The costs and benefits of providing feedback during learning

MATTHEW JENSEN HAYS
University of California, Los Angeles, California

NATE KORNELL
Williams College, Williamstown, Massachusetts

AND

ROBERT A. BJORK
University of California, Los Angeles, California

It seems uncontroversial that providing feedback after a test, in the form of the correct answer, enhances learning. In real-world educational situations, however, the time available for learning is often constrained—and feedback takes time. We report an experiment in which total time for learning was fixed, thereby creating a trade-off between spending time receiving feedback and spending time on other learning activities. Our results suggest that providing feedback is not universally beneficial. Indeed, under some circumstances, taking time to provide feedback can have a negative net effect on learning. We also found that learners appear to have some insight about the costs of feedback; when they were allowed to control feedback, they often skipped unnecessary feedback in favor of additional retrieval attempts, and they benefited from doing so. These results underscore the importance of considering the costs and benefits of interventions designed to enhance learning.

One of the goals of education—perhaps the most fundamental goal—is maximizing long-term retention and transfer. Toward achieving that goal, one key manipulation is to induce the retrieval of to-be-learned information, which has been shown in many studies to enhance later recall more than does providing additional study trials (e.g., Roediger & Karpicke, 2006). Thus, in the domain of language learning, for example, students' frequent use of flash cards to study vocabulary items (Kornell, 2009; Kornell & Bjork, 2007) is wise because it induces retrieval (e.g., retrieving the English word *soul* when cued by the Swahili word *roho*). From a principles-of-learning standpoint, flash cards would seem to be a particularly effective study tool for another reason as well: They provide feedback in a natural way. After trying, successfully or unsuccessfully, to retrieve *soul*, for example, a student needs only to flip the card over to see *roho–soul*.

Intuitively, feedback should be universally beneficial. However, several studies have identified conditions in which feedback provides no benefit. For example, Pashler, Cepeda, Wixted, and Rohrer (2005) taught their participants Luganda–English word pairs by presenting the pairs once and testing them twice. For example, their participants first saw *leero–today* and later had two separate opportunities to retrieve *today* when provided with *leero–___*. Feedback (e.g., *leero–today*) was provided for some items but not for others. Providing immediate feedback yielded no benefit to long-term retention when the retrieval attempt had been successful. Karpicke and Roediger (2007, 2008) reported similar results. Their participants derived no benefit from re-reading information that they had successfully retrieved (although re-*retrieving* information doubled its later recall). This pattern of results is neither new (Guthrie, 1971) nor limited to particular instruction formats (see, e.g., Anderson, Kulhavy, & Andre, 1971).

There is no doubt that feedback can be beneficial. In the study mentioned above, Pashler et al. (2005) found that immediate feedback after an incorrect response improved final recall by 494%. Further, feedback after a correct response can be helpful—if the feedback is delayed (e.g., Kulhavy & Anderson, 1972; but see Metcalfe, Kornell, & Finn, 2009). To our knowledge, however, the only evidence that immediate feedback might be beneficial after a correct response comes from Butler, Karpicke, and Roediger (2008). When their participants provided correct responses with low confidence ratings, immediate feedback substantially improved their recall. However, Butler et al.'s participants were responding to multiple-choice questions. As a result, some of their low-confidence correct responses may have reflected some guessing (i.e., after participants' having eliminated one or two answer choices), in which case it would not be surprising if feedback were to be beneficial. Thus, Butler et al.'s results do not necessarily challenge the idea that feedback has few,

M. J. Hays, matt@hayslab.com

if any, benefits for materials that are already accessible in memory.

Prior studies in which feedback has yielded no benefit were designed so that feedback also incurred no cost. What if, though, total time on task had been controlled in these studies, as is usually the case in educational contexts? Students make decisions such as "I'll work on my Spanish for an hour" or "I'll study these flash cards for 20 minutes." When time is limited, feedback incurs a possible cost that may, in some circumstances, offset its benefits: It takes time away from activities that are potentially more productive, such as additional retrieval attempts (see, e.g., Karpicke & Roediger, 2008). In cases where providing feedback costs time but yields no benefit, doing so may have a net negative effect on learning.

The present study was conducted to examine the consequences of providing feedback when total time is fixed. The participants were presented with a set of word pairs and were then asked to retrieve them repeatedly. In one condition, feedback was provided after every retrieval attempt. In another condition, feedback was skipped after some retrieval attempts. Because selectively skipping feedback permitted more total retrieval attempts, we expected that such skipping would improve final-test recall.

We also examined whether our participants would manage selective feedback appropriately. We did so by manipulating whether the skipping of feedback was controlled by the computer or by the participants themselves. Given prior findings (e.g., those of Pashler et al., 2005), we programmed the computer to skip feedback after correct responses. Because learners are typically unaware that retrieval provides more benefit than does additional study (Karpicke, 2009; Kornell & Bjork, 2008; Kornell & Son, 2009), we expected our participants to skip feedback in nonoptimal ways, which, in turn should have yielded lower final-test recall.

METHOD

Participants

The participants were 68 undergraduate students at the University of California, Los Angeles. They received partial course credit as compensation for their participation.

Materials

The materials were a pool of 32 Swahili–English pairs (e.g., *roho–soul*) selected from norms collected by Nelson and Dunlosky (1994).

Procedure and Design

The 32 pairs were divided into two lists of 16 pairs each. The participants first learned one list, then the other. Learning began with a 5-sec presentation of each pair in the list. The order in which the pairs were presented was randomly generated for each participant.

After every pair in a list had been presented once, the first cycle of retrieval practice began. In each retrieval-practice cycle, the participants received one test trial for each pair in the list. A test trial consisted of a retrieval prompt for the English word (e.g., *roho–___*). Each test trial lasted 5 sec and was followed by a 2-sec pause. The order of test trials was determined randomly for each cycle for each participant.

After a test trial for a given pair, feedback—if it provided— consisted of a presentation of the intact pair (e.g., *roho–soul*). Each feedback trial lasted 5 sec and was followed by a 2-sec pause.

Whether the participants received a feedback trial after a test trial depended on the condition to which they and the list they were learning had been assigned. The experiment had two independent variables. The first independent variable was *skipping*, which determined whether skipping feedback was permitted or forbidden. If skipping was *forbidden*, a feedback trial followed each test trial. If skipping was *permitted*, the feedback trial after a test trial could be provided or skipped. Skipping was manipulated within subjects and between lists; if skipping was permitted on the first list, it was forbidden on the second, and vice versa. Random assignment determined whether the participants first encountered the skipping-permitted or skipping-forbidden list.

The second independent variable, nested within the skipping-permitted condition, was *controller*—that is, whether skipping was controlled by the participant or the computer. When the *participants* controlled skipping, they could choose, during the 2-sec pause after each test trial, to skip the feedback. If they wanted to skip, they pressed the space bar, which moved them directly to the next item in the list. If they did not press the space bar, they encountered a feedback trial for the pair that had just been tested.

When the *computer* controlled skipping, the participants were asked, during the 2-sec pause at the end of each test trial, whether they thought they had just provided the correct response. If they thought they had answered correctly, they pressed the space bar. Whether they did or did not receive feedback was, however, controlled by whether their response was actually incorrect or correct, respectively. (Correctness of participants' responses was judged via a strict, but not verbatim, letter-match algorithm.) Thus, the frequency of space-bar presses was similar in the two controller conditions, but only in the participant-controlled condition did pressing the space bar determine whether feedback would be skipped.

On each list, the total number of trials was held constant. On the skipping-forbidden list, each pair received four test trials and four feedback trials. In other words, there were four cycles through the skipping-forbidden list. On the skipping-permitted list, skipping a feedback trial allowed time for an additional trial, which constituted the beginning of a fifth cycle through the list. The pair tested in this additional trial was randomly determined (i.e., skipping *roho–soul* would probably not yield another test for *roho*). Skipping more feedback trials allowed time for more trials to populate the fifth cycle— and beyond. Feedback trials could be skipped on these later cycles as well. Indeed, had every feedback trial been skipped, there would have been eight test trials for each pair. When the participants controlled skipping, they were told: "You have a fixed amount of time to study each list, so by skipping feedback you will leave yourself more time for future study trials." Including the initial study phase, learning a list took just over 16 min.

After learning the second of the two lists (one in each skipping condition), the participants were dismissed. After a delay of 24 h, the participants were contacted via e-mail with a login to a Web site. This Web site administered an untimed final test on all the pairs from both lists (e.g., *roho–___*), which the participants completed after 24–48 h.

RESULTS AND DISCUSSION

We begin by analyzing the overall effect of skipping on recall. We then explore the participants' skipping behavior and its effects on memory.

Skipping Feedback and Its Effects on Memory

When skipping was forbidden, each item received four test trials. When skipping was permitted, each item received an average of 5.88 test trials ($SD = 0.90$). This difference was reliable [$t(67) = 17.17, p < .001$]. Thus, as was our intention, permitting skipping allowed a greater number of retrieval attempts.

Table 1
Proportion of Tests on Which the Participants Provided the Correct Answer and/or Pressed the Space Bar

Test	Skipping Forbidden			Skipping Permitted (Human)			Skipping Permitted (Computer)		
	Correct	Space Bar	Observations	Correct	Space Bar	Observations	Correct	Space Bar	Observations
Cycle 1	.27		2,272	.33	.27	1,072	.27	.35	1,200
Cycle 2	.49		2,272	.50	.38	1,072	.42	.50	1,200
Cycle 3	.65		2,272	.66	.61	1,072	.63	.70	1,200
Cycle 4	.76		2,272	.80	.75	1,072	.75	.79	1,200
Cycle 5			0	.86	.78	508	.90	.92	812
Cycle 6			0	.89	.93	255	.99	.99	563
Cycle 7			0	.92	.96	92	1.00	.98	270
Cycle 8			0	–	–	0	1.00	1.00	35
Final	.63			.74			.71		

Note—The proportion of feedback trials skipped on each cycle can be computed from the data in this table. When the computer controlled skipping, feedback was skipped whenever a correct response was provided. When the participant controlled skipping, feedback was skipped whenever the participant pressed the space bar.

Table 1 presents the participants' recall on the final test as a function of permission and controller. Final recall was reliably greater when skipping was permitted (combined across controller conditions: $M = .72$, $SD = .25$) than when it was forbidden ($M = .63$, $SD = .28$) [$t(67) = 3.88$, $p < .001$]. Thus, as we hypothesized, when the time to learn is constrained, as in typical real-world learning, providing feedback can be more a curse than a blessing.

How Did the Participants Decide to Skip Feedback?

On the skipping-permitted list, feedback could be skipped anywhere from zero to eight times per pair. Overall, the computer skipped feedback an average of 3.65 times per pair ($SD = 1.70$). The participants skipped feedback an average of 3.85 times per pair ($SD = 1.92$). This difference was not reliable [$t(66) < 1$, $p = .648$]. Thus, there was no overall difference in skipping frequency as a function of controller. (In the remainder of this section, our analyses of skipping behavior are based on only the first four cycles because of the drop-off in the number of observations in later cycles shown in Table 1.)

However, the participants' decisions about when to skip diverged from the computer's decisions in two ways not reflected in the overall skipping frequency data. First, the participants chose to skip feedback after an average of 15% ($SD = 16\%$) of their incorrect responses. The computer never skipped this feedback. This difference in behavior was reliable [$t(32) = 5.28$, $p < .001$]. Second, the participants chose to skip feedback after an average of 66% ($SD = 23\%$) of their correct responses. In other words, the participants *failed* to skip feedback after roughly one out of every three correct responses. The computer always skipped this feedback. This difference in behavior was reliable [$t(32) = 8.37$, $p < .001$]. Previous studies (e.g., Pashler et al., 2005) suggested that both types of deviation would result in better recall when skipping was computer controlled than when it was participant controlled.

One explanation for these divergences in behavior is that the participants were not accurately able to assess the correctness of their responses. That is, if the participants did not know when they were correct, they would have been unable follow the computer's skipping algorithm even if that had been their goal. The participants' ability to assess the accuracy of their responses can be derived from their space-bar-pressing behavior when the computer controlled skipping (i.e., when they pressed space bar if they thought they had provided the correct answer). The participants made two kinds of errors. First, they mistakenly assessed as correct 22% ($SD = 24\%$) of their incorrect responses. The computer never made this mistake. This difference in diagnostic accuracy was reliable [$t(34) = 5.38$, $p < .001$]. Second, the participants mistakenly assessed as incorrect 10% ($SD = 14\%$) of their correct responses, another mistake the computer did not make, and a difference that was also reliable [$t(34) = 4.36$, $p < .001$]. Taken together, these results indicate that the participants diverged from the computer substantially more often than their assessments were incorrect. Thus, errors in assessment cannot have been fully responsible for the participants' decisions about skipping that we believed were nonoptimal.

In the analyses immediately above, much of the difference between assessment errors and skipping-behavior divergence was driven by the participants' treatment of correct responses. On the computer-controlled list, the participants mistakenly assessed as incorrect 10% of correct responses. On the participant-controlled list, the participants failed to skip feedback after 34% of correct responses. This difference was reliable [$t(66) = 5.18$, $p < .001$]. (The corresponding difference in the participants' behaviors after incorrect responses was not reliable: $p = .171$.) Thus, the participants sometimes chose to view feedback even when they believed they had provided the correct response. This interpretation is consistent with the data in Table 1, which show that the participants consistently skipped less often than they were correct. Only by the sixth cycle did the participants skip approximately as often as they were correct, and many participants never skipped feedback on enough trials to reach Cycle 6.

We explored the data further to determine why the participants requested feedback when they apparently knew they were correct. Across cycles, there was a substantial increase in the participants' willingness to skip feedback after correct responses. They skipped feedback after fewer than half of their correct responses on Cycle 1 ($M = .49$, $SD = .36$). They were reliably more likely to skip feedback after a correct response on Cycle 2 ($M = .57$, $SD = .35$) [$t(32) = 2.08$, $p = .045$]. This value again reliably

increased from Cycle 2 to Cycle 3 ($M = .80$, $SD = .26$) [$t(32) = 3.89$, $p < .001$]. The increase from Cycle 3 to Cycle 4 ($M = .85$, $SD = .20$) was marginal [$t(32) = 1.85$, $p = .074$]. However, on the fourth cycle, the participants were skipping feedback after only 15% of correct responses—approximately as often as they mistook them for incorrect responses. That is, by the fourth cycle, the participants appeared to be attempting to control feedback in much the same way as was programmed into the computer.

How Did Control of Skipping Affect Final Recall?

As can be seen in Table 1, recall on the final test was slightly greater when skipping was controlled by the participants ($M = .74$, $SD = .25$) than when it was controlled by the computer ($M = .71$, $SD = .26$). This difference was not reliable [$t(66) < 1$, $p = .695$]. Thus, even though the participants' assessments were not perfect, and the participants diverged from the computer algorithm (overall) in their decisions about skipping feedback, their recall did not suffer. The conjecture that this divergence did cause a detriment, but that it was masked by an initial difference in the two groups' recall performance, is untenable given that the initial difference was not reliable [$t(66) = 1.37$, $p = .176$]. The most likely explanation appears to be that, by eventually adopting a skipping algorithm similar to that used by the computer, the participants offset the effects of their initial divergences in skipping behavior we deemed optimal.

GENERAL DISCUSSION

The main finding of this study is that feedback is not always beneficial—and that it can even be costly. When time was allocated to retrieval attempts instead of immediate feedback after correct responses, learning was enhanced. Two additional findings were notable. First, when participants were allowed to skip feedback, their decisions became broadly similar to those of the computer across cycles through the list. Second, skipping feedback was equally beneficial whether the decisions about when to skip were made by the participant or by the computer.

Explaining the Benefits of Skipping Feedback

The benefit to recall yielded by skipping feedback in the present study supports the broader notion that little is gained from being presented with information that is easily accessible in memory (Bjork & Allen, 1970; Pyc & Rawson, 2009). This notion is explicit, for example, in the new theory of disuse (Bjork & Bjork, 1992). In this framework, the magnitude of a learning increment is inversely proportional to the accessibility of the to-be-learned information. This framework predicts that feedback following a successful retrieval should have virtually no effect on learning. This prediction is borne out directly by previous studies (e.g., Pashler et al., 2005) and is consistent with the results presented here.

Further, this prediction is consistent with the two cases in which feedback after a correct response can be useful.

One such case is that of correct multiple-choice selections made with low confidence (Butler et al., 2008). In this case, low confidence in a correct selection may indicate that the information was not readily accessible in memory. The other case holds when feedback after a correct response is delayed. In this case, the spacing effect manifests as increased later recall of the retrieved information (Smith & Kimball, 2010). In both of these cases, feedback is beneficial because the to-be-retrieved information is not fully accessible when feedback is provided.

Deciding to Skip Feedback

When participants were allowed to skip feedback, they requested feedback on a significant number of items on which they were able to provide the correct response. There are at least two ways to interpret this pattern of results. One is that our participants were demonstrating metacognitive sophistication. That is, perhaps they requested feedback when they thought they were correct but wanted to verify their assessment. Given that the participants' assessments were imperfect, there must have been some trials on which the participants thought they were correct, requested feedback anyway, and found out that they had been wrong. Had the participants instead assumed they were correct and skipped feedback on such trials, they would have sacrificed the substantial benefit of corrective feedback, thereby dramatically reducing their chance to learn that item on that cycle (see, e.g., Kornell & Bjork, 2007; Kornell & Son, 2009). They might also have been inclined to skip feedback for that item on later cycles, which would have robbed them of subsequent learning opportunities.

In this light, the participants may have been demonstrating awareness of the fallibility of their assessments. By viewing feedback when they thought they were correct, they guarded themselves against overconfidence and likely improved their ability to assess their responses. This interpretation is supported by the finding that participants viewed feedback after correct trials progressively less frequently across retrieval-practice cycles. By the fourth cycle, they were doing so approximately as often as they mistook correct responses for incorrect responses. Interpreted this way, our results suggest that the participants had a remarkable understanding of their memories' strengths and weaknesses. Further, they were able to apply this understanding to effectively manage feedback during instruction.

On the other hand, it may be that the properties of our paradigm conspired to make our participants *appear* metacognitively sophisticated. That is, the participants may have tried to adopt the computer's algorithm, but for the wrong reasons. Kornell and Bjork's (2008) participants behaved in a way that supports this alternative interpretation. Like our participants, their participants were allowed to skip learning events in a fixed-study-time paradigm. Also like our participants, their participants chose to skip some items in order to make time for additional study of other items. In their case, however, the skipped events included tests as well as feedback. As a result, their participants learned *less* when they were allowed to skip than

when they were not (cf. Karpicke & Roediger, 2007), but opted to skip anyway. It may be that learners are willing to skip learning events that appear to be repetitive because they think of "learning" as acquiring new information. That approach would have been beneficial in the present paradigm because the consequence of skipping was positive.

Although this bleak interpretation is less appealing than the more charitable explanation above, it is consistent with other evidence that learners are often poor stewards of their own learning (e.g., Son, 2005). In the present experiment, the computer was programmed to skip feedback after correct responses in order to make way for more retrieval attempts. Learners, however, are not aware of the benefits of retrieval (Kornell & Son, 2009; Roediger & Karpicke, 2006), nor are they aware that a presentation after a successful retrieval confers minimal benefit to later recall (e.g., Karpicke, 2009). As a result, the parallels between the behaviors of the participants and the computer may have been motivated by decisions having little to do with metacognition or metacognitive sophistication.

These two interpretations are not mutually exclusive; perhaps elements of both are valid. The participants may have improved their ability to assess their responses—and reduced potential overconfidence—by viewing feedback on items they thought they knew. However, they may not have understood the powerful effects of retrieval, and so might have skipped more beneficial feedback in a different paradigm.

Concluding Comment

Across the history of laboratory research on learning and memory, there has been a tendency to ignore time on task. Comparisons between two or more conditions of instruction frequently equate the number of trials per item or the number of cycles through a given list without controlling total time. From an educational perspective, controlling time spent learning may be a more appropriate constraint. Restricted time is a reality in classrooms, where lessons last for a fixed duration. It is also a reality for students, who do not study for unlimited amounts of time. Busy schedules, extracurricular activities, and basic needs like food and sleep make time much more scarce in the dormitory than in the laboratory. It is important to consider not only the principles we want to examine, but also the questions that educators and learners want answered (Daniel & Poole, 2009). Recognizing limitations such as time on task, and comparing conditions of learning within those limitations, will lead to theoretical advances that can be coupled with practical advice.

REFERENCES

ANDERSON, R. C., KULHAVY, R. W., & ANDRE, T. (1971). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology*, **63**, 186-188. doi:10.1037/h0032653

BJORK, R. A., & ALLEN, T. W. (1970). The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning & Verbal Behavior*, **9**, 567-572. doi:10.1016/S0022-5371(70)80103-7

BJORK, R. A., & BJORK, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35-67). Hillsdale, NJ: Erlbaum.

BUTLER, A. C., KARPICKE, J. D., & ROEDIGER, H. L., III (2008). Correcting a metacognitive error: Feedback increases retention of low confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **34**, 918-928. doi:10.1037/0278-7393.34.4.918

DANIEL, D., & POOLE, D. (2009). Learning for life: An ecological approach to pedagogical research. *Perspectives on Psychological Science*, **4**, 91-96. doi:10.1111/j.1745-6924.2009.01095.x

GUTHRIE, J. T. (1971). Feedback and sentence learning. *Journal of Verbal Learning & Verbal Behavior*, **10**, 23-28. doi:10.1016/S0022-5371(71)80088-9

KARPICKE, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General*, **138**, 469-486. doi:10.1037/a0017341

KARPICKE, J. D., & ROEDIGER, H. L., III (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory & Language*, **57**, 151-162. doi:10.1016/j.jml.2006.09.004

KARPICKE, J. D., & ROEDIGER, H. L., III (2008). The critical importance of retrieval for learning. *Science*, **319**, 966-968. doi:10.1126/science.1152408

KORNELL, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, **23**, 1297-1317. doi:10.1002/acp.1537

KORNELL, N., & BJORK, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, **14**, 219-224.

KORNELL, N., & BJORK, R. A. (2008). Optimising self-regulated study: The benefits—and costs—of dropping flashcards. *Memory*, **16**, 125-136. doi:10.1080/09658210701763899

KORNELL, N., & SON, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, **17**, 493-501. doi:10.1080/09658210902832915

KULHAVY, R. W., & ANDERSON, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology*, **63**, 505-512. doi:10.1037/h0033243

METCALFE, J., KORNELL, N., & FINN, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition*, **37**, 1077-1087. doi:10.3758/mc.37.8.1077

NELSON, T. O., & DUNLOSKY, J. (1994). Norms of paired-associate recall during multitrial learning of Swahili–English translation equivalents. *Memory*, **2**, 325-335. doi:10.1080/09658219408258951

PASHLER, H., CEPEDA, N. J., WIXTED, J. T., & ROHRER, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 3-8. doi:10.1037/0278-7393.31.1.3

PYC, M. A., & RAWSON, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory & Language*, **60**, 437-447. doi:10.1016/j.jml.2009.01.004

ROEDIGER, H. L., III, & KARPICKE, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, **17**, 249-255. doi:10.1111/j.1467-9280.2006.01693.x

SMITH, T. A., & KIMBALL, D. R. (2010). Learning from feedback: Spacing and the delay–retention effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **36**, 80-95. doi:10.1037/a0017407

SON, L. K. (2005). Metacognitive control: Children's short-term versus long-term study strategies. *Journal of General Psychology*, **132**, 347-363. doi:10.3200/genp.132.4.347-364