How and when do students use flashcards?

Kathryn T. Wissman [a] , Katherine A. Rawson [a] & Mary A. Pyc [b]

[a] Department of Psychology, Kent State University, Kent, OH, USA

[b] Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA

PLEASE SCROLL DOWN FOR ARTICLE

How and when do students use flashcards?

Kathryn T. Wissman[1], Katherine A. Rawson[1], and Mary A. Pyc[2]

[1]Department of Psychology, Kent State University, Kent, OH, USA
[2]Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA

Previous survey research has documented students' use of self-regulated study strategies, with a particular interest in self-testing. These surveys indicate that students frequently use flashcards to self-test and that self-testing is primarily used as a way to monitor learning. Whereas previous surveys provide information about whether and why students self-test, they provide minimal information about how and when students choose to self-test. Accordingly, the primary purpose of the current survey was to explore how and when students engage in self-testing. We surveyed 374 undergraduates about the amount of practice and the timing of practice, two factors that strongly affect the efficacy of self-testing. Results indicate that students understand the benefits of practising to higher criterion levels (amount of practice) but do not typically implement or understand the benefits of practising with longer lags (timing of practice). We discuss practical implications for supporting more successful student learning.

Keywords: Self-testing; Criterion learning; Lag effects; Survey.

A wealth of previous research has shown that testing is beneficial for learning and memory (for recent reviews, see Rawson & Dunlosky, 2011; Roediger & Butler, 2011). However, the effectiveness of self-testing depends on several factors, including the amount and timing of practice. Although the majority of research on testing effects has been conducted in laboratory settings in which practice tests are under experimental control, in real-world contexts decisions of how and when to self-test are left up to students. Thus, the actual benefits of self-testing in most learning environments will depend on the extent to which students' choices reflect those conditions that facilitate learning. Given the sizeable amount of laboratory research exploring conditions under which self-testing is most effective, it is perhaps surprising that little is known about how students

implement self-testing on their own. Accordingly the main focus of the current survey was to explore how and when students self-test, with a particular interest in students' self-reported use of flashcards (given that students often report using flashcards during self-regulated learning).

HOW SHOULD STUDENTS IMPLEMENT SELF-TESTING?

Two critical factors that affect the efficacy of self-testing are the amount of practice and the timing of practice. Concerning the amount of practice, increasing the number of times an item is correctly recalled during practice benefits later retention (e.g., Karpicke, 2009; Karpicke & Roediger, 2008; Pyc & Rawson 2009; Vaughn &

Rawson, 2011). For example, Vaughn and Rawson (2011) presented learners with Lithuanian–English word pairs for test–restudy practice until items were correctly recalled one to five times. Correctly recalling an item more than once during encoding significantly improved final cued recall two days later (e.g., 31% versus 71% for items correctly recalled once versus four to five times). In addition to effects of the amount of practice within a given learning session, increasing the number of learning sessions benefits long-term retention. For example, Rawson and Dunlosky (2011) presented students with key concept definitions for test–restudy practice until items were correctly recalled during initial learning, followed by anywhere from one to five relearning sessions on subsequent days. Final cued recall 1 month later improved substantially as the number of relearning sessions increased (e.g., 35% versus 62% after one versus four to five relearning sessions).

Concerning the timing of practice, self-testing is most effective when using longer versus shorter lags between practice trials for a given item, both within and between sessions (e.g., Bahrick & Hall, 2005; Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Karpicke & Bauernschmidt, 2011; Kornell, 2009; Pashler, Zarow, & Triplett, 2003; Pyc & Dunlosky, 2010; Pyc & Rawson, 2007, 2009, 2012). Pyc and Rawson (2009) presented learners with Swahili–English word pairs for test–restudy practice trials that were separated by either 6 or 34 other items. Items practised with a longer versus shorter lag were more likely to be recalled on a final test 25 minutes later (76% versus 55%) or one week later (30% versus 5%). In sum, self-testing is most effective when implemented with higher criterion levels within and across sessions and longer lags between practice trials.

HOW DO STUDENTS IMPLEMENT SELF-TESTING?

In contrast to the sizeable amount of laboratory research in which practice tests are under experimenter control, minimal research has explored how students implement self-testing on their own. A few studies have used experimenter-controlled environments to examine students' decisions about the amount or timing of self-testing (Karpicke, 2009; Kornell & Bjork, 2008a; Kornell & Metcalfe, 2006). Kornell and Bjork (2008a,

Experiment 3) gave learners 10 minutes to learn 20 Swahili–English word pairs (in anticipation of a final cued recall test after a 5-minute delay). After initial study, items were presented for test–restudy trials. After each trial learners were given the option to remove a word pair or keep it in the list for further practice. Learners dropped 63% of word pairs after one correct recall rather than practising to a higher criterion. These results appear to suggest that students do not understand the benefits of higher criterion levels. However, a time limit of 10 minutes might have compelled students to drop items after only one correct recall to ensure enough time was left to practise unlearned items. More generally, all previous experimental studies have only given students control over limited aspects of self-testing in which much of the implementation of testing is still under experimenter control.

Survey methods provide a complementary approach to investigating students' use of self-testing (Carrier, 2003; Hartwig & Dunlosky, 2012; Karpicke, Butler, & Roediger, 2009; Kornell & Bjork, 2007; Kornell & Son, 2009; McCabe, 2010). Surveys avoid the disadvantage of limited control in experimental studies by providing students with opportunities to report decisions used in real-world contexts. Several recent survey studies have established that students use self-testing and frequently report using flashcards in particular. Hartwig and Dunlosky (2012) had students report which study strategies they use on a regular basis, and 62% of learners reported using flashcards. Likewise Karpicke et al. (2009) had students freely list strategies used during study; 40% of learners reported using flashcards. Prior surveys also provide information about why students self-test. Kornell and Son (2009) asked learners why they choose to self-test during study; 66% of learners reported self-testing to determine how well they knew the information, whereas only 20% said because they learn more than with restudy. Similarly, in surveys by Hartwig and Dunlosky (2012) and Kornell and Bjork (2007) 54% and 68% of learners reported self-testing to figure out how well they had learned information, and only 27% and 18% said because they learned more that way than rereading. These findings suggest that students primarily self-test to monitor how well target information has been learned rather than as a learning strategy.

Although prior survey research is informative about whether and why students choose to self-test, prior survey studies provide minimal

information about *how* and *when* students engage in self-testing. Given that the efficacy of self-testing is highly dependent on how and when it is used, the present survey focuses on the details of these factors in students' flashcard use.

METHOD

Surveys were administered via computer as part of six larger studies. Across all six studies, 374 undergraduates from Kent State University completed the survey in return for course credit. Each participant was given 20 minutes to complete the survey. Due to the time limit some participants did not respond to every question. Furthermore the particular set and sequence of questions varied somewhat across studies. Accordingly the number of respondents varies across questions; sample size for each question is indicated in the tables below.

For the majority of the survey questions, response format was open-ended. The few questions involving an alternate response format are denoted in the tables below. For scoring purposes, all open-ended responses were coded by one of two raters. Data were initially coded by frequency of response exemplars. The tally for each response exemplar included verbatim or close paraphrases of a particular idea. For example, Question 1A asked learners why they use flashcards to study for classes; responses of "easy for me to learn", "learning is easier this way", and "easiest way to learn" were all coded under "easy to learn". Two authors then reviewed the list of response exemplars and decided which exemplars could be collapsed into response categories. Response categories were developed for exemplars that indicated a similar kind of information (e.g., for Question 2C, we collapsed responses of "quickly recalled" and "easily recalled" into one response category) or to combine infrequent responses (e.g., for Question 2G, we created the category "6+ times"). Overall, response categories were used infrequently; all responses presented in the tables are exemplars unless otherwise denoted as a response category.

RESULTS AND DISCUSSION

Before reporting the primary outcomes of interest concerning how and when students self-test, we briefly discuss relevant background information

about respondents and replications of findings from prior survey studies.

Background information

Appendix A of the online supplement reports background information about the survey respondents. Most students were freshmen, and a majority of the sample maintained above a 3.00 grade point average. A significant number of students reported that teachers and/or professors taught them study strategies (A1–3). Interestingly, although a large majority of students reported having received instruction from teachers about how to study, students reported that only 23% of instructors recommended flashcards, and only 25% recommended completing practice problems or quizzes.

Whether and why students use flashcards

Replicating prior survey research, Table 1 shows that a large number of students report using flashcards to study (Question 1A). In contrast to previous findings, our results differ somewhat in regard to reasons students gave for why they use flashcards. Whereas the most common reason students reported self-testing in prior studies was to evaluate how well target information has been learned, the most common reason reported in the current study was to help memorise information (although see Question 2J for some evidence that students revisit dropped flashcards for purposes of checking memory). These divergent outcomes may be due to methodological differences between studies. For example, whereas we asked students why they use flashcards in particular, prior surveys asked students why they choose to quiz themselves more generally (which might or might not involve the use of flashcards). Additionally, the three prior survey studies all used a multiple choice format in which possible reasons were provided but that only permitted selection of one response option. In contrast we used an open-ended question. Given the open-ended nature of the responses it is unclear whether students meant the act of testing in and of itself helps memorisation or if testing helps determine what information still needs to be memorised. Further clarification on this distinction awaits future research.

TABLE 1
Whether and why students use flashcards

1A. Do you use flashcards to study for your classes? ($n = 247$)

Yes	67.6%
No	32.4

Why or why not? ($n = 154$)*

	Why ($n = 114$)	Why Not ($n = 40$)
Helps to memorise	60.5%	
Easy way to learn	22.8	
Lots of vocabulary	14.9	
Convenient to carry	8.8	
Isolates information	6.1	
Helps organisation	2.6	
Breaks up information	1.7	
Forces rereading	1.7	
Helps/forces rewriting	0.9	
Not helpful		60.0%
Not useful		35.0
Didn't need to study		5.0

1B. Explain what type of information you put on them ($n = 216$)*

Vocabulary	82.9%
Factual information from notes (people/places/dates)	29.2
Key concepts	28.2
Formulas	7.4
Difficult information	3.7
Information from study guide	3.7

1C. In reference to the classes you took last semester, how often did you use flashcards to study? ($n = 264$) +

0 classes	35.6%
1 class	11.4
2 classes	12.1
3 classes	22.3
4+ classes	18.6

*For these questions the sum of response percentages exceeds 100% because participants could report more than one response.
+For this question participants were given five corresponding response buttons.

Extending beyond prior research, Table 1 provides information about what kinds of materials students use with flashcards—an overwhelming 80% of students report using flashcards to learn vocabulary (Question 1B)—and also shows that a majority of students (53%) use flashcards in more than one class (Question 1C).

How students use flashcards: Amount of practice

Learning to criterion. Table 2 includes questions about how students use flashcards, focusing

TABLE 2
How student use flashcards: Amount of practice

2A. Imagine that you are studying with a stack of flashcards. How do you decide when you have studied a given flashcard enough? ($n = 327$)*

I recall it correctly	39.4%
I recall it correctly more than once	25.7
I don't have to look at the other side	17.4
I don't have to think about it	15.6
I recall it quickly	12.5
I understand the information	8.6
Mix-up and still recall	5.5
Other	3.1

2B. Imagine you are studying a stack of flashcards. Which of the following would influence how many times you studied an item? ($n = 272$)*+

How easily you can remember the item	80.0%
How many times you can correctly recall item	68.7
Importance of the particular item	58.9
How long you planned to study that day	40.1
Other	4.0

2C. Do you ever remove cards from the stack of flashcards when you are studying? ($n = 327$)

Yes	73.7%
No	26.3

If yes, please explain when and why you would drop out cards ($n = 296$)*

Need to focus on unknown	40.0%
Remembered multiple times	33.8
Don't need to study what is known	29.8
Quickly/easily recalled †	12.0
Ones I need to study more	8.4
Helps divide concepts	3.1
Helps organise	2.7
Other	1.8

If no, please explain why you would not drop out cards ($n = 72$)*

Need to know all the material	55.6%
Will remember better more times studied	36.1
Messes up the order	8.3
Don't see the point	8.3

2D. When you take a card out, how often is it because you tested yourself and were able to remember it correctly? ($n = 133$) ^

0.0–19.9%	8.3%
20.0–39.9%	3.8
40.0–59.9%	13.5
60.0–79.9%	21.8
80.0–100.0%	52.6

2E. When you take a card out, how often is it because it's too hard and you don't feel that you can learn it? ($n = 135$) ^

0.0–19.9%	65.0%
20.0–39.9%	16.3
40.0–59.9%	5.9
60.0–79.9%	8.9
80.0–100.0%	3.0

2F. When you take a card out, how often is it because it's too easy and you don't feel you need to practice it? ($n = 135$) ^

0.0–19.9%	20.7%

TABLE 2 (*Continued*)

20.0–39.9%	11.1
40.0–59.9%	8.1
60.0–79.9%	25.2
80.0–100.0%	34.8

2G. When using flashcards to self-test your memory, how many times do you think you should be able to correctly recall an item during flashcard practice so that you will be able to remember it later for an exam? (*n* = 243)

1 time	0.8%
2 times	4.5
3 times	22.2
4 times	8.2
5 times	22.0
6+ times †	29.6
Non-numeric response †	12.7

2H. Imagine that you are self-testing using flashcards and you correctly recall an item once. How much more likely would you be to remember that item later if you continued practising until you correctly recall that item again a second time? (*n* = 143) ^

0.0–19.9%	2.1%
20.0–39.9%	7.7
40.0–59.9%	9.1
60.0–79.9%	32.3
80.0–100.0%	49.0

2I. Do you think there is a point at which correctly recalling an item more times during practice would no longer benefit your ability to remember that item later? If so, what do you think that point is? (*n* = 143)

No, there is not a point	44.8%
Yes, there is a point	55.2

If so, what is the point? (*n* = 78)*

I recall it X number of times	52.6%
I know answer without thinking about it	30.8
I remember the item the next day	12.8
I start to become unfocused	11.5
Other	2.6

2J. If you were to drop out cards from your stack of flashcards when you are studying, once an item is dropped from practice, does it ever get put back in the stack of flashcards within that same session? (*n* = 315)

Yes	84.1%
No	15.9

Why or why not?*

	Why (*n* = 257)	Why Not (*n* = 37)
To check memory	41.0%	
Always useful to restudy	37.7	
Need to keep all to learn all	14.4	
Helps to understand completely	7.0	
Forget otherwise	7.0	
I know it		64.9%
Use it in different session		37.9
Need to focus on unknown		13.5
Confuses if re-added		8.1

TABLE 2 (*Continued*)

2K. Suppose you were using flashcards to prepare for an upcoming exam, and you wanted to get an Ā on the exam. On how many different days would you study with your stack of flashcards? (*n* = 126)

1 day	4.0%
2 days	13.5
3 days	15.1
4 days	13.5
5 days	15.1
6 days	4.8
7 days	14.3
8–10 days †	12.7
Other/non-numeric †	7.2

2L. When using flashcards to self-test, how do you decide if you have correctly recalled an item? (*n* = 274)*

Self check by looking at external information	48.9%
Self check in mind	27.4
Continuously recall correctly	16.4
Answer correctly without hesitation	12.0
Partner check	3.6
Write down correctly	2.5

*For these questions the sum of response percentages exceeds 100% because participants could report more than one response.

+For this question participants were provided with the list of options.

^For Questions 3D–3F participants were given were given a sliding scale with endpoints marked as ''0% of the time'' and ''100% of the time''. For Question 3H participants were given a sliding scale with endpoints marked as ''100% more likely'' to ''0% more likely''.

†Response category: see text for details.

specifically on students' self-reported use and beliefs about amount of practice. Question 2A demonstrates that a large majority of students (65%) report practising until items are correctly recalled (collapsing across ''I recall it correctly'' and ''I recall it correctly more than once''; these two response exemplars were treated as mutually exclusive for coding purposes). Although the number of correct recalls was not always specified, this finding suggests that many students learn to a criterion of at least one correct recall when using flashcards. With that said, a disconcerting number of students report making decisions without checking to make sure they are correct (''I don't have to look at the other side'', 17%) or based on cue familiarity (''I don't have to think about it'', 16%) and/or fluency of retrieval (''I recall it quickly'', 13%). Although fluency of retrieval can accurately predict subsequent

performance in some situations, fluency can be misleading in other cases (e.g., if the fluently retrieved information is incorrect, or if fluency is due to retrieval from working memory rather than long-term memory). Relying only on the fluency with which information comes to mind without evaluating the correctness of the fluently retrieved information may be particularly problematic if it leads students to incorrectly believe they know more than they do, which is particularly likely for more complex material (e.g., Dunlosky, Hartwig, Rawson, & Lipko, 2011; Rawson, O'Neil, & Dunlosky, 2011).

Nonetheless most students seem to understand the importance of correctly recalling target information during study, which is one key component of maximising the benefits of self-testing. Converging evidence for this conclusion comes from Question 2B, with a majority of students (69%) reporting that they practice to criterion when studying with flashcards. Additionally, Question 2D shows that students commonly report dropping flashcards in their stack because they reached criterion. Along the same lines Question 2E indicates that students are not likely to drop cards prior to reaching criterion. Students also reported removing items from practice when the item was easy (Question 2F), but these results are somewhat harder to interpret given that students might have interpreted "too easy" to mean they were able to recall target material correctly, quickly, or based on some other source (e.g., cue familiarity).

Higher criterion is better. Results thus far indicate that students typically choose to reach criterion before dropping flashcards from study. The next question of interest concerns the extent to which students practise beyond a single correct recall. In Question 2A, 26% of students spontaneously reported that they practise until information is correctly recalled more than once. Question 2C indicates that of the 74% of students who report dropping flashcards from their stack during study, 34% of these students explicitly stated that they practised recalling information multiple times before dropping it. However, these outcomes may underestimate the prevalence of practising to higher criterion levels, in that some students might simply not have thought to specify a particular criterion level. Indeed, when asked directly, almost every single student indicated belief that practising beyond one correct recall will help later memory of target material for an exam (Question 2G). With that said, by asking "how many times", the phras-

ing of Question 2G might have biased some students to report more than one correct recall versus only one (although students' preferences for more versus fewer correct recalls are harder to explain as due to biasing from the prompt).

Question 2H shows that a majority of students expect to benefit substantially from correctly recalling information for a second time. Question 2I suggests that some students do not anticipate diminishing returns from continuing to correctly recall target material (although a majority of students do correctly believe that once information has been recalled a certain number of times, there is no need to continue to practise it). Finally, Question 2J indicates almost all students (84%) report putting dropped flashcards back into their stack during a practice session. Moreover, when students report not putting a flashcard back into their stack, it is interesting to note that a common reason is because they intend to practise it at a later time (38%). Lastly, Question 2K indicates that students believe it is important to restudy flashcards on more than one day in order to obtain an A on an exam, suggesting that students understand the benefit of engaging in multiple learning sessions.

Concerning criterion learning and amount of practice, one related aspect is important to consider. Students report learning to criterion and that a higher criterion is beneficial, but how do students know when they have correctly recalled information? Question 2L shows that fewer than half of students report checking answers against external information (e.g., turning the flashcard over or checking against notes). Importantly, even when students check their answers against external information, they often still overestimate the correctness of their own answer (e.g., Dunlosky et al., 2011; Lipko et al., 2009; Rawson et al., 2011). Even more disconcerting is the fact that 27% of students reported checking answers "in their mind", given that research has shown that evaluating the correctness of a response without comparison to external information produces even higher levels of overconfidence (Dunlosky & Rawson, in press; Lipko et al., 2009). In either case, overconfidence can produce detrimental effects on learning because students do not effectively regulate practice (Dunlosky & Rawson, in press; Rawson et al., 2011). For present purposes this outcome suggests that students may not actually achieve their intended criterion level if they prematurely terminate practice due to overconfidence in the accuracy of their responses.

TABLE 3
How students use flashcards: Timing of practice

3A. Imagine that you have a total of 60 flashcards for a class. You need to know all 60 items for the next exam. When studying the flashcards, how many cards would you practise at a time? In other words, would you practise all 60 in one big stack, or would you divide them up into smaller stacks and then study one stack at a time? Below indicate how many flashcards you would practise at a time ($n = 279$)+

60	28.7%
30	9.0
20	21.0
15	14.0
10	18.3
Other	9.0

Why? ($n = 253$)*

	10–15 ($n = 87$)	20–30 ($n = 72$)	60 ($n = 74$)
Need to know them all	4.6%	11.1%	63.5%
Easier to remember a smaller stack	67.8	44.4	2.7
Less overwhelming	34.5	40.3	–
Divides into ones I know/don't know	–	15.3	20.3
Influences how information is encoded/learned	4.6	4.2	8.1
Needs to be somewhat difficult	1.1	6.9	6.8
Confusing to divide	1.1	1.4	2.7
Split into categories/chapters †	5.7	2.8	1.3
See card more times	4.6	4.2	–

3B. Do you think having a larger or small stack of flashcards is better for studying? ($n = 263$)*

Small	72.2%
Medium	1.9
Large	16.3
Doesn't matter	6.8
Depends	4.2

Why? ($n = 229$)*	Small ($n = 176$)	Medium/Large ($n = 42$)
Less to remember	53.7%	9.5%
Learn quicker	23.9	4.8
Narrows focus	19.1	11.9
See card more times	8.5	2.4
Gets confusing	8.0	–
Helps to feel accomplished	6.4	4.8
Brain only holds so much	6.4	–
Need to know all the information	3.2	64.3
Time between cards better	1.1	4.8

3C. If you were going to study a stack of flashcards, what do you think would work best and why? Study them one at a time repeatedly (for example, 1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4). Study them mixed together (for example, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4). ($n = 140$)*

Mixed	80.9
Repeated	17.1
Other	3.9

*For these questions the sum of response percentages exceeds 100% because participants could report more than one response.
+For this question participants were given six corresponding response buttons.
†Response category: see text for details.

When students use flashcards: Timing of practice

Table 3 includes questions about when students use flashcards with specific respect to the timing of practice. Most strikingly, several outcomes indicate that students do not understand the benefits of longer lags during study. For example, Question 3A indicates that a majority of students prefer to use a smaller versus larger stack of flashcards when studying. Of those students who reported a preference for using a smaller stack, a

substantial percentage (68%) said the reason was because it is easier. In addition, even when students said they would use a larger stack, none of them reported doing so because it would improve learning. Question 3B provides converging evidence showing that when students are specifically asked whether using a smaller versus larger stack of flashcards is better for studying, an overwhelming 72% believe smaller is superior, whereas research has shown that retention is better after practising with a larger versus smaller stack (e.g., Kornell, 2009). Furthermore, not a single student indicated that the reason for their preference on stack size was because it is better for retention. However, although students do not appear to understand the advantage of long versus short spacing, Question 3C shows that at least most students (81%) believe that spaced study is better than massed study.

Although students' understanding of lag effects was of primary interest, Appendix A of the online supplement provides results for questions that are relevant to another aspect of timing. Of interest, a recently emerging issue in the testing effect literature focuses on interleaving versus blocking of practice for items from different categories (e.g., Kornell & Bjork, 2008b; Taylor & Rohrer, 2010). Despite frequent reports of using flashcards for more than one class (Question 1C and A1–4), Question A1–5 indicates that almost no one said they would mix flashcards of different topics when studying.

Putting the pieces together: Calendar question

Tables 2 and 3 report outcomes from fine-grain questions that prompted students to report either on amount of practice or on timing of practice. Table 4 reports outcomes from a more global survey question that was designed to explore how students would make decisions about both the amount and timing of practice and how flashcard use might be combined with the use of other strategies. For this question students were shown a visual calendar display for the month of February along with the informational prompt and questions shown in Table 4.

Not surprisingly almost all students said that they would study the night before an exam (93%) and a fair number reported studying the morning of an exam (38%). Although prior research has shown

that massed study can enhance performance on immediate tests, cramming before an exam is not sufficient for longer-term retention (e.g., Balota, Duchek, & Paullin, 1989; Karpicke & Roediger, 2007; Rawson & Kintsch, 2005). Responses to Question 4A are consistent with outcomes described above, in that almost all students reported that they would study on more than one day. However, results also extend the finding that students do not understand lag effects. For students who reported they would study in two or more sessions, the majority reported that they would only separate their study sessions by about one day.

Responses to Question 4B indicate that a majority of students (65%) said they would use two or three different strategies when studying for the exam. However, students also reported that they would rely more heavily on restudying than on self-testing. Restudying notes had the highest reported percentage for strategies used the morning of an exam, the night before an exam, or any other day, and 52% of students said they would restudy notes during every study session. In contrast only 34% said they would use flashcards the morning of an exam and 38% said they would use flashcards the night before the exam. Likewise only 34% said they would use flashcards on other days, and only a minority (24%) said they would use flashcards on every other day they studied. Although the percentage of students who said they would use flashcards is lower than in other survey questions described above (e.g., 68% in Question 1A), note that Question 4B asked students specifically about preparing for an exam in General Psychology. To revisit Question 1B, a majority of students reported using flashcards for vocabulary (83%) versus information from notes or key concepts (29% and 28%, respectively). Given that General Psychology typically involves learning key concepts and related material rather than vocabulary (as in foreign language classes), the lower percentage of students who reported they would use flashcards in Question 4B is consistent with the percentage who reported using flashcards for key concepts in Question 1B.

DISCUSSION

The current survey extends beyond previous survey research by exploring how and when students self-test, with an emphasis on flashcard use. Of primary interest, two key patterns emerged: First, most students report engaging in criterion learning and

TABLE 4
Responses to calendar question

Here is a calendar for the month of February. Imagine that today is Feb 1st and you have an exam in General Psychology on Feb 27th. Suppose you wanted to get an A on this exam. In the field to the right, answer these questions in as much detail as possible:

4A. Which day or days would you study for the exam (write exact dates)? ($n = 286$)

Participants reporting they would study the morning of the exam	38.5%
Participants reporting they would study the night before the exam	93.0

Not counting the morning of the exam or the night before the exam, number of other days participants reported they would study:

0 days	6.3%
1 day	13.0
2 days	9.1
3 days	13.3
4 days	8.7
5 days	7.3
6 days	9.8
7–8 days	10.8
9+ days	21.7

For participants who reported studying on two or more sessions ($n = 273$), the mean lag between sessions:

1 day (1.0–1.5)	66.3%
2 days (1.6–2.5)	21.2
3 days (2.6–3.5)	7.0
4 days (3.6–4.7)	3.3
7 days (6.5–7.5)	2.2

4B. What exactly would you be doing in each of these study sessions to prepare for the exam? Next to each date you wrote down, write what you would be doing during that session to prepare for the exam ($n = 269$)
Number of strategies participants reported using to study across sessions:

1 strategy	20.0%
2 strategies	32.6
3 strategies	31.9
4 strategies	11.8
5 strategies	3.3
6 strategies	0.4

For participants who indicated they would study the morning of the exam ($n = 108$), how did say they would study? For participants who indicated they would study the night before the exam ($n = 247$), how did say they would study? For participants who indicated they would study on at least one other day ($n = 252$), how did they say they would study?*

	Morning	Night before	Other day(s)
Restudy notes	70.4%	68.8%	65.6%
Restudy textbook	30.6	36.4	40.0
Flashcards	34.3	38.4	34.0
Practice test/ problems	15.7	14.6	10.2
Other	21.3	25.5	21.2

TABLE 4 (*Continued*)

How many sessions on other day(s) participants reported they would use flashcards, restudy the textbook, and restudy notes to prepare for exam ($n = 252$)

	Flashcards	Restudy textbook	Restudy notes
Every session	24.2%	30.6%	51.6%
No session	55.6	48.8	23.0
Some but not all sessions	20.2	20.6	25.4

*This question only includes participants who provided responses for both how and when they would study. The sum of response percentages exceeds 100% because participants could report more than one response

understand the benefits of practising to higher criterion levels. Second, students do not appear to implement or understand the benefits of longer lags when self-testing with flashcards. Concerning learning to criterion, a majority of students choose to reach at least one correct recall during practice before dropping a given flashcard (Questions 2A, 2B, and 2D), although it is interesting to note that students do not always check their answer against external information (Question 2L). Concerning criterion level, some students spontaneously reported going beyond one correct recall during practice (Questions 2A and 2C) and almost all students believe that going beyond one correct recall is important (Questions 2G and 2H). In addition a majority of students report using more than one practice session (Question 2K). Concerning lag, responses indicate a majority of students opt to use smaller versus larger stack sizes when self-testing (Questions 3A and 3B). Of critical importance is that not one student (regardless of preferred stack size) indicated awareness of the relationship between lag and learning (Questions 3A and 3B). Overall the results suggest that when students self-test with flashcards they make advantageous decisions regarding the amount of practice but detrimental decisions regarding the timing of practice.

Note that some questions in the current survey tapped students' beliefs whereas others tapped students' behaviours. We used this multi-faceted approach because what students believe about effective study strategies and what they actually do when studying likely do not completely align (e.g., due to time pressure). For example, nearly all students believe that correctly recalling an item more than once has a substantial benefit on later retrieval (Question 2H), whereas only 34% of students spontaneously reported practising to a

higher criterion when studying (Question 2C). Additionally, if students do not report engaging in a particular behaviour, questions about beliefs help to diagnose whether suboptimal behaviour is due to a metacognitive knowledge deficit or an implementation deficit. For example, 71% of students report dividing up information into smaller stacks (Question 3A); the finding that 72% of students believe using a smaller versus larger stack of flashcards is better (Question 3B) suggests that the suboptimal behaviour is primarily due to a metacognitive knowledge deficit. More generally, we would advocate for the inclusion of both belief and behaviour questions in future survey research.

Our findings concerning how and when students use flashcards have several important implications for educational practice. Perhaps most importantly the results showing a general lack of understanding about how lag affects learning, and that relatively few students effectively implement lag either within or between practice sessions, indicate that students should be educated about the benefits of lag and how to effectively implement lag when self-testing. Another unanticipated but important outcome concerns the extent to which students reported using flashcards for a relatively restricted range of materials (in Question 1B, 83% of students report using flashcards for vocabulary whereas only 29% report using flashcards for key concepts). Thus students would also likely benefit from education about the utility of flashcards for learning a wider range of material. In addition, responses concerning how students evaluate the correctness of their answers (Question 2A) suggest that students should also be educated about the importance of explicitly checking their answers, particularly when using flashcards to learn complex material.

Beyond implications for more effective flashcard use in particular, the current survey results have practical implications for supporting successful student learning with the use of self-testing more generally. Although a majority of students choose to self-test during study, the current results suggest that students would benefit from more information or training about how to effectively implement self-testing. One straightforward approach would be for instructors to teach students about the importance of the amount and timing of self-testing. Unfortunately the current results suggest that relatively few instructors advise students to use self-testing as a learning strategy (Question A1–3), much less how to self-test effectively. Given that few instructors recommend

self-testing to students, educators themselves may be unaware of the robust effects of self-testing and the factors that influence the efficacy of self-testing. Ideally teacher training programmes could educate instructors on the benefits of self-testing, factors that influence the effectiveness of self-testing, and ways to incorporate self-testing methods into curricula. Another approach to enhance the effective use of self-testing is through study skills classes or workshops increasingly offered by universities to support student achievement. Ideally these programmes would target first-semester freshmen, who may have never been taught about the use of self-testing. The goal would be to educate students on how and when to self-test in ways that facilitate learning (e.g., higher criterion levels and longer lags). More generally, promoting students' awareness of the benefits of self-testing and providing knowledge of how to effectively self-test in real-world contexts will support more successful student learning.

REFERENCES

Bahrick, H. P., & Hall, L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language*, 52, 566–577.

Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag, and retention interval. *Psychology and Aging*, 4, 3–9.

Carrier, M. L. (2003). College students' choices of study strategies. *Perceptual and Motor Skills*, 96, 54–56.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, 19, 1095–1102.

Dunlosky, J., Hartwig, M. K., Rawson, K. A., & Lipko, A. R. (2011). Improving college students' evaluation of text learning using idea-unit standards. *Quarterly Journal of Experimental Psychology*, 64, 467–484.

Dunlosky, J., & Rawson, K. A. (in press). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*.

Hartwig, M. K., & Dunlosky, J. D. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, 19, 126–134.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval

during learning. *Journal of Experimental Psychology: General*, *138*, 469–486.

Karpicke, J. D., & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory,and Cognition*, *37*, 1250–1257.

Karpicke, J. D., Butler, A. C., & Roediger, H. L. III. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory*, *17*, 471–479.

Karpicke, J. D., & Roediger, H. L. III. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 704–719.

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

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, *23*, 1297–1317.

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. (2008a). Optimising self-regulated study: The benefits – and costs – of dropping flashcards. *Memory*, *16*, 125–136.

Kornell, N., & Bjork, R. A. (2008b). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science*, *19*, 585–592.

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

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, *17*, 493–501.

Lipko, A. R., Dunlosky, J., Hartwig, M. K., Rawson, K. A., Swan, K., & Cook, D. (2009). Using standards to improve middle-school students' accuracy at evaluating the quality of their recall. *Journal of Experimental Psychology: Applied*, *15*, 307–318.

McCabe, J. (2010). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, *39*, 462–476.

Pashler, H., Zarow, G., & Triplett, B. (2003). Is temporal spacing of tests helpful when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 1051–1057.

Pyc, M. A., & Dunlosky, D. (2010). Toward an understanding of students' allocation of study time: Why do they decide to mass or space their practice? *Memory & Cognition*, *38*, 431–440.

Pyc, M. A., & Rawson, K. A. (2007). Examining the efficiency of schedules of distributed retrieval practice. *Memory and Cognition*, *35*, 1917–1927.

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 and Language*, *60*, 437–444.

Pyc, M. A., & Rawson, K. A. (2012). Why is test–restudy practice beneficial for memory? An evaluation of the mediator shift hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 737–746.

Rawson, K. A., & Dunlosky, J. (2011). Optimising schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, *140*, 283–302.

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend upon time of test. *Journal of Educational Psychology*, *97*, 70–80.

Rawson, K. A., O'Neil, R. L., & Dunlosky, J. (2011). Accurate monitoring leads to effective control and greater learning of patient education materials. *Journal of Experimental Psychology: Applied*, *17*, 288–302.

Roediger, H. L. III., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Science*, *15*, 20–27.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, *24*, 837–848.

Vaughn, K. E., & Rawson, K. A. (2011). Diagnosing criterion-level effects on memory: What aspects of memory are enhanced by repeated retrieval? *Psychological Science*, *22*, 1127–1131.

APPENDIX A

Background information and outcomes of secondary interest.

A1–1. What is your college standing, based on number of credit hours completed? ($n = 372$)

Freshman	66.1%
Sophomore	15.9
Junior	8.9
Senior	9.1

A1–2. What was your approximate GPA in high school? ($n = 370$) What is your approximate GPA in college? ($n = 308$)

	High school	College
Below 2.00	0.8%	2.9%
2.00 – 2.49	4.1	10.7
2.50 – 2.99	18.4	25.3
3.00 – 3.49	35.1	32.5
3.50 – 4.00	38.4	28.6
4.00 +	3.2	n/a

A1–3. Has anyone ever given you advice or taught you how you should study? ($n = 135$)

Yes	66.7%
No	33.3

When was it (what grade were you in)? ($n = 67$)*

Grade school	13.4%
Middle school	11.0
High school	47.8
College	40.3
All grades	6.0

Who was it (for example, a teacher or a friend)? ($n = 86$)*

Teacher/professor	75.6%
Friend	23.3
Parent	19.8
Other	7.0

APPENDIX A (Continued)

What did they tell you to do when studying? ($n = 73$)*

	Overall	Teacher/professor ($n = 40$)
Flashcards	37.0%	22.5%
Make outline	27.4	27.5
Read/reread material	27.4	30.0
Practice problems/quizzes	24.7	25.0
Find out best method for self	17.8	17.5
Make up system (rhymes, colour code)	12.3	12.5
Work in groups	6.8	5.0
Make charts/graphs	4.1	5.0
Other	4.1	5.0

A1–4. Have you ever been in a situation in which you were using flashcards to study two topics at the same time (for example, might have had two exams, in Spanish and Chemistry)? ($n = 149$)

Yes	59.1%
No	40.9

A1–5. Suppose you had sets of flashcards for two different classes. Would you mix the cards from the two topics and study them together or would you keep the cards from the two topics separate? ($n = 150$)

Don't mix	97.8%
Mix	2.2

If you said don't mix, explain why you would study them separately ($n = 141$)*

Confusing	63.1%
Different subjects	53.9
Mixed hurts memorisation	17.0
Separate helps memorisation	6.4
Makes more challenging	3.5
Classes coincide	2.1

*For these questions the sum of response percentages exceeds 100% because participants could report more than one response.