delay from training, to learn novel pairs consisting of high-frequency stimuli.

N-back task At the end of the final session participants performed a working-memory task using the same Chinese characters that allowed us to test our hypothesis that more familiar stimuli consume less working memory resources. We adopted the *N*-back task, which does not require building new knowledge structures – it uses working memory resources instead (Braver et al., 1997; Kirchner, 1958). Participants were shown a series of individual Chinese characters one at a time at a fixed presentation rate of 2.5 seconds and had to indicate whether the current stimulus matched the stimulus that appeared *N* presentations before, where *N* varied from 1 to 3 in different blocks of trials. Half of the blocks for each level of *N*-back used low-frequency Chinese characters and the other half used the high-frequency characters. The 1-back task is not demanding because one only compares the current stimulus with the immediately preceding item; the 3-back task, however, involves holding three stimuli in working memory so that the identities of the stimuli that are "3-back," "2-back," and "1-back" can be updated with each presentation, as well as simultaneously determining the correct response and pushing the button. There were separate blocks of 17 trials for each *N*-back level for both familiarity levels, with four blocks of trials for each of these six conditions. The order of blocks was randomly determined for each participant.

Results

We analyzed the accuracy data via logistic mixed-effects regressions and reaction times via linear mixed-effects regressions, both with participants and items as random intercept effects (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). *P* values were obtained by likelihood ratio tests of the model with the effect in question compared with an identical model that lacked only the effect in question; all effects were added to the models in the order they are reported.

We excluded from the analyses cases with RTs more than 3 median absolute deviations above or below the median RT, calculated separately for each participant, session and condition (2.26 %). RTs were log-transformed because residual plots revealed a lack of homoscedasticity. For the RT analyses, we considered only trials with correct responses (8.19 % error).

Analyses of the visual search task

Participants performed better on the visual search task for trials involving more familiar characters, becoming faster, ΔAIC = -5,075, LLR χ^2 (1) = 5,076.716, p < .001, and more accurate, ΔAIC = -1,080, LLR χ^2 (1) = 1,082.134, p < .001, over the 12 sessions (see Fig. 2). In addition, participants identified high-frequency characters in the displays more quickly, ΔAIC = -63, LLR χ^2 (1) = 64.812, p < .001, and with greater accuracy, ΔAIC = -179, LLR χ^2 (1) = 180.835, p < .001. The accuracy and speed improved over weeks for both frequency conditions as all the characters became more familiar. Some improvement may be due to practice at the task; however, performance was always better for the high-frequency characters, so practice at the task is not the entire story. Moreover, the improvement in speed with time was greater for low-frequency stimuli, ΔAIC = -8, LLR χ^2 (1) = 10.392, p = .001.

Analyses of the paired-associate/cued-recall test

Not only did participants need to learn different character pairs each week, new, different English words were also assigned to each new character pair each week. Accuracy at recalling the words to a new character pair was the measure of performance. The left panel of Fig. 3 shows mean performance for each week's test as a function of character frequency. Across weeks of tests, performance was significantly better for high-frequency characters, ΔAIC = -10, LLR χ^2 (1) = 12.019, p < .001, and accuracy in both frequency conditions increased with additional familiarization across weeks, ΔAIC = -42, LLR χ^2 (1) = 43.494, p < .001. The performance difference between frequency conditions increased from Week 1 (0.01) to Weeks 2 and 3 (0.07 and 0.08); ΔAIC = -1, LLR χ^2 (1) = 2.675, p = 0.1. These results are consistent with our hypothesis that it is easier to learn new information when the constituents are more familiar.

The right panel of Fig. 3 shows the effect of frequency on the delayed tests across each of the five rounds of learning. This familiarity advantage for learning new associations was still significant after a two-to-four-week delay, ΔAIC = -36, LLR χ^2 (1) = 39.605, p < .001.

Analyses of the *N*-back task

Figure 4 shows the performance on the *N*-back task. Performance declined, ΔAIC = -495, LLR χ^2 (2) = 499.122, p < .001, as working memory load increased (from 1-back to 3-back), as expected from earlier results using this paradigm (e.g., Braver et al., 1997; Kirchner, 1958; see the Method section). Accuracy was better for high-frequency characters compared to low-frequency characters, ΔAIC = -27, LLR χ^2 (1) = 19.252, p < .001, but there was no significant interaction between frequency and working memory load, ΔAIC = 17, LLR χ^2 (2) = 1.328, p = .515 (high- to low-frequency odds ratio per condition – 1.59, 1.40, and 1.27, respectively, for the 1-back, 2-back, and 3-back conditions).
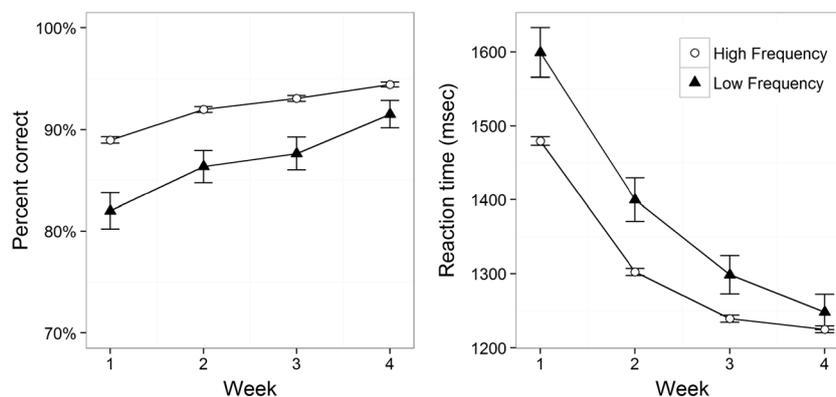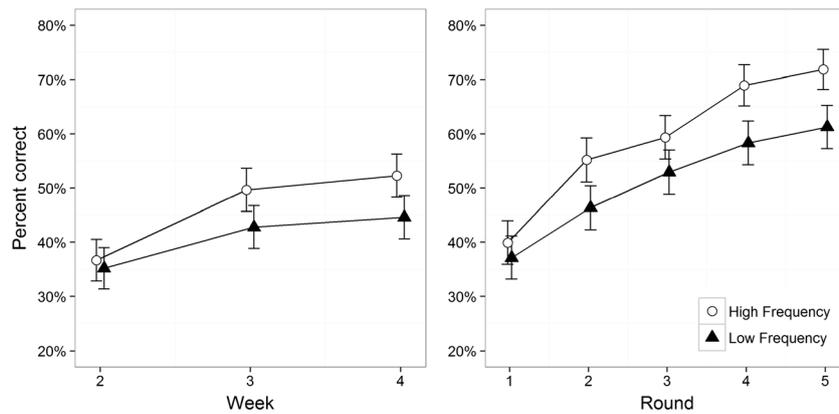

Fig. 2 Mean performance on visual search task trials for low and high-frequency over weeks of training. Left panel plots accuracy, right panel plots reaction times. Error bars indicate +/- 2 SEs

Fig. 3 Left panel plots accuracy for the three weekly associative learning tests. Right panel plots accuracy on the delayed test (between 2 to 4 weeks post training) of associative learning over successive rounds of testing.

Note that the frequency of the character-pairs is constant across weeks and conditions. Only the constituent characters within a pair differ in frequency. Error bars reflect +/- 2 SEs

Discussion

It has long been known that learning is better with more practice. The novel finding is that it is easier to learn new facts that are composed of more familiar elements. Previously, we knew that information that can be chunked can be more easily retained, but chunks were treated as all or nothing, not as graded entities. This study demonstrated that with each week of training, learning improved for completely new associations of English words to a *new combination* of two Chinese characters. Given that the pairing of characters changed each week, one might have thought that performance would be hurt from such interference. Quite the contrary; learning became faster. Furthermore, this improvement was not due to practice at the task per se, because the characters that experienced more familiarization each week were always learned faster in new combinations than the less familiarized characters, week after week.

Fig. 4 Accuracy on N-Back task for low and high-frequency characters for the three levels of N-back difficulty. Error bars reflect +/- 2 SEs

The reason for this improvement, we contend, is that less working memory is required to encode more familiar stimuli. This was demonstrated in the *N*-back task and also in better performance for the more familiar characters on the visual search task, session after session. Indeed, there is close correspondence between improvement (speed and accuracy) in the visual search task and improvement in learning new knowledge structures (see Figs. 2 and 3). It is important to stress that there were new pairings of characters for each week's test. The familiarity of the characters varied as a function of week of training and frequency of exposure during training, but a given pair was always new for each test. Nevertheless, new pairings consisting of more familiar characters were learned more easily. This effect cannot be attributed to ease of encoding per se because when lists contain both low-frequency and high-frequency pairs, the advantage dissipates. This mixed-list effect also occurs with real words of different frequencies.
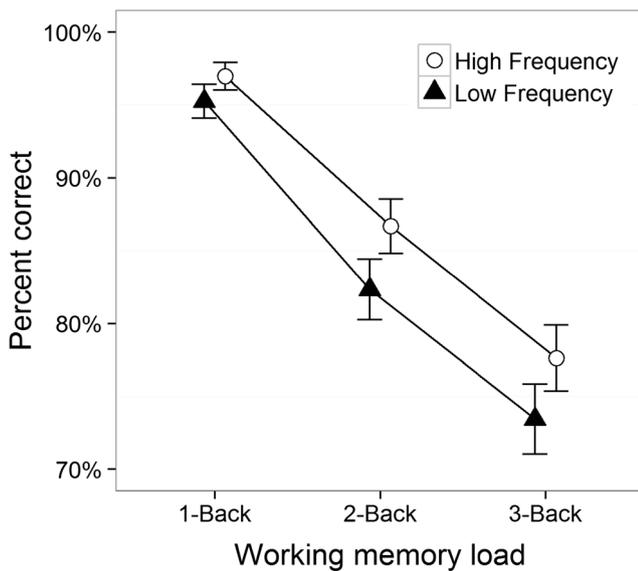
While we have long known that more information can be encoded when it can be chunked into higher level knowledge units, the conventional wisdom has been that chunks are all or none. We demonstrated that success in forming new, complex knowledge structures depends critically on the strength of each constituent chunk that needs to be integrated into that complex structure. In addition, the relationship between the familiarity of the chunks and working memory is one that had not been previously explored, but is one that is important to note in order to fully understand why the level of familiarity of components matters when acquiring new and complex knowledge.

While the thesis and evidence we put forth here are novel, our stance is consistent with the results of numerous other studies. For example, Nelson and Shiffrin (2013) also highlighted the importance of familiarity on learning new stimuli. They used an extended training paradigm to differentially familiarize subjects with previously unknown Chinese characters and then demonstrated effects of this differential exposure on perceptual and memory tasks for the individual characters. Our work builds on theirs by experimentally

demonstrating that familiarity not only affects single item performance but also influences one's ability to form complex *novel* associations based on these items.

Our research also challenges the common assumption that children have less working memory than adults (Elman, 1993). It may be that children have at least as much capacity as adults, but they have fewer chunks and those are weaker. Encoding or manipulating weak chunks or information that is not yet chunked consumes more working memory, making children seem to have less capacity than adults.

The pedagogical implications of this research are clear: as teachers and as learners, we need to be more mindful of the amount of novel information we try to present and digest at once. When trying to teach a new skill such as a language, music, or math, the components must be introduced slowly so that the lower level elements become strong chunks before instruction that requires many new elements to be combined. For example, teaching algebra or calculus requires introducing many symbols that are initially unfamiliar to the student. If there is too much unfamiliar information to be processed, learning will fail or be imperfect, at best. Familiarizing the elements so that they become strong chunks, or "bricks," will allow for successful construction. Otherwise the noncohesive elements will only enable a structure made of sand that quickly collapses.

Author contributions L.R. developed theory, study concept, and design. All authors contributed to data analyses and write-up.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412.

Bird, C. P., Nicholson, A. J., & Ringer, S. (1978). Resistance of the spacing effect to variations in encoding. *The American Journal of Psychology, 91,* 713–721.

Blumenfeld, R. S., & Ranganath, C. (2006). Dorsolateral prefrontal cortex promotes long-term memory formation through its role in working memory organization. *Journal of Neuroscience, 26,* 916–925.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage, 5,* 49–62.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Criss, A. H., Aue, W. R., & Smith, L. (2011). The effects of word frequency and context variability in cued recall. *Journal of Memory and Language, 64*(2), 119–132.

DeWitt, M. R., Knight, J. B., Hicks, J. L., & Ball, B. H. (2012). The effects of prior knowledge on the encoding of episodic contextual details. *Psychonomic Bulletin & Review, 19*(2), 251–257.

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 805–815.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology (H. A. Ruger & C. E. Bussenius, Trans.)*. New York, NY: Dover.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*(1), 71–99.

Ericsson, K. A., Krampe, R. T., & Teschromer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13,* 8–20.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*(2), 95–112.

Greene, R. L. (1989). Spacing effects in memory—Evidence for a 2-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 371–377.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77–99). Potomac, MD: Erlbaum.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319,* 966–968.

Kirchner, W. K. (1958). Age-differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55,* 352–358.

Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language, 63*(1), 46–63.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Multhaup, K. S., Balota, D. A., & Cowan, N. (1996). Implications of aging, lexicality, and item length for the mechanisms underlying memory span. *Psychonomic Bulletin & Review, 3*(1), 112–120.

Nelson, A., & Shiffrin, R. (2013). The co-evolution of knowledge and event memory. *Psychological Review, 120,* 356–394.

Nimmo, L. M., & Roodenrys, S. (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied PsychoLinguistics, 23*(04), 643–659.

Poppenk, J., Kohler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1321–1330.

Postman, L., & Knecht, K. (1983). Encoding variability and retention. *Journal of Verbal Learning and Verbal Behavior, 22*(2), 133–152.

Reder, L. M., Angstadt, P., Cary, M., Erickson, M. A., & Ayers, M. (2002). A reexamination of stimulus-frequency effects in recognition: Two mirrors for low- and high-frequency pseudowords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 138–152.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 294–320.

Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., & Dickison, D. (2007). In B. Ross & A. S. Benjamin (Eds.), *The psychology of learning and motivation* (pp. 271–312). New York, NY: Academic Press.

Reder, L. M., Victoria, L. W., Manelis, A., Oates, J. M., Dutcher, J. M., Bates, J. T., … Gyulai, F. (2012). Why it's easier to remember seeing a face we already know than one we don't: Pre-existing memory representations facilitate memory formation. *Psychological Science, 24*(3), 363–372.

Ross, B. H., & Landauer, T. K. (1978). Memory for at least one of two items: Test and failure of several theories of spacing effects. *Journal of Verbal Learning and Verbal Behavior, 17*(6), 669–680.

Simon, H. A. (1974). How big is a chunk? *Science, 183*(4124), 482–488.

Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term-memory encoding. *Psychonomic Bulletin & Review, 2,* 387–390.

Tulving, E., Markowitsch, H. J., Craik, F., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex, 6,* 71–79.