Mem Cogn
DOI 10.3758/s13421-015-0517-3

Cue generation: How learners flexibly support future retrieval

Jonathan G. Tullis · Aaron S. Benjamin

© Psychonomic Society, Inc. 2015

Abstract The successful use of memory requires us to be sensitive to the cues that will be present during retrieval. In many situations, we have some control over the external cues that we will encounter. For instance, learners create shopping lists at home to help remember what items to later buy at the grocery store, and they generate computer file names to help remember the contents of those files. Generating cues in the service of later cognitive goals is a complex task that lies at the intersection of metacognition, communication, and memory. In this series of experiments, we investigated how and how well learners generate external mnemonic cues. Across 5 experiments, learners generated a cue for each target word in a to-be-remembered list and received these cues during a later cued recall test. Learners flexibly generated cues in response to different instructional demands and study list compositions. When generating mnemonic cues, as compared to descriptions of target items, learners produced cues that were more distinct than mere descriptions and consequently elicited greater cued recall performance than those descriptions. When learners were aware of competing targets in the study list, they generated mnemonic cues with smaller cue-to-target associative strength but that were even more distinct. These adaptations led to fewer confusions among competing targets and enhanced cued recall performance. These results provide another example of the metacognitively sophisticated tactics that learners use to effectively support future retrieval.

Keywords Cue generation · Metacognition · Control

J. G. Tullis (✉)
Department of Educational Psychology, University of Arizona, 1430 E. Second Street, Tucson, AZ 85721, USA
e-mail: jonathantullis@gmail.com

A. S. Benjamin
Department of Psychology, University of Illinois at Urbana–Champaign, Champaign-Urbana, IL, USA

When taking notes in meetings, making to-do lists, outlining readings, and naming computer files, learners support later retrieval by generating cues for their future selves. Learners create external cues to reduce the demands placed on their limited memories. The ability of a learner to remember target information may fade over time; a good external cue can sustain memory retrieval in the face of considerable forgetting. Generating external cues to support later retrieval plays a vital role in many real-world activities and carries serious consequences. For example, patients often set up memory cues, like pill boxes and cell phone reminders, to aid in their adherence to medications. In the current experiments, we examined how learners generate cues to support later memory retrieval. More specifically, we analyzed how the characteristics of self-generated cues differed under various instructional conditions and with various stimulus sets. In the first trio of experiments (Experiments 1–3), we examined how learners intentionally generate cues for themselves and evaluated the characteristics of intentional mnemonic cues. In the second series of experiments (Experiments 4–5) we analyzed whether learners adjust their cues based upon the characteristics of the to-be-remembered set of stimuli.

Although there is not a large body of research on cue generation in mnemonic tasks, learners must exercise accurate metacognitive control when generating good mnemonic cues in order to support later memory. Therefore, we will address metacognitive control and the qualities of good mnemonic cues before describing the current experiments.

Metacognitive control

Learners exercise strategic control over their memories during both encoding (Benjamin, 2008; Kornell & Metcalfe, 2006; Tullis & Benjamin, 2011, 2012) and retrieval (Goldsmith & Koriat, 2007). In fact, individuals' differing mnemonic

Published online: 17 March 2015

abilities may arise largely from differences with which learners exercise strategic control over encoding and retrieval processes (Benjamin, 2008; Fiechter, Benjamin, & Unsworth, 2015; Finley, Tullis, & Benjamin, 2009). Learners deliberatively and strategically choose encoding tactics (Allen, 1968), study activities (Finley & Benjamin, 2012; Karpicke, 2009; Tullis, Benjamin, & Fiechter, 2015), study time allocation policies (Tullis & Benjamin, 2012; Tullis, Benjamin, & Liu, 2014; Son & Metcalfe, 2000), and study schedules (Benjamin & Bird, 2006; Son, 2004) to boost mnemonic performance. However, how learners exercise metacognitive control over testing circumstances is largely unexplored.

Learners can often strategically control their testing circumstances by generating external cues. Good external cues remove the challenging demands of memory from the mind and offload them to the physical environment. Harris (1980) surveyed learners about the external cues used to aid retrieval and found widespread use of a variety of different external aids, including putting items in special places to remind them of something, writing on a calendar, making notes to oneself, and asking others to remind them. The effectiveness with which learners use such external support has been most thoroughly examined with respect to prescription adherence (Caranasos et al., 1974), and has shown that learners often fail to utilize effective external cues (Haynes, McKibbon, & Kanani, 1996; Park, Morrell, Frieske, & Kincaid, 1992; Piette, Weinberger, Kraemer, & McPhee, 2001). Generating effective cues requires that learners set up circumstances to enhance the probability of being reminded of the appropriate action (Tullis, Braverman, Ross, & Benjamin, 2014; Tullis, Benjamin, & Ross, 2014) rather than merely increase the current memory strength of the target item. To do so effectively, learners must predict their future cognitive context and the future state of the world, an act that may be very challenging.

Few studies have examined how learners intentionally control their retrieval environment, and they have yielded mixed results of giving learners control over testing circumstances. In one study, learners chose how they would be tested on individual items (Finley & Benjamin, 2015a). Learners studied targets with either rhyming or semantic cues and then chose what type of cue (a rhyme or semantic associate) they wanted to receive during the test. Learners did not preferentially select testing circumstances that matched their encoding conditions and so failed to effectively control their testing circumstances to take advantage of transfer-appropriate processing (Fisher & Craik, 1977).

However, another study showed that learners effectively select cues from a set of options to support later memory performance (Finley & Benjamin, 2015b). In this study, learners selected two out of four possible cues for each target to receive during the memory test. Sometimes learners' selections were honored and the selected cues were presented at test; other times two random cues out of the possible four were presented instead. Learners recalled more targets when they received the cues that they chose compared to when they received random cues. Learners thus effectively set up advantageous testing conditions by choosing cues to support later memory performance. We extend these results by examining how effectively learners generate cues on their own.

Related studies have examined how learners generate descriptions of target items and have shown that learner-generated descriptions support very high levels of memory performance, even for long lists of items (Hunt & Smith, 1996; Mäntylä, 1986). Learners generate descriptions that take advantage of their own idiosyncratic understanding and conceptualizations of target items, and using these idiosyncratic descriptions during a cued recall test is very beneficial for memory. However, in this prior work, learners never generated mnemonic cues—they only generated descriptions of the target items. If learners' metacognition about retrieval cues is accurate and efficient, the mnemonic cues that learners generate may differ substantially from the descriptions that they generate.

Qualities of good cues

Learners' self-generated cues often fail to support retrieval in everyday life. People struggle to understand their notes from a class, do not know what computer file a file name refers to, and forget to take their medication. These failures reveal that the cues spontaneously utilized by learners do not flawlessly support retrieval. Weekly phone calls, pill boxes, and beepers have been implemented to improve medication adherence, and these interventions often succeed in increasing the rates with which patients consistently take their prescription (Lachowsky & Levy-Toledano, 2002; Park et al., 1992; Piette et al., 2001). This evidence shows that the quality of the cue makes a big difference in whether target information is recalled and desired tasks are completed.

Research in both retrospective and prospective memory suggests that good cues have the following three properties: They are strongly associated to the target, they are distinctive, and they are consistent across encoding and retrieval. First, the association between the target and cue influences how much is recalled. Cued recall performance increases as the cue-to-target associative strength in a word pair increases (Feldman & Underwood, 1957; Koriat & Bjork, 2006). Similarly, cues in a prospective memory task that are more strongly related to the intended action result in higher prospective memory than ones less closely related (Einstein, Holland, McDaniel, & Guynn, 1992; McDaniel, Guynn, Einstein, & Breneiser, 2004).

The second characteristic of cues that determines their effectiveness is how distinctively they enable target processing within a context. Different theories outline what distinctive processing entails, with some arguing that it is the unique processing of an item at encoding that enhances the discriminability of that item at retrieval (Jacoby & Craik, 1979) and others suggesting that it is the encoding of differences among elements that are similar on some dimension (Hunt, 2006, 2012). Nairne (2006) argues that distinctiveness specifies a particular stored event to the exclusion of others; therefore, distinctiveness is a function of both how strongly associated the cue is to the target and how many other possible targets the cue is associated with. Distinctive processing of cues supports memory because unusual features can guide access or direct retrieval to the distinct target item to the exclusion of other possible targets (Hunt, 2006; Waddill & McDaniel, 1998). The efficacy of a retrieval cue is lessened when it points to a large number of targets (Anderson, 1974; Watkins & Watkins, 1975). Unique cues that point to attributes of targets that are not shared by other studied targets provides distinctive processing of targets and support better later recall (Hunt & Smith, 1996). Furthermore, research in prospective memory suggests that unusual cues support prospective memory better than common words because they have fewer extra-list associations (Einstein & McDaniel, 1990). Some argue that a cue's distinctiveness is the single most important attribute of a cue for determining if a target memory is recalled or forgotten (Nairne, 2002).

Third, the consistency between cues at encoding and retrieval determines how much is recalled. Cues present during both encoding and retrieval are much more beneficial for retrieval than cues that are present only during one phase (Tulving & Osler, 1968). In fact, the effectiveness of a cue may depend much less upon the cue-to-target associative strength than upon the match between encoding and retrieval. In one famous example (Tulving & Thompson, 1973), learners encoded cue-target word pairs with weak cue-to-target associative strengths, but those weak cues were much more effective than strongly associated (but unstudied) cues at supporting cued recall. In another set of experiments, during a cued recall test, learners received descriptions of target items that they or other subjects generated during encoding. Learners recalled far fewer words when they received descriptions generated by another learner compared to when they received their own descriptions (Andersson & Ronnberg, 1997; Hunt & Smith, 1996; Mäntylä, 1986). Interpretations during encoding are idiosyncratic among learners, and learners' benefit from receiving retrieval cues consistent with their own encoding processes. Throughout this paper, cue-to-target associative strength, distinctiveness, and match between encoding and retrieval will be considered in order to describe the types and effectiveness of cues that learners generated.

Operationalizing qualities of good cues

Before describing the individual experiments, the theoretical constructs of cue-to-target associative strength, distinctiveness, and match between encoding and retrieval must be operationalized. Cue-to-target associative strength was determined using the normative cue-to-target associative strength found in the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998). Although idiosyncratic cue-to-target associative strength may play a larger role than normative cue-to-target associative strength in recall, idiosyncratic cue-to-target associative strength cannot be easily determined for individual learners. The normative cue-to-target associative strength reported in this database is the best approximation we currently have available.

Distinctive processing, the processing of features that serve to restrict the later search space, was operationalized using a variety of related measures from the University of South Florida Free Association Norms. Distinctiveness was calculated using the database because the targets were collected from it, and therefore, it serves as a possible context or list of potential target items. Our first operationalization of distinctiveness was the number of targets associated to each cue, with more distinctive cues associated to fewer targets. To characterize the strength of all cue-to-target relationships, the total cumulative cue-to-target associative strength was also calculated. The total cumulative cue-to-target associative strength is the sum of the cue-to-target associative strengths from a cue to all associated targets in the database. The larger the total cumulative strength from a cue to all possible targets, the less distinct a cue is. Finally, if a cue was not located in the South Florida Free Association Norms, it was assumed to be (relatively) distinctive.

The quality of match between the internal and external cues present at encoding and retrieval was manipulated by providing cues generated by oneself or by others during the cued recall test. Even when highly associated distinctive cues are generated, those cues are likely to be less effective when they are not consistent with the learner's interpretation and encoding of the target. Consequently, cues generated by oneself are more likely to be consistent across encoding and retrieval than ones generated by others, even when those generated by others are normatively more effective.

Experiments

In these studies, learners were explicitly asked to generate cues to help their later recall. Despite the fact that this situation arises quite often in daily life, oddly, no prior study has examined how learners do so, and how successful they are. Instead, across the few prior studies with self-generated cues (Hunt & Smith, 1996; Mäntylä, 1986), learners have been asked to generate descriptions of a set of nouns without any foreknowledge of the upcoming cued recall task. We examined the more metacognitively relevant case where learners generate cues specifically in anticipation of using those cues to support their later retrieval. We specifically examined whether learners value distinctiveness and cue-to-target associative strength, and how the balance between those factors shifted across different conditions. In general, one would think that learners should be very effective at generating mnemonic cues to support future cued recall because they can rely upon their rich, idiosyncratic knowledge to produce stable, distinctive cues with high cue-to-target associative strength.

Experiment 1

In the first experiment, we evaluated whether learners intentionally generate mnemonic cues that are more effective than descriptions of to-be-remembered items. In other experiments, learners have generated their own cues (Bäckman & Mäntylä, 1988; Hunt & Smith, 1996; Mäntylä, 1986; Mäntylä & Nilsson, 1983, 1988), but in all of those experiments, learners had no foreknowledge that those descriptions would be presented as cues during a later memory test. Further, all extant studies using this procedure except one (Bäckman, Mäntylä, & Erngrund, 1984) have utilized incidental learning conditions and surprise cued recall tests to measure the effectiveness of self-generated descriptions. Here we assessed whether knowing that items will be used as cues during a future cued recall test impacted how learners generate their cues. In a between-subjects manipulation, learners were either asked to generate descriptions of each target item, like in prior studies (Mäntylä, 1986), or were asked to generate cues that would be given back to them during a later memory test. Learners were always informed of the upcoming memory test for the target words. If learners are effective at generating cues to support memory performance, learners in the cue generation condition should show better cued recall performance than learners in the description generation condition, even though both groups expect a final memory test.

Participants

Fifty introductory psychology students at the University of Illinois at Urbana–Champaign participated for partial course credit.

Materials

Sixty nouns that college-aged subjects would likely have some personal experiences with were collected from the University of South Florida Free Association Norms (Nelson et al., 1998). Examples of words included *dancing, haircut*, and *roommate.* The Thorndike-Lorge written frequencies ranged from 27 to 2,218, with a mean of 536 and a standard deviation of 510.

Procedure

The experiment was presented using the Psychophysics toolbox in MATLAB (Brainard, 1997) on personal computers in individual testing rooms. Instructions about the memory task were first presented on personal computer screens; these indicated that subjects would study a list of 60 words and would later be tested on their memory for these words. Subjects were assigned to either the *description generation* or the *cue generation* condition according to a counterbalancing scheme that evenly distributed learners across conditions. The counterbalancing scheme further yoked half of the subjects to prior subjects in their same condition and half to prior subjects in the other condition. The first subject on each of the six different computers received all of the cues that they generated. All other subjects were yoked to the immediately prior subject on that same computer. Half of the cues that these subjects received during the test were generated by them and half were generated by the previous subject.

The 26 subjects in the *description* condition received the following instructions:

> For each target word, you will need to generate some aspect of the word that constitutes an appropriate description of the target item. This aspect or description can be created according to your own life experiences. Your memory for the target words will be tested at the end of the experiment. Once again, for each target in the list, we ask that you type in one word that, according to your own experiences, describes the target word or is an aspect of the target word. You can use any one word description. However, the target word cannot serve as its own description.

The 24 subjects in the cue generation condition received the following instructions:

> For each target word, you will generate some type of cue to help you remember the target word at a later test. At the time of the test, you will be given the cue word you generated and will be asked to remember the specific target word for each cue. Please generate the cues that will be most useful to you in remembering the target words. You can use any one word cue. However, the target word cannot serve as its own cue”

Subjects studied the to-be-remembered targets one at a time in a random order in black 25-point Arial font on the right side of the computer screen. To the left of each target item, an empty cue box was presented. For each word, subjects typed a single word into the cue box and pressed the return key. If the subjects attempted to type a space, the program reminded learners that they should only generate one word per target item and did not record the space. After generating cues for all of the items, subjects completed an unrelated face memory task for approximately 10 minutes. Subjects finally completed a cued recall test of the target items. Subjects were informed, "For some cue-target pairs, you will be given the cue word that another person generated for that target. We still ask you to try your best to recall the target for every cue word." The targets and cue types were randomly ordered during the recall test. Subjects were presented with a cue on the left side of the screen and were asked to type the corresponding target item in the empty response box on the right side of the screen. As during the study list, subjects proceeded through the test at their own rate.

Results

Cued recall performance in all experiments was assessed using strict scoring, such that only the responses that matched the target identically were counted as correct. For this and for Experiment 2, the data from the first subjects in each room were not included in the data analysis. These subjects were excluded because they received only the cues they generated (and never cues from others).

For each experiment, we analyzed characteristics of the cues between conditions, cued recall performance, and the characteristics of effective cues. Learners utilized a variety of different cue generation strategies to enable future target retrieval. The strategies subjects used to generate cues across all experiments are categorized in Appendix A. For the analyses presented across all experiments, we included cues generated using all of strategies in order to avoid any potential item effects; however, all conclusions hold if we only analyze cues using the one-word strategy.

Cue characteristics

The cue-to-target associative strength and the distinctiveness of the cues generated between conditions are presented in Table 1. Significant differences in the distinctiveness of the cues were found between conditions. The cues generated by learners in the cue generation condition were associated to fewer possible targets, $t(42) = 2.55$, $p = .01$, $d = 0.79$, had smaller cumulative cue-to-target associative strength to all possible targets, $t(42) = 2.42$, $p = .02$, $d = 0.75$, and were less likely to be in the South Florida database, $t(42) = 2.55$, $p = .01$, $d = 0.79$) than cues generated in the description condition. Cue uniqueness, which we believe to be a measure of cue idiosyncrasy, was calculated by measuring the number of different cues that were provided for each target divided by the total number of subjects in the condition. A greater proportion of cues were unique under cue generation instructions ($M = 0.72$) than under description generation instructions ($M = 0.63$), $t(59) = 6.29$, $p < .001$, $d = 0.82$. The cue-to-target associative strength did not differ between conditions, $t(42) = 0.63$, $p = .53$, $d = 0.017$. The Pearson intercorrelation matrix for these measures is displayed in Appendix E.

Memory performance

Cued recall performance is displayed in Fig. 1. Learners remembered more items when they received their own cues rather than another subject's cues, $t(43) = 15.54$, $p < .001$, $d = 1.95$. Further, learners who received their own *cues* remembered more than learners who received their own *descriptions,* $t(43) = 3.70$, $p < .001$, $d = 1.13$. Learners who received others' *cues* also remembered numerically more than learners who received others' *descriptions*, though this effect was not significant, $t(43) = 1.67$, $p = .11$, $d = 0.51$.

Characteristics of effective cues

The characteristics of cues or descriptions that led to correct recall were compared with those that led to failed recall; the results are displayed in Table 2. The characteristics of effective cues depended upon the cue originator. The following

Table 1 Characteristics of cues generated under different instructions in Experiment 1. Boxes highlighted in gray show significant differences between instruction conditions

	Description	Cue generation
Cue-to-target associative strength	0.051 (0.025)	0.056 (0.027)
Number of targets associated to cue	10.10 (1.81)	8.23 (2.86)
Cumulative associative strength from cue	0.58 (0.08)	0.49 (0.16)
Proportion of cues in the database	0.73 (0.10)	0.61 (0.20)
Cue uniqueness	0.63 (0.14)	0.72 (0.15)

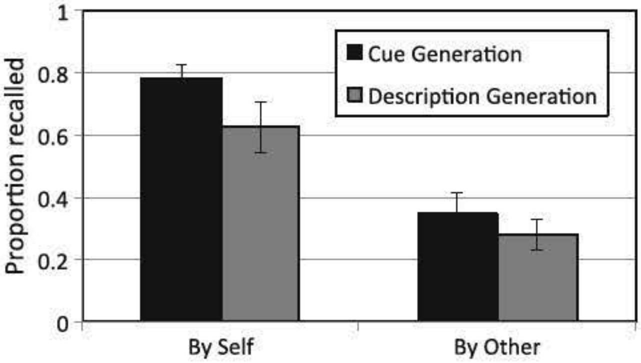

Fig. 1 Cued recall performance by cue originator and the cue originators' instruction condition for Experiment 1. Error bars indicate 95 % confidence intervals for the mean of each column

analyses are correlational in nature and are corrected for multiple comparisons using Bonferroni corrections such that only analyses significant at an alpha level of less than 0.005 are considered significantly different. Successful cues generated by oneself were more highly associated to the target and were more distinctive than unsuccessful cues. Effective cues generated by oneself had higher cue-to-target associative strength, $t(44) = 6.55$, $p < .001$, $d = 1.00$; fewer targets associated with them, $t(44) = 4.16$, $p < .001$, d = 0.63); smaller cumulative associative strength to all normed targets, $t(44) = 4.49$, $p < .001$, $d = 0.68$; and were less likely to be listed in the South Florida database, $t(44) = 4.43$, $p < .001$, $d = 0.67$, than unsuccessful cues. Although both cue-to-target associative strength and distinctiveness were beneficial when cues were generated by oneself, only cue-to-target associative strength impacted the effectiveness of a cue when it was generated by another. Effective cues by others had higher cue-to-target associative strength than ineffective cues, $t(42) = 5.26$, $p < .001$, $d = 0.83$. Effective and ineffective cues by others did not differ in measures of distinctiveness: The number of targets associated to a cue was similar, $t(42) = 1.72$, $p = .09$, $d = 0.27$; the total cumulative associated strength was similar, $t(42) = 0.41$, $p = .68$, $d = 0.06$; and the likelihood of being in the South Florida database was similar, $t(42) = 0.06$, $p = .95$, $d = 0.01$.

The characteristics of effective cues were further analyzed using a hierarchical linear model (HLM), as shown in Appendix B. A hierarchical linear model provides some advantages over the traditional method outlined above (Baayen, Davidson, & Bates, 2008). First, hierarchical linear models simultaneously consider the contributions of the different explanatory variables. The HLM can show which cue characteristics meaningfully contribute to predictions about recall performance and which cue characteristics overlap with others and do not improve the predictive ability of the model. This statistical analysis also lessens concerns about multiple comparisons that are espoused above (Gelman, Hill, & Yajima, 2012). Second, the HLM can easily analyze the interactions between cue originator and the effectiveness of each cue without computing multiple ANOVAs. Finally, the HLM increases the power of the analysis because it considers each target item a unit of analysis rather than each subject. The conclusions of the HLM analysis replicate those found in the traditional analysis: Cue generation instructions increased eventual recall, self-generated cues elicited higher recall than other-generated cues, greater cue-to-target associative strength enhanced recall regardless of cue originator, and the influence of cue distinctiveness depended upon cue originator.

Discussion

Learners made deliberate and purposeful choices when generating cues to support their later memory performance that were different than those made when providing descriptions of the targets. Mnemonic cues were more distinct than descriptions. Under cue generation instructions, cues were associated to fewer possible target items, had smaller cumulative cue-to-target associative strength, and were less likely to be found in the association database. Learners selected cues to support memory that were more idiosyncratic than learners who generated descriptions of items, as shown by a measure of cue uniqueness. Instructions not only changed the types of cues that learners generated, but also led to differences in cued

Table 2 Characteristics of cues that led to successful and unsuccessful retrieval split by cue originator in Experiment 1. Significant differences between effective and ineffective cues are highlighted in gray. Numbers in parentheses indicate standard errors of the mean

	Cue by oneself		Cue by other	
	Correct	Incorrect	Correct	Incorrect
Cue-to-target associative strength	0.06 (0.04)	0.02 (0.03)	0.11 (0.09)	0.04 (0.02)
Number of targets associated to cue	8.39 (2.90)	10.30 (3.74)	8.71 (2.63)	9.33 (3.08)
Cumulative associative strength from cue	0.50 (0.15)	0.61 (0.22)	0.56 (0.14)	0.54 (0.16)
Proportion of cues in the database	0.63 (0.19)	0.75 (0.27)	0.69 (0.17)	0.67 (0.20).

recall. Cues generated under cue generation instructions supported higher levels of cued recall than cues generated under the description generation instructions, especially when those cues were given back to the cue generator.

Learners' self-generated cues supported retrieval more than others' cues. This result highlights the importance of match between encoding and retrieval. Learners generated a variety of different cues during encoding according to their own idiosyncratic interpretation of the target. Learners' reliance on idiosyncratic knowledge prevents their cues from being as beneficial when presented to others. Other learners have not encoded the targets in the same context and, therefore, do not benefit as much when receiving another subject's cues during retrieval.

Experiment 2

In Experiment 2, the cue generation condition from Experiment 1 was replicated with a new set of items in order more powerfully analyze the qualities of successful and unsuccessful memory cues and to generalize results to a greater variety of stimuli. The impact that cue-to-target associative strength and cue distinctiveness had on memory performance was more cleanly assessed without feeding cues generated under different instructions to both conditions.

Participants

Forty-two introductory psychology students at the University of Illinois at Urbana–Champaign were run across six different computers.

Materials

A new set of target words was used to ensure generalization of the findings across stimuli. The word list comprised nouns, verbs, and adjectives and was not selected with the intention that students had experiences with the words. Example words included *banjo, humid*, and *grasp*. With the Thorndike–Lorge written frequencies ranging from 2 to 392, with a mean of 113 and a standard deviation of 101, these stimuli had significantly lower frequencies than those used in the prior experiment.

Method

The procedure of Experiment 2 was identical to that of Experiment 1, except that all subjects received cue generation instructions. The data from each of the first subject on each of the six computers were not analyzed because they received only their own generated cues during the test. Therefore, analyses include only data from 36 subjects.

Results

Memory performance

Subjects recalled significantly more targets when they received their own cues (M = 0.83, SD = 0.12) than another subject's cues (M = 0.40, SD = 0.18), $t(35) = 14.42$, $p < .001$, $d = 2.48$. The relationships among the measures of cue characteristics are displayed in Appendix E.

Characteristics of effective cues

The cue-to-target associative strength and distinctiveness of effective cues were compared to those of ineffective cues, and all results replicated those found in Experiment 1. The mean values as a function of cue originator are displayed in Table 3. Cues generated by oneself benefited from higher cue-to-target associative strength and distinctiveness; cues generated by others benefited only from higher cue-to-target associative strength. Effective cues generated by oneself had higher cue-to-target associative strength, $t(29) = 5.42$, $p < .001$, $d = 1.04$[1]; were associated to fewer targets, $t(29) = 2.52$, $p = .02$, $d = 0.47$; had a smaller cumulative associative strength to all targets, $t(29) = 2.89$, $p = .007$, $d = 0.54$; and were less likely to be in the South Florida database, $t(29) = 3.52$, $p = .001$, $d = 0.66$, than ineffective cues. Effective cues generated by others had higher cue-to-target associative strengths, $t(35) = 6.53$, $p < .001$, $d = 1.10$, than ineffective cues. No measure of distinctiveness differed between effective and ineffective cues generated by others: the number of targets associated to a cue, $t(35) = 1.18$, $p = .24$, $d = 0.20$; the cumulative associative strength from a cue to all targets, $t(35) = 0.85$, $p = .40$, $d = 0.14$; and the likelihood of being in the South Florida database, $t(35) = 0.12$, $p = .90$, $d = 0.02$, did not differ between effective and ineffective cues. For the corresponding HLM analysis of the cue characteristics that support recall, see Appendix C. Again, the central results of the HLM analysis replicated those found in the traditional analysis.

Discussion

Results replicated those from Experiment 1. Learners' mnemonic performance benefited greatly when learners received their own cues at the time of the test, indicating both the importance of the match between encoding and retrieval and learners' idiosyncratic encoding of target items. Cue-to-target associative strength and cue distinctiveness significantly moderated the effectiveness of a cue, as would be expected given

[1] Six subjects recalled all of the targets associated with their own cues. Therefore, they contribute no data to the ineffective cue conditions, and the degrees of freedom for these analyses are reduced.

Table 3 Characteristics of cues that led to successful and unsuccessful retrieval for cues generated by oneself and by another subject (Experiment 2). Significant differences between successful and unsuccessful cues are highlighted by gray backgrounds. Standard deviations of the mean are indicated in the parentheses

	Cue by oneself		Cue by other	
	Correct	Incorrect	Correct	Incorrect
Cue-to-target associative strength	0.026 (0.021)	0.005 (0.009)	0.076 (0.06)	0.009 (0.01)
Number associated to cue	9.68 (3.26)	11.53 (3.63)	10.06 (2.87)	10.75 (2.62)
Cumulative associative strength from cue	0.56 (0.17)	0.69 (0.21)	0.63 (0.16)	0.61 (0.14)
Cue in the database	0.69 (0.21)	0.86 (0.24)	0.76 (0.19)	0.75 (0.17)

prior research. Cues generated by oneself benefited from both increased cue-to-target associative strength and cue distinctiveness. Interestingly, cues generated by others benefited from greater cue-to-target associative strength, but were unaffected by a cue's distinctiveness. Other-generated distinctive cues may have been too distinctive and idiosyncratic to help a different learner who has never experienced the cue with the target.

Experiment 3

In Experiments 1 and 2, we cannot tease apart the contributions to memory played by differentially effective cues or by differentially effective encoding processes. Generating a cue may engender deeper encoding processes than generating a description, and these differences in encoding may cause the differences seen in cued recall. "Cues" and "descriptions" presented at the time of the cued recall test may not support memory differentially; the mnemonic differences may be partially or entirely caused by differentially effective encoding processes. We tested this hypothesis in Experiment 3 by utilizing the same study/generation procedure, but tested learners' memories with a free recall procedure.

Participants

Forty-four introductory psychology students at Indiana University participated for partial course credit.

Materials

The items used in Experiment 3 were the same as those used in Experiment 1.

Method

The study phase of this experiment proceeded just as in Experiment 1. Learners were alternatively assigned to a "cue generation" or "description generation" condition. Unlike Experiments 1 and 2, at the time of the test, learners were asked to type in as many of the targets as they could. Learners were not presented with any cues or descriptions during the free recall test.

Results

Cue characteristics

The cue-to-target associative strength and the distinctiveness of the cues generated between conditions were analyzed and are presented in Table 4. Significant differences in the distinctiveness of the cues were found between conditions, and entirely replicated those found in Experiment 1. The cues generated by learners in the cue generation condition were associated to fewer possible targets, $t(42) = 3.58$, $p < .001$, $d = 1.10$; had smaller cumulative cue-to-target associative strength to all possible targets, $t(42) = 3.12$, $p = .003$, $d = 0.96$; and were less likely to be in the South Florida database, $t(42) = 3.14$, $p < .001$, $d = 0.97$, than cues generated in the description condition. Furthermore, as in Experiment 1, a greater proportion of cues were unique under cue generation instructions ($M = 0.78$) than under description generation instructions ($M = 0.67$), $t(59) = 7.65$, $p < .001$, $d = 1.00$. The cue-to-target associative strength did not differ between conditions, $t(42) = 0.78$, $p = .44$, $d = 0.24$.

Memory performance

Learners in the "cue generation" condition recalled a smaller proportion of targets ($M = 0.15$, $SD = 0.13$) than learners in the "description generation" condition ($M = 0.19$, $SD = 0.12$),

Table 4 Characteristics of cues generated under different instructions in Experiment 3. Boxes highlighted in gray show significant differences between instruction conditions. Standard deviations of the mean are indicated in the parentheses

	Description	Cue generation
Cue-to-target associative strength	0.056 (0.027)	0.048 (0.037)
Number of targets associated to cue	10.31 (2.38)	7.46 (2.78)
Cumulative associative strength from cue	0.58 (0.11)	0.45 (0.17)
Proportion of cues in the database	0.73 (0.14)	0.57 (0.20)
Cue uniqueness	0.67 (0.13)	0.78 (0.12)

$t(42) = 1.00$, $p = .32$, $d = 0.32$. This difference between conditions is in the opposite direction of that seen in Experiment 1, thus ruling out the possibility that encoding differences caused the memory benefit seen in Experiment 1 and 2. Rather, it seems that the benefit that arises from using one's own cues derives from the cues themselves.

Characteristics of effective cues

The characteristics of cues or descriptions that led to correct recall were compared with those that led to failed recall. If the cue generation process affects later free recall performance, differences in the characteristics of generated cues that led to successful recall compared to unsuccessful recall should be apparent. If cue generation processes do not affect later free recall, no differences in the characteristics of cues that led to successful and unsuccessful recall should occur. As shown in Table 5, we found no significant differences in the cue characteristics between cues that led to successful and unsuccessful recall.

Table 5 Characteristics of effective and ineffective cues generated in Experiment 3. No significant differences were found between effective and ineffective cues in this free recall paradigm. Standard deviations of the mean are indicated in the parentheses

	Cue by oneself	
	Correct	Incorrect
Cue-to-target associative strength	0.05 (0.06)	0.05 (0.03)
Number of targets associated to cue	9.41 (3.75)	8.77 (2.88)
Cumulative associative strength from cue	0.55 (0.19)	0.51 (0.15)
Proportion of cues in the database	0.69 (0.27)	0.64 (0.19)

Discussion

Cue generation did not increase the amount of items recalled on a free recall test. Therefore, differences in recall between the cue generation and description conditions did not arise from differential encoding processes. The differences in cued recall apparent in Experiments 1 and 2, then, must have arisen from differences in the cues that learners generated. When generating cues, learners consistently generated words that were more distinctive than descriptions. However, cues and descriptions did not differ in cue-to-target associative strength. Whether learners are aware of the benefits and limitations of cue-to-target associative strength remains unsettled. In the next two experiments, increases in cue-to-target associative strength may actually cause confusions among target items and hinder overall cued recall. We examined, then, if learners can choose to decrease the cue-to-target associative strength of their cues in order to benefit their recall. To do this, we presented learners with lists that include some highly related targets.

Learners often need to recall information in the face of distracting competitors. Cueing memory to distinguish between similar targets is an essential skill needed to successfully navigate those competitors. If a cue leads to recall of a competing item, negative consequences may ensue. For instance, a cue to pick up your daughter from soccer practice that does not distinguish between an array of possible practice fields may fail as an effective cue and produce an angry daughter. Similarly, notes about applying permutations and combinations that do not clearly delineate the circumstances under which each is applicable may not enhance learners' grades. In the next pair of experiments, reducing the cue-to-target associative strength may enhance learners' performance. We examined if learners modulate the cue-to-target associative strength and distinctiveness of cues when generating them for a set of related and competing targets.

Evidence from the interpersonal communication literature suggests that speakers consistently consider competing referents when crafting messages. In fact, a referent's contrast set may have one of the largest influences on communication, as speakers need to identify a referent uniquely (Olson, 1970). For example, speakers provide subordinate level categorical information (e.g., collie) only when referents have competitors at the basic level (e.g., other dogs; Sedivy, 2003). Speakers use scalar adjectives (modifiers that reference an object's size) when drawing contrasts among competing referents (Yoon & Brown-Schmidt, 2013). By using subordinate-level categorical information or scalar adjectives, learners deliberately increase the distinctiveness of their message when competitors are present. Whether consideration of competitors plays a similar role in generating cues to communicate with a future self was explored in the next two experiments.

Across the next two experiments, learners generated cues to remember sets of related triplets. Learners read a set of three words and then generated a cue for each specific item in the set. To manipulate whether learners were likely to notice the relationships among competing targets, related triplets were either presented simultaneously (*together* condition) or spaced out in time (*apart* condition). These conditions were varied within-subjects in Experiment 4 and between-subjects in Experiment 5. We examined how effectively learners generated cues to differentiate among related targets, as well as the characteristics of the cues and the resultant cued recall performance.

Experiment 4

Participants

Thirty-four introductory psychology students at the University of Illinois at Urbana–Champaign participated for partial course credit.

Materials

Twenty triplets of synonyms were collected from the South Florida Free Association Norms (Nelson et al., 1998). Each to-be-remembered target then had two related to-be-remembered competitors. The triplets included both nouns (e.g., *quiz, exam, test*) and verbs (e.g., *irritate, annoy, bother*). The average associative strength among the members of each related triplet was 0.18.

Method

This experiment utilized a within-subjects design, with related triplets presented on the same screen simultaneously or presented across three different sets of items. Subjects were given the same cue generation instructions that were utilized in the prior two studies. Unlike the prior studies, however, three target items were displayed on the screen at once, as shown in Fig. 2. The three different targets appeared in a column in random order on the right-hand side of the screen for 6 seconds before any subject response was allowed. Presenting the targets together for 6 seconds before responses could be made was done to encourage subjects to read all three targets before generating cues. After 6 seconds, the first response box appeared next to the top target item and subjects typed in the first cue. Subjects then entered a cue for the middle target and finally entered a cue for the bottom target.

Three target items were always presented on each cue generation screen. In the together condition, the three target items came from the same related triplet. In the apart condition, three random unrelated target items were presented on screen together, and the related triplet items were randomly distributed across three different screens. Conditions were randomly assigned across presentation screens. After subjects completed the cue generation phase, they immediately took the cued recall memory test. One at a time, cues were presented on the left side of the screen in a random order, and subjects typed in their response in a box on the right-hand side of the screen.

Results

Memory performance

Cued recall performance is displayed in the left side of Fig. 3 and did not differ between the together and apart conditions, $t(33) = 0.86$, $p = .17$, $d = 0.13$. The number of confusions within each condition and within each subject was also calculated. An incorrect response was considered a confusion when the response was a competing target from the related triplet. Subjects made more confusions when cues were generated in the apart condition ($M = 2.53$) than when cues were generated in the together condition ($M = 1.38$), $t(33) = 2.19$, $p = .04$, $d = 0.44$.

Cue characteristics

The characteristics of the cues generated as a function of generation condition are displayed in the left columns of Table 6.

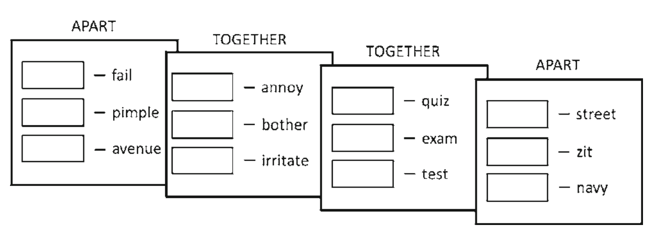

Fig. 2 The cue generation procedure screens utilized for Experiments 4 and 5

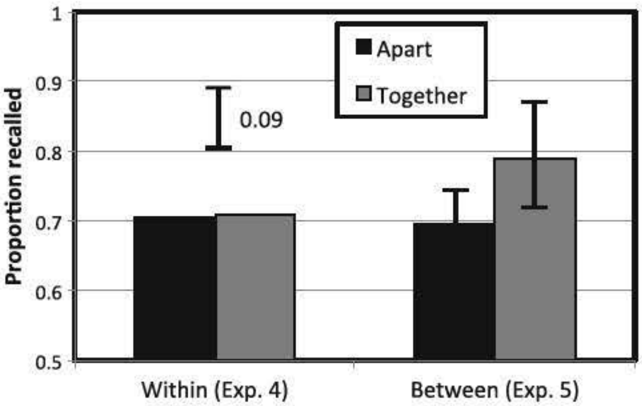

Fig. 3 Cued recall performance for Experiments 4 and 5. The error bar over Experiment 3 reflects the 95 % confidence interval of the difference between apart and together in Experiment 4 (Loftus & Masson, 1994). The error bars in Experiment 5 show the 95 % confidence intervals of the means of each condition

Subjects took more time to generate cues when related triplets were presented together than when they were presented apart ($t(33) = 2.85$, $p = 0.007$, $d = 0.50$). Subjects used the other items from the related triplet as cues for 28 % of the targets when the triplets were presented apart, and for only 10 % of the targets when triplets were presented together ($t(33) = 6.69$, $p < 0.001$, $d = 1.17$). The cue-to-target associative strength was greater in the apart condition than in the together condition ($t(33) = 4.39$, $p < 0.001$, $d = 0.95$).

Table 6 Dependent measures across Experiments 4 (left columns) and 5 (right columns). All comparisons displayed reached significance ($p < .05$). Numbers in parentheses indicate standard deviations of the means

	Experiment 4 (within subjects)		Experiment 5 (between subjects)	
	Apart	Together	Apart	Together
Time to generate cue (sec)	7.37 (3.66)	8.88 (5.01)	6.00 (1.99)	10.01 (4.19)
Assoc. strength from cue to target	0.10 (0.07)	0.05 (0.04)	0.13 (0.05)	0.05 (0.04)
Number of items associated from cue	7.94 (3.07)	6.88 (2.84)	9.45 (1.08)	6.60 (2.80)
Proportion of cues that were competitors	0.28 (0.20)	0.10 (0.12)	0.41 (0.16)	0.07 (0.11)
Assoc. strength from cue to competitors	0.05 (0.03)	0.03 (0.03)	0.08 (0.02)	0.02 (0.02)
Proportion of cues repeated in list	0.04 (0.05)	0.02 (0.04)	0.07 (0.04)	0.01 (0.01)
Number of confusions	2.53 (2.07)	1.38 (1.57)	7.79 (2.76)	1.95 (2.20)
Cue uniqueness	0.69 (0.14)	0.82 (0.13)	0.54 (0.16)	0.79 (0.13)

In the together condition, learners increased the distinctiveness of a cue compared to all normed items and relative to the other targets in the study list. The cues generated in the together condition were associated to fewer items in the South Florida database than those generated in the apart condition ($t(33) = 3.59$, $p = 0.001$, $d = 0.62$). Further, the associative strength between a cue and the two wrong (but related) targets was greater in the apart condition than in the together condition ($t(33) = 4.75$, $p < 0.001$, $d = 0.83$). The number of cues that a subject used for more than one target was also calculated. In the apart condition, subjects repeated cues more frequently (4 %) than in the together condition (2 %; $t(33) = 2.23$, $p = 0.03$, $d = 0.38$). Finally, cue overlap was calculated and showed that cues were more unique in the together condition than in the apart condition ($t(59) = 5.61$, $p < 0.001$, $d = 0.73$).

Discussion

When learners were likely to notice the competing target items, learners generated different cues than when they were not likely to notice the competing target items. For items presented with their competitors, learners took more time to generate cues, used fewer competitors as cues, used cues with smaller cue-to-target associative strength to the competitors, and generated cues that were associated to fewer possible target items. Learners recognized the difficulty of generating effective cues for the competing items when presented together and altered their cue generation process for these items. Although simultaneous and sequential presentation of the triplets changed the cue generation process, final cued recall performance did not benefit—though we will revisit this finding in Experiment 5.

Experiment 5

In Experiment 4, differences in cue generation between the together and apart conditions may have been artificially reduced because the manipulation was entirely within-subjects. After viewing related triplets presented simultaneously on screen, learners may have noticed the relationships among the triplets presented sequentially. Then learners may have adjusted their cue generation strategies and applied stricter criteria for cues from the apart items, which may have eliminated performance differences between together and apart items. It may also have influenced the retrieval strategies applied at test. To reduce the risk of this homogenization of these processes across conditions, conditions were varied between-subjects in Experiment 5. The between subjects manipulation utilized in Experiment 5 should increase the differences between conditions and provide a stronger test of the effects of that manipulation on accuracy.

Participants

Thirty-nine introductory psychology students at the University of Illinois at Urbana-Champaign participated for partial course credit.

Materials

The twenty triplets utilized here were the same as those in Experiment 4.

Method

The only change between the procedure utilized here and that of Experiment 4 was that the together and apart conditions were varied between-subjects. Subjects were alternatively assigned to the together and apart conditions. For subjects in the together condition, all cue generation screens presented related triplets simultaneously. For subjects in the apart condition, related triplets were always distributed across three different screens.

Results

Memory performance

Cued recall performance is displayed in Fig. 3. Learners in the together condition recalled significantly more targets than learners in the apart condition ($t(37) = 2.08$, $p = 0.04$, $d = 0.68$). As in the prior experiment, subjects made more confusions when cues were generated in the apart condition ($M = 7.79$) than when cues were generated in the together condition ($M = 1.95$; $t(37) = 7.13$, $p < 0.001$, $d = 2.34$)).

Cue characteristics

The characteristics of the generated cues are displayed in the right columns of Table 6, and amplify the differences found between conditions in the prior experiment. Subjects took more time to generate cues when related triplets were presented together than when presented apart, $t(37) = 3.68$, $p < .001$, $d = 1.21$. Subjects used the other items from the related triplet as cues for 41 % of the targets when the triplets were presented apart, and for only 7 % of the targets when triplets were presented together, $t(37) = 7.39$, $p < .001$, $d = 2.42$. The associative strength between a cue and the correct target was greater in the apart condition than in the together condition, $t(37) = 5.99$, $p < .001$, $d = 1.97$. However, the associative strength between a cue and related (but wrong) targets was greater in the apart condition than in the together condition, $t(37) = 8.30$, $p < .001$, $d = 2.73$. Cues generated in the together condition were associated to fewer items in the database than those generated in the apart condition, $t(37) = 4.05$, $p < .001$, $d = 1.33$. Furthermore, in the apart condition, subjects repeated cues more frequently (7 %) than in the together condition (1 %), $t(37) = 5.80$, $p < .001$, $d = 1.90$. Finally, cues were more unique in the together condition than in the apart condition, $t(59) = 10.66$, $p < .001$, $d = 1.39$.

Effective cue characteristics

As in Experiments 1 and 2, we conducted an additional HLM analysis to ascertain what characteristics of cues across Experiments 4 and 5 effectively supported retrieval. These analyses are presented in more depth in Appendix D and show that cue-to-target associative strength and cue distinctiveness both support memory performance even when cue generation condition is considered. This hints that cue generation condition affects the characteristics of the generated cues and consequently the characteristics of the generated cues affect later recall.

Discussion

When information about competitors was available, learners generated cues to prevent confusions among targets and improve cued recall performance. Awareness of competitors led learners to decrease the cue-to-target associative strength, increase the distinctiveness of their cues, and improve performance. As in Experiment 4, cues in the together condition had smaller cue-to-target associative strength than cues in the apart condition, were less likely to be competing targets, and were less strongly associated to the competing targets. Differences in cues led learners in the together condition to correctly recall more items and have fewer confusions than those in the apart condition. Learners were better able to tailor their cues to distinguish among competitors when knowledge of the competitors was available. Just like the case of communicating with others, learners consider the context of the to-be-remembered items when making messages for their future self.

The between-subjects manipulation in Experiment 5 revealed a significant difference in cued recall performance, while the within-subjects manipulation in Experiment 4 did not. The subjects in the within-subjects apart condition were likely aware that there were relationships among separated targets and altered their cue generation strategies to try to distinguish among the competing targets. However, even though learners likely knew there would be related competitors across the within-subjects study list, they could not exactly predict what those related words might be. Not knowing what the competitors would be prevents these learners from generating cues that are as beneficial, as they can in the together condition. The differences between the cue characteristics from the together and apart conditions were greatly exaggerated in Experiment 5 when compared to Experiment 4. The average effect size of the differences in cue characteristics was over twice as big in Experiment 5 ($M_{effect\ size}$ = 1.85) compared to the average in Experiment 4 ($M_{effect\ size}$ = 0.74). The differences in effect sizes seems to be largely driven by differences in the apart conditions across experiments, which suggests that subjects in the between-subjects apart condition had laxer criteria for these items than subjects in the within-subjects apart condition.

When targets were presented with competing targets, learners generated cues with smaller cue-to-target associative strength, across both Experiments 4 and 5. In order to prevent confusions among the related targets, learners reduced how strongly a cue was associated to the target and its related competitors. Learners, then, are aware of the inherent benefits and limitations of cue-to-target associative strength and modulate it to effectively support future retrieval.

Furthermore, presenting the related items together allowed learners to generate cues that focused on the differences among targets in the context of similarity (Hunt, 2006). In both experiments, simultaneous presentations resulted in fewer confusions among the targets. Learners generated cues that specifically directed later access to a particular target to the exclusion of the incorrect targets. The better recall of items following combined processing of similarity and differences (as when the triplets are presented simultaneously) not only echoes extant results (Begg, 1978; Epstein, Phillips, & Johnson, 1975) but further suggests that learners can and naturally do generate distinctive cues that focus on unusual features to effectively constrain search.

General discussion

Across the set of experiments, learners used sophisticated tactics to generate cues that supported consistently high levels of cued recall performance. Cue-to-target associative strength, cue distinctiveness, and the match between encoding and retrieval played significant roles in fostering future retrieval. Learners generated cues that were compatible with their own rich, idiosyncratic knowledge, and they thrived when they received their own cues during the test. Regularities emerged in the types of cues learners utilized. When intentionally generating cues (instead of descriptions), learners increased the distinctiveness of their generated cues. Similarly, when learners were aware of related competitors in the to-be-remembered list, they increased the distinctiveness of the cues at the expense of the cue-to-target associative strength. Increasing the distinctiveness of the cues both enhanced recall performance and reduced confusions among competing targets under these task demands. In other situations, however (Tullis & Benjamin, in press), learners can improve cued recall by increasing their cues' cue-to-target associative strength at the expense of the cues' distinctiveness. Learners do not seek to maximize either cue-to-target associative strength or cue distinctiveness (or both) when generating cues. Rather, learners flexibly modulate the characteristics of generated cues to fit the particular demands of the task, and by doing so, bolster their memory performance.

Metacognitive control over cue characteristics

These results add to the growing literature that suggests that learners expertly utilize strategic metacognitive control beyond control of encoding to improve their learning. A growing body of literature is beginning to show how learners effectively exercise metacognitive control not just during encoding but also during retrieval (Finley & Benjamin, 2012; Halamish, Goldsmith, & Jacoby, 2012; Tullis et al., 2015; Tullis & Benjamin, in press). Variability in strategic control during retrieval may have just as large of an impact as variability in control over the circumstances of learning.

According to the memory-as-skilled-cognition perspective, variance in the use of memory arises from differences in the exercise of strategic control. If learners routinely create good retrieval cues, they are likely to exhibit great memory, especially outside of the laboratory where artificial constraints are removed. The ability to choose mnemonic cues may be an example of the "efficient action of higher-level decision making on the inputs and outputs from memory" that determines the quality of our memory use (Benjamin, 2008, p 177).

Learners can choose and utilize mnemonic cues effectively because they can base study choices upon their idiosyncratic cognitive state and metacognitive monitoring of learning (Jameson, Nelson, Leonisio, & Narens, 1993). Learners have privileged access to their idiosyncratic mental states which allows them to make more effective choices for themselves than could be determined by an outsider or by aggregate data. A learner's personal choices about cues are beneficial because they rely upon a learner's rich and idiosyncratic mental environment. Learners may generate cues that are idiosyncratically strongly associated to the target but would be meaningless for other learners. Learners can connect a memory goal with their rich, personal knowledge to increase the likelihood of accessing the target, decrease the costs of retrieval, and enable high levels of memory performance. Learners' choices about effective cues are guided by their idiosyncratic knowledge, their beliefs about how memory works, and their conceptions about how their internal and external environment will change in the future. Given that generating a supportive mnemonic cue requires the combination of those different types of knowledge, it is amazing how successful learners are at offloading their cognition to external mnemonic cues. Creating effective retrieval cues remains another example of how learners successfully use metacognitive control and privileged access to their mental states to enhance future retrieval.

Appendix A

Strategies of learners

We analyzed the types of cues learners generated across experiments to get an idea of what strategies learners were using. After an initial overview of the types of cues utilized across all experiments, six distinct categories of strategies were created. The categories included one-word cues, two-word cues (although learners were told to generate one-word cues and were not allowed to use the space bar, e.g., "femalecologne" as a cue for "perfume"), rearrangement of the letters in the target (e.g., "honep" as a cue for "phone"), the beginning of the target (but not the entire word, e.g., "pos" as a cue for "positive"), and adding a letter to the end of the target (e.g., "shoes" as a cue for "shoe"). Cues were coded blind to the condition in which they were generated. As shown in Table 7, learners mostly generated one-word cues across all experiments. All statistical analyses conducted throughout the paper were conducted on all of the cues generated, regardless of strategy, in order to avoid possible item effects. However, we also conducted analyses on only the one-word cues, and none of the statistical results changed.

Two interesting patterns of data arose. First, in Experiment 1, learners in the description condition used almost exclusively one-word cues, while learners in the cue generation condition relied less upon one-word cues. Learners were instructed to generate a one-word cue, so subjects may resort to generating two-word cues when they are relatively more concerned about recalling the target in the future. They shift away from the default one-word strategy more often when generating mnemonic cues than when generating descriptions. Second, in Experiment 5, learners relied upon two-word cues more when generating cues for the competing targets in the together condition than in the apart condition. Learners may shift cue generation strategies when recognizing the inherent difficulty of distinguishing among three related concepts. Learners may shift from the default one-word cue instructions when most concerned about remembering the target or distinguishing among competing targets.

Appendix B

The log odds of a correct recall on each trial were predicted by the different stimulus characteristics using a multilevel logit model. The model included the fixed effects of instruction condition and the simple interactions of cue originator condition with each of the following variables: cue-to-target associative strength, number associated from cue, total cumulative strength from cue, and cue in database. These were crossed with random intercepts for subjects and items. The model was fit in the R software package (R Development Core Team, 2008) with Laplace estimation using the *lmer()* function of the *lme4* package (Bates, Maechler, & Dai, 2008). Using backward elimination, interactions and variables that were least influential in the model were removed until removing predictors produced a model that fit the data significantly worse. The parameters of the least complex best fitting model are displayed below. The beta weights (and corresponding *Z* values) are shown in Table 8.

This model of performance shows that cue originator has a very strong effect on recall. The odds of correctly responding to their own were 13.2 times greater than correctly responding to another's cues. Additionally, the instructions given to the cue originator make a significant difference in the probability of recall. The odds of correctly recalling a target in response to a mnemonic cue were 1.67 times the odds of correctly recalling a target in response to a description. Furthermore, the stronger the cue-to-target associate strength, the more likely the target is recalled. The cue originator interacts with the cumulative total strength from the cue, which suggests that when the cue was created by another learner, total cumulative strength is less important than when it was created by oneself. When it was created by oneself, lesser cumulative associative strength promotes higher memory performance.

Table 7 Proportion of strategies utilized across Experiments 1–5. Comparisons of interest are shaded in the same color

	One-word cues	Two-word cues	Rearranged letters from targets	Beginning of target	Target with extra letter
Exp. 1 (Cue Gen)	0.90	0.05	0.04	0	0
Exp. 1 (Description)	0.99	0.01	0	0	0
Exp. 2	0.93	0.02	0	0	0.05
Exp. 3 (Cue Gen)	0.92	0.04	0	0.03	0
Exp. 3 (Description)	0.97	0.02	0	0	0
Exp. 4 (Together)	0.83	0.03	0.03	0.11	0
Exp. 4 (Apart)	0.85	0.03	0.03	0.10	0
Exp. 5 (Together)	0.89	0.09	0.01	0.01	0
Exp. 5 (Apart)	0.98	0.01	0	0	0

Table 8 Beta weights from the logit model

	β	*Z* value
Intercept	-0.16	0.70
Cue originator	2.58	15.06
Instructions	-0.53	4.45
Cue-to-target associative strength	4.59	9.36
Total strength	-0.39	2.10
Cue originator * total strength	-1.12	4.39

Table 9 Beta weights derived from the logit model

	β	*Z* value
Intercept	-0.34	1.61
Cue originator	3.08	10.78
Cue-to-target associative strength	16.14	6.71
Cue in database	-1.79	2.64
Total strength	1.54	1.85, $p = 0.07$
Cue originator * total strength	-1.06	2.72

The HLM corroborates the major results from the traditional analyses: cue generation instructions influence recall, self-generated cues show a higher likelihood of recall than other-generated cues, greater cue-to-target associative strength enhances likelihood of recall regardless of cue originator, and the influence of cue distinctiveness depends upon cue originator. A major difference between the HLM and the standard prior analysis is the influence of whether a cue is in the database and the raw number of potential targets associated to the cues (two measures of distinctiveness). The HLM shows that the number of targets associated to the cues did not contribute to the model's ability to predict memory performance. These two variables likely overlap largely with the other measures of distinctiveness (cumulative total strength); therefore, they did not contribute new information to the model and do not appear in it.

Appendix C

The log odds of a correct recall on each trial were predicted by the different stimulus characteristics using a multilevel logit model in Experiment 2. The model included the fixed effects of simple interactions of cue originator condition with each of the following variables: cue-to-target associative strength, number associated from cue, total cumulative strength from cue, and cue in database. These were crossed with random intercepts for subjects and items. The model was fit in the R software package (R Development Core Team, 2008) with Laplace estimation using the *lmer()* function of the *lme4* package (Bates, Maechler, & Dai, 2008). Using backward elimination, interactions and variables that were least influential in the model were removed until removing predictors produced a model that fit the data significantly worse. The parameters of the least complex best fitting model are displayed below. The beta weights (and corresponding *Z* values) are shown in Table 9.

As the prior hierarchical model showed, this model of performance shows that cue originator has a very strong effect on recall. If a learner receives their own cues, the odds of correctly recalling the target are 21.76 times the odds if they receive another's cues. Further, the stronger the cue-to-target associate strength, the more likely the target is recalled. Measures of distinctiveness have a more complex relationship to recall. When the cue is located in the database, learners are less likely to recall the target. Further, the cue originator interacts with the cumulative total strength from the cue. When the cue was created by another learner, total cumulative strength is less important than when it was created by oneself. When it was created by oneself, lesser cumulative associative strength promotes higher memory performance.

As in the prior hierarchical model, learners are more likely to recall the targets when they generated the cue than when someone else did. Furthermore, increased cue-to-target strength is beneficial to recall. As in the prior model, the effects of distinctiveness interact with cue originator. Finally, the number associated from the cue has no significant impact on probability of recalling the target. Once again, the information contained in this variable likely largely overlap with the other measures of distinctiveness. The one difference in this model contrasted with the prior one, is that whether the cue is in the database or not impacts the probability of recall significantly.

Appendix D

The log odds of a correct recall on each trial were predicted by the different stimulus characteristics using a multilevel logit model in Experiments 4 and 5. The model included the fixed effects of simple interactions of together condition with each of the following variables: cue-to-target associative strength, number associated from cue, total cumulative strength from cue, cue in database, whether the cue was a competitor, and the total associative strength from the cue to the competitors. These were crossed with random intercepts for subjects, items, and related trios. The model was fit in the R software package (R Development Core Team, 2008) with Laplace estimation using the *lmer()* function of the *lme4* package (Bates, Maechler, & Dai, 2008). Using backward elimination, interactions and variables that were least influential in the model

Table 10 Beta weights derived from the logit model

	β	*Z* value
Intercept	1.63	9.87
Cue-to-target associative strength	3.08	8.72
Total strength	-1.06	8.98
Cue is a competing target?	-0.55	4.49

were removed until removing predictors produced a model that fit the data significantly worse. The parameters of the least-complex best fitting model are displayed below. The beta-weights (and corresponding *Z* values) are shown in Table 10.

This model shows that the cue-to-target associative strength is strongly associated with cued recall, while the total cumulative associative strength is negatively associated with cued recall. When the cue is a competing target, the odds of incorrectly recalling the target are 1.73 times the odds if the cue is not a competitor. Noticeably absent from this list of significant predictors is the study condition (together vs. apart), which had a final beta weight of 0.01 (*Z* value of 0.094). The model suggests, then, that the study condition affects the kinds of cues generated, and the kinds of cues generated affect final cued recall.

Appendix E

Table 11 shows the Pearson intercorrelation matrix between the measured cue characteristics from Experiments 1, 2, and 3. The matrix reveals that all measures are positively correlated, and the three measures of distinctiveness are, in particular, very highly correlated with each other ($r > .75$).

Table 11 Intercorrelation matrix of cue characteristics from Experiments 1, 2, and 3

	Cue-to-target associative strength	Number associated to cue	Cumulative strength from cue	Cue in database
Cue-to-target associative strength	1			
Number associated to cue	0.11	1		
Cumulative strength from cue	0.28	0.80	1	
Cue in database	0.27	0.85	0.98	1

References

Allen, M. (1968). Rehearsal strategies and response cueing as determinants of organization in free recall. *Journal of Verbal Learning and Verbal Behavior, 7,* 58–63.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451–474.

Andersson, J., & Ronnberg, J. (1997). Cued memory collaboration: Effects of friendship and type of retrieval cue. *European Journal of Cognitive Psychology, 9,* 273–287.

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,* 390–412.

Bäckman, L., & Mäntylä, T. (1988). Effectiveness of self-generated cues in younger and older adults: The role of retention interval. *International Journal of Aging and Human Development, 26,* 241–248.

Bäckman, L., Mäntylä, T., & Erngrund, K. (1984). Optimal recall in early and late adulthood. *Scandinavian Journal of Psychology, 25,* 306–314.

Bates, D., Maechler, M., & Dai, B. (2008). *lme4: Linear mixed-effects models using S4 classes (Version 0.999375-20) [Computer software]*. Available at http://cran.r-project.org/.

Begg, I. (1978). Similarity and contrast in memory for relations. *Memory and Cognition, 6*, 509–517.

Benjamin, A. S. (2008). Memory is more than just remembering: Strategic control of encoding, accessing memory, and making decisions. In A. S. Benjamin & B. H. Ross (Eds.), *The psychology of learning and motivation: Skill and strategy in memory use* (Vol. 48, pp. 175–223). London, England: Academic Press.

Benjamin, A. S., & Bird, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language, 55,* 126–137.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Caranasos, G. J., Stewart, R. B., & Cluff, L. E. (1974). Drug-induced illness leading to hospitalization. *Journal of the American Medical Association, 228,* 713–717.

Einstein, G. O., Holland, L. J., McDaniel, M. A., & Guynn, M. J. (1992). Age-related deficits in prospective memory: The influence of task complexity. *Psychology and Aging, 7,* 471–478.

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 717–726.

Epstein, M. L., Phillips, W. D., & Johnson, S. J. (1975). Recall of related and unrelated word pairs as a function of processing level. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 149–152.

Feldman, S. M., & Underwood, B. J. (1957). Stimulus recall following paired-associate learning. *Journal of Experimental Psychology, 53,* 11–15.

Fiechter, J. L., Benjamin, A. S., & Unsworth, N. (2015). The metacognitive foundations of effective remembering. In J. Dunlosky & S. K. Tauber (Eds.), *Handbook of metamemory* (pp. 62–77). Oxford: Oxford University Press.

Finley, J. R., & Benjamin, A. S. (2012). Adaptive changes in encoding strategy with experience: Evidence from the test expectancy paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 632–652.

Finley, J. R., & Benjamin, A. S. (2015b). *Retrieval cue variability: When and why are two meanings better than one?* Manuscript under review.

Finley, J. R., & Benjamin, A. S. (2015a). *Metacognitive control of transfer-appropriate processing.* Manuscript in preparation.

Finley, J. R., Tullis, J. G., & Benjamin, A. S. (2009). Metacognitive control of learning and remembering. In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Cognition, computers and collaboration in education*. New York: Springer Science & Business Media.

Fisher, R. P., & Craik, F. I. (1977). Interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 3,* 701–711.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness, 5,* 189–211.

Goldsmith, M., & Koriat, A. (2007). The strategic regulation of memory accuracy and informativeness. *Psychology of Learning and Motivation, 48*, 1–60.

Halamish, V., Goldsmith, M., & Jacoby, L. L. (2012). Source constrained recall: Strategic control of production quality. *Journal of Experimental Psychology: Human Learning and Memory, 38,* 1–15.

Harris, J. E. (1980). Memory aids people use: Two interview studies. *Memory & Cognition, 8,* 31–38.

Haynes, R. B., McKibbon, K. A., & Kanani, R. (1996). Systematic review of randomized trials of interventions to assist patients to follow prescriptions for mediations. *The Lancet, 348,* 383–386.

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3–25). New York, NY: Oxford University Press.

Hunt, R. R. (2012). Distinctive processing: The co-action of similarity and difference in memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 56, pp. 1–46). Oxford, England: Elsevier.

Hunt, R. R., & Smith, R. E. (1996). Accessing the particular from the general: The power of distinctiveness in the context of organization. *Memory & Cognition, 24,* 217–225.

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ: Erlbaum.

Jameson, A., Nelson, T. O., Leonesio, R. J., & Narens, L. (1993). The feeling of another person's knowing. *Journal of Memory and Language, 32,* 320–335.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138,* 469–486.

Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1133–1145.

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32,* 609–622.

Lachowsky, M., & Levy-Toledano, R. (2002). Improving compliance in oral contraception: "The reminder card.". *The European Journal of Contraception and Reproductive Health Care, 7*(4), 210–215.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

Mäntylä, T. (1986). Optimizing cue effectiveness: Recall of 500 and 600 incidentally learned words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 66–71.

Mäntylä, T., & Nilsson, L. G. (1983). Are my cues better than your cues? Uniqueness and reconstruction as prerequisites for optimal recall of verbal materials. *Scandinavian Journal of Psychology, 24,* 303–313.

Mäntylä, T., & Nilsson, L. G. (1988). Cue distinctiveness and forgetting: Effectiveness of self-generated retrieval cues in delayed recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 502–509.

McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. (2004). Cue-focused and reflexive-associative processes in prospective memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 605–614.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10,* 389–395.

Nairne, J. S. (2006). Modeling distinctiveness: Implications for general memory theory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 27–46). Oxford, NY: Oxford University Press.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms.* Retrieved from http://w3.usf.edu/FreeAssociation/.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review, 77,* 257–273.

Park, D. C., Morrell, R. W., Frieske, D., & Kincaid, D. (1992). Medication adherence behaviors in older adults: Effects of external cognitive supports. *Psychology and Aging, 7,* 252.

Piette, J. D., Weinberger, M., Kraemer, F. B., & McPhee, S. J. (2001). Impact of automated calls with nurse follow-up on diabetes treatment outcomes in a department of Veterans Affairs health care system: A randomized controlled trial. *Diabetes Care, 24,* 202–208.

R Development Core Team (2008). *R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing*. Available at www.R-project.org.

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research, 32,* 3–23.

Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 601–604.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204–221.

Tullis, J. G., & Benjamin, A. S. (2011). On the effectiveness of self-paced learning. *Journal of Memory and Language, 64,* 109–118.

Tullis, J. G., & Benjamin, A. S. (2012). Consequences of restudy choices in younger and older learners. *Psychonomic Bulletin & Review, 19,* 743–749.

Tullis, J. G., & Benjamin, A. S. (in press). Generating mnemonic cues for others. *Memory & Cognition*.

Tullis, J. G., Benjamin, A. S., & Fiechter, J. (2015*). The efficacy of learners' testing choices.* Manuscript under review.

Tullis, J. G., Benjamin, A. S., & Liu, X. (2014a). Self-pacing study of faces of different races: Metacognitive control over study does not eliminate the cross-race recognition effect. *Memory & Cognition, 42*(6), 863–875.

Tullis, J. G., Benjamin, A. S., & Ross, B. H. (2014b). The reminding effect: Presentation of associates enhances memory for related words in a list. *Journal of Experimental Psychology: General, 143*(4), 1526–1540.

Tullis, J. G., Braverman, M., Ross, B. H., & Benjamin, A. S. (2014c). Remindings influence the interpretation of ambiguous stimuli. *Psychonomic Bulletin & Review, 21*(1), 107–113. doi:10.3758/s13423-013-0476-2.

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology, 77,* 593–601.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Waddill, P. J., & McDaniel, M. A. (1998). Distinctiveness effects in recall: Differential processing or privileged retrieval? *Memory & Cognition, 26,* 108–120.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 442–452.

Yoon, S. O., & Brown-Schmidt, S. (2013). Lexical differentiation in language production and comprehension. *Journal of Memory and Language, 69,* 397–416.