Reinventing Flashcards to Increase Student Learning

Article *in* Psychology Learning and Teaching · July 2017

DOI: 10.1177/1475725717719771

4 authors:

Sawa Senzaki
University of Wisconsin - Green Bay
13 PUBLICATIONS **82** CITATIONS

Jana Hackathorn
Murray State University
44 PUBLICATIONS **278** CITATIONS

Drew Appleby
Indiana University-Purdue University Indianapolis
28 PUBLICATIONS **111** CITATIONS

Regan A. R. Gurung
University of Wisconsin - Green Bay
115 PUBLICATIONS **4,718** CITATIONS

Article

PSYCHOLOGY LEARNING & TEACHING

Psychology Learning & Teaching
0(0) 1–16

DOI: 10.1177/1475725717719771
journals.sagepub.com/home/plat
SAGE

Reinventing Flashcards to Increase Student Learning

Sawa Senzaki
University of Wisconsin-Green Bay

Jana Hackathorn
Murray State University

Drew C. Appleby
Indiana University-Purdue University Indianapolis

Regan A. R. Gurung
University of Wisconsin-Green Bay

Abstract

Two studies examined the effectiveness of a flashcard-based study strategy, *Flashcards-Plus*, in an ecologically valid context. The strategy requires students to create flashcards designed to increase their ability to retain, comprehend, and apply textbook material to exams. In Studies 1a ($n = 73$) and 1b ($n = 62$), we introduced all students to the *Flashcards-Plus* method and compared their exam scores. Students who used this strategy scored significantly higher than those who did not. In Study 2 ($n = 434$), we randomly assigned six introductory psychology courses to either receive a classroom lecture with the *Flashcards-Plus* strategy (i.e., three experimental courses) or no lecture (i.e., three control courses). Students in the experimental courses scored significantly higher than those in the control courses after the lecture. The results from all three studies demonstrate that students who were introduced to the *Flashcards-Plus* study strategy scored significantly higher on exams following the lecture than students who were not. These findings suggest that this easily implemented teaching strategy can help students achieve deeper levels of processing (i.e., comprehension and application) in a self-directed manner, which benefit students' performance.

Keywords

Flashcards, study strategies, pedagogy, exams

Corresponding author:
Sawa Senzaki, UW-Green Bay, 2420 Nicolet Drive, MAC C315, Green Bay, WI 54311, USA.
Email: senzakis@uwgb.edu

Introduction

The study techniques students use can affect their exam performance (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Hattie, 2012, p. 265), but not all studying techniques are equally effective (Balch, 1998; Dunlosky et al., 2013; Gurung & McCann, 2012; Hackathorn et al., 2012). We focused on flashcards, a commonly used study method (Hartwig & Dunlosky, 2012) by studying the effectiveness of a new way to use flashcards, the *Flashcards-Plus (FP)* technique (Appleby, 2013), in two different naturalistic classroom environments. *FP* expands the use of flashcards to help students understand and prepare for three types of multiple-choice questions commonly found on psychology tests: those that measure retention, comprehension, and application.

In reviewing the efficacy of 10 commonly used studying techniques, Dunlosky and colleagues (2013) identified practice testing (or practice retrieval) and distributed practicing (or spaced practice) as the two most effective techniques. Students can engage in both of these high utility techniques through the use of flashcards (Wissman, Rawson, & Pyc, 2012). In the basic form of flashcards, a term or concept appears on one side of an index card and its definition appears on the other side. Students look at the term and test whether they know the answer, flipping the card to check. Surprisingly little classroom research on college student flashcard use exists and has yielded mixed results (Golding, Wasarhaley, & Fletcher, 2012; Hartwig & Dunlosky, 2012).

Although flashcards are relatively easy to create and commonly used by college students (Wissman et al., 2012), the effectiveness of student-created flashcards in a classroom is unclear. In an introductory psychology class, Golding et al. (2012) found that students who used flashcards for the first exam scored higher than students who did not. However, the use of flashcards was not beneficial on the second exam, and the effect was only marginal on the third exam. Furthermore, Hartwig and Dunlosky (2012) demonstrated that student-created flashcards use was unrelated to students' grade point average (GPA). These results suggest that retrieval practice and distribution of practice increases retention of information, but the way in which flashcards are typically used may not be sufficient to be successful in college courses.

One possible reason for the ineffectiveness of the traditional use of flashcards is that students (70%) use self-testing techniques such as flashcards to determine how well they have learned the current material instead of to learn more (18%, Karpicke, Butler, & Roediger, 2009). Flashcards can facilitate repetitive learning, but repetition is only a superficial level of processing (Brown, Roediger, & McDaniel, 2014). While 82.9% of students in a study indicated they used flashcards to remember vocabulary, no students reported using flashcards to develop deeper understanding of or the application of concepts (Wissman et al., 2012).

When students use deep-level processing, they focus on substance and the underlying meaning of the information, make a personal commitment to understanding, and concentrate on relationships between different pieces of information they encounter (Entwistle, 2009). Students also strive to apply what they are learning to their everyday lives and integrate and synthesize information with their prior learning, essentially moving up to the next level of Bloom's (revised) hierarchy of learning (Anderson & Krathwohl, 2001). Moreover, deep learning is associated with more enjoyable learning, whereas the surface approach tends to be less satisfying (Biggs, 2003; Tagg, 2003). The *FP* method goes beyond the traditional use of flashcards to facilitate deeper understanding and increase learning.

The *FP* strategy begins in the same way as traditional flashcards. Students identify boldfaced terms from the textbook and write them on one side of a notecard. Students write the textbook definition on the other side. Students memorize the definition from the textbook (or the instructor) to increase retention. While most students stop here with traditional flashcards, *FP* creates deeper levels of learning by having the student perform two more actions. In the next step, students write a definition for the same key term in their own words. Rephrasing the term helps students understand the material that will increase comprehension. Finally, students generate a realistic example of the key term from their own lives that will increase application.

Appleby (2013) originally designed the *FP* strategy as a study aid for multiple-choice question exams to promote active learning in students, particularly those enrolled in large introductory courses. The strategy employs a scaffolding technique (Vygotsky, 1978), during which students learn how to move beyond their current developmental skill set and master more difficult tasks. While most exam questions assess students' retention, comprehension, and application (Gronlund, 1998; Palmoba & Banta, 1999), many students need support to move beyond simple retention to understand and apply the material (Appleby, 2008, 2013). The *FP* strategy offers a clear explanation of the importance of the higher-level understanding and a strategy to achieve this goal in a self-directed manner. Instructors can find full details on how to use the *FP* strategy online (with downloadable PowerPoint slides): http://teachpsych.org/Resources/Documents/otrp/resources/appleby13flashcard.pdf/. The introduction of the slides takes approximately 10 minutes in class.

The *FP* strategy involves one of the most well-known memory-facilitating processes. Remembering is enhanced when the meaning of material is fully processed at the time of encoding (e.g., Craik, 1979, 2002; Craik & Lockhart, 1972; Paller, Kutas, & Mayes, 1987). For example, students who are asked to create their own explanations of a concept recalled that information better than those who were given explanations passively (Pressley, McDaniel, Turner, Wood, & Ahmad, 1987). Furthermore, information processed in relation to one's self results in particularly strong recall (Klein & Loftus, 1988; Rogers, Kuiper, & Kirker, 1977).

It is, however, also well known that engaging in practice retrieval significantly increases long-term retention of knowledge (Karpicke & Blunt, 2011; Karpicke et al., 2009; Karpicke & Roediger, 2010; Kornell, Hays, & Bjork, 2009). Interestingly, researchers suggest that an essential component of retrieval-based learning is remembering the context of learning; that is, the act of thinking back to what occurred in a particular place and time in association to the concept to be learned (Karpicke, Lehman, & Aue, 2014; Tulving, 1983). Taken together, retrieval effects can be even stronger if the information is related to students' own experiences.

Finally, a large number of studies have shown that explicit instruction to provide basic tools for academic skills is a crucial characteristic of pedagogical scaffolding (Archer & Hughes, 2011). The *FP* strategy is an ideal application of these empirical studies to help students engage in meaningful learning that can be implemented in the classroom. With *FP,* students: (1) receive an explicit instruction of what is expected; (2) develop an example and apply the concepts they are required to learn; and (3) engage in practice retrieval of all three levels (i.e., retention, comprehension, and application) of the assigned material.

In the current two studies, we tested the efficacy of the *FP* strategy. Studies 1a and 1b used quasi-experimental designs comparing students who volunteered to use the method with those who did not. In Study 2, we randomly assigned students to either receive instruction regarding the *FP* method or not. In these two studies, we hypothesized that students who used the method (Study 1) and students who were exposed to the *FP* method (Study 2) would learn better, which we operationalized as higher scores on their course exams.

Study 1

Study 1 had three goals. We sought to obtain an initial assessment of the effectiveness of the *FP* strategy on the exam performance, whether this strategy was related to any other studying techniques, and whether students' perceptions of using *FP* influenced their performance (Study 1a). In Study 1, students chose to use the *FP* or not as extra credits, and we expected students who used the *FP* method to improve their performance on the exam after implementing the studying strategy. In Study 1b, we aimed at replicating the results by controlling for the timing of the introduction of the *FP* studying method, and also gather qualitative data to further understand students' experience with the *FP* method.

Study 1a

Method

Participants. Seventy-three undergraduate students from two sections of an upper level course in Human Development at a midsized Midwestern public university participated in the study. The first author taught both classes. Students were predominantly women ($n = 62$) and juniors ($n = 40$), followed by seniors ($n = 20$) and finally sophomores ($n = 13$). The mean age was 21.5 years (standard deviation (SD) = 3.25, range = 19–41) and the mean self-reported GPA was 3.0 ($SD = 0.51$).

Materials and Procedure. The study took place over the course of one semester, and we measured performance on each exam as the outcome. There were three non-cumulative exams, each consisting of approximately 30 multiple-choice questions and four short-answer questions. All students took the first exam before we delivered instructions for the *FP* strategy. After Exam 1, students in Section 1 were given access to Appleby's original PowerPoint slides (Appleby, 2013) via an online course management system, Desire 2 Learn. After Exam 2 the instructor gave students in both sections detailed instruction on the *FP* strategy with the PowerPoint slides in class, to encourage active exposure. The instructor also emphasized the rationale behind the flashcards studying strategy and provided extra credit if they used the *FP* strategy to study for Exam 3. Students submitted their flashcards to the instructor to ensure they followed the three steps as described. The same instructor taught both sections, and the only difference between sections being when the *FP* method was introduced.

At the end of the semester students completed a studying strategy survey, assessing *FP* method use (e.g., hours per week and perceptions of the helpfulness and difficulty). The survey also asked about the use of twelve other studying techniques: summarizing material; highlighting/underlining; memorizing keywords using mnemonics; using imagery for text; explaining material to yourself; rereading; practice testing; spreading out study sessions (rather than cramming); studying material with a friend; making up examples to help understand; applying material to their lives; and taking notes on the book (modified from Bartolewski & Gurung, 2014). The reliability of this self-report scale was relatively high (Cronbach's $\alpha = 0.79$). Finally, we asked students to think back to the beginning of the

Table 1. Means (and Standard Deviations (SDs)) of Exam Scores and Grade Point Average (GPA) Separated by Sections in Studies 1a and 1b.

		Exam 1 *Mean (SD)*	*Exam 2* *Mean (SD)*	*Exam 3* *Mean (SD)*	*GPA* *Mean (SD)*
Study 1a	Section 1	0.78 (0.10)	0.74 (0.12)	0.85 (0.10)	3.09 (0.46)
	Section 2	0.74 (0.12)	0.74 (0.14)	0.83 (0.12)	3.00 (0.57)
Study 1b	Section 1	0.77 (0.15)	0.78 (0.11)	0.83 (0.10)	3.25 (0.49)
	Section 2	0.72 (0.12)	0.75 (10)	0.78 (.14)	3.01 (0.51)

semester and report how motivated they were to take the course at that time. Students provided a signed informed consent to complete the survey.

Results and Discussion. A similar proportion of students reported using flashcards across two sections of the course, $\chi^2(1) = 1.78$, $p = 0.18$. Furthermore, exam average scores did not differ across two sections, Exam 1, $t(71) = 1.48$, $p = 0.15$, Exam 2, $t(70) = 0.11$, $p = 0.91$, and Exam 3, $t(71) = 0.38$, $p = 0.70$ (see Table 1 for descriptive statistics of the preliminary analyses). Therefore, we combined the data from Sections 1 and 2 for the main analyses. Combining both sections, approximately half ($n = 39$) of the students reported using the *FP* studying strategy to study for Exam 3, and the other half of the students ($n = 34$) reported that they did not use this method. Students who choose to use the *FP* method and those who did not had similar GPA, $t(71) = 0.91$, $p = 0.37$ and initial motivation, $t(71) = 1.21$, $p = 0.23$. Because the scores are in percentages, an arcsine transformation was conducted for analyses, although we used percentage scores for descriptions of means.

We examined whether exam performance improved significantly as a result of *FP* use. A mixed-model analysis of variance (ANOVA) with Exam as the repeated-measure variable and *FP* (Users vs. Non-users) the between-subject variable yielded a main effect of Exam, $F(1, 71) = 48.44$, $p < 0.001$, $\eta_p^2 = 0.41$. Both *FP* users and non-users scored significantly higher on Exam 3 (mean $(M) = 84.49\%$, $SD = 11.07$) than on Exam 1 ($M = 75.99\%$, $SD = 11.12$). There was no main effect of *FP* strategy use, $F(1, 71) = 1.07$, $p = 0.31$, but there was a significant interaction effect between Exam and *FP*, $F(1, 71) = 9.36$, $p = 0.003$, $\eta_p^2 = 0.12$ (Figure 1). Although both *FP* users and non-users improved on their exam scores from Exam 1 to Exam 3, this effect was stronger for *FP* users, $t(38) = 7.34$, $p < 0.001$, $d = 1.11$, than non-users $t(33) = 2.67$, $p = 0.012$, $d = 0.47$.[1]

Only two other studying methods significantly differentiated between *FP* users and non-users: "making up examples to help understand" and "practice testing." *FP* users ($M = 3.33$, $SD = 1.11$) reported a significantly higher rate of making up examples than non-users ($M = 2.65$, $SD = 1.13$), $t(71) = 2.62$, $p = 0.011$. *FP* users ($M = 2.56$, $SD = 1.23$) also reported significantly higher rate of practice testing than non-users ($M = 2.00$, $SD = 1.07$), $t(71) = 2.07$, $p = 0.042$ (see Table 2 for a complete list of analyses).

Students' perception of the easiness of the *FP* strategy ($M = 1.97$, $SD = 0.78$) significantly correlated with their exam score, $r(39) = 0.53$, $p < 0.001$. Students who indicated that the use of *FP* method was easy scored higher than students who indicated that the method was difficult. Students' perception of the helpfulness ($M = 3.23$, $SD = 0.71$) and the hours they spent on the method ($M = 3.33$, $SD = 2.46$; range = 0–10 hours) were not correlated with their exam scores.

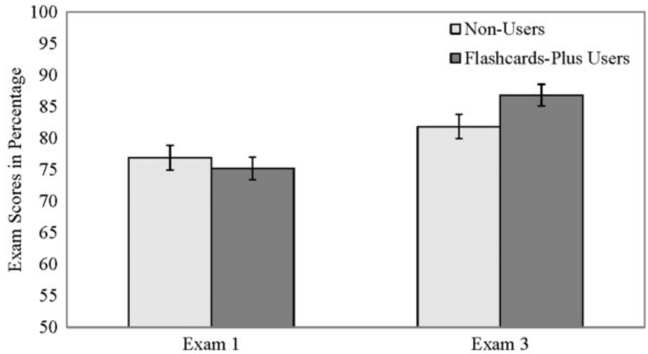

Figure 1. Mean (and standard errors) exam scores on Exams 1 and 3 from Flashcards-Plus users and non-users in Study 1a.

Table 2. Motivation Level and Studying Methods in Study 1a.

	Users		*Non-Users*		
	Mean (M)	*Standard Deviation* (SD)	M	SD	*t-test*
Motivation	4.15	0.84	3.85	1.28	1.21
Summarizing	3.54	0.913	3.18	0.968	1.64
Highlighting/underlining	3.97	0.986	3.53	1.331	1.64
Mnemonics	3.72	1.099	3.65	1.125	0.27
Imagery	2.97	0.811	2.82	1.086	0.68
Explaining to yourself	3.64	0.903	3.50	1.261	0.55
Rereading	3.79	0.894	3.50	1.187	1.21
Practice Test	2.56	1.231	2.00	1.073	2.07*
Spacing	3.31	1.080	2.91	1.083	1.56
With a friend	2.24	1.149	1.97	1.058	1.02
Examples	3.33	1.108	2.65	1.125	2.62*
Applying	3.51	0.970	3.06	1.205	1.78**
Notes	3.10	1.170	2.85	1.160	0.92

*$p < 0.05$; **$p = 0.079$.

Study 1b

In Study 1b we better controlled the timing of the introduction of this strategy to the students. We also asked whether students have used flashcards to study in the past and to further indicate their experience with the *FP* studying method in an open-ended format.

Method

Participants. Sixty-two students in two sections of an upper level course in Human Development at a midsized Midwestern public university participated in the study. Students were predominantly women ($n = 52$) and sophomores ($n = 32$), followed by juniors

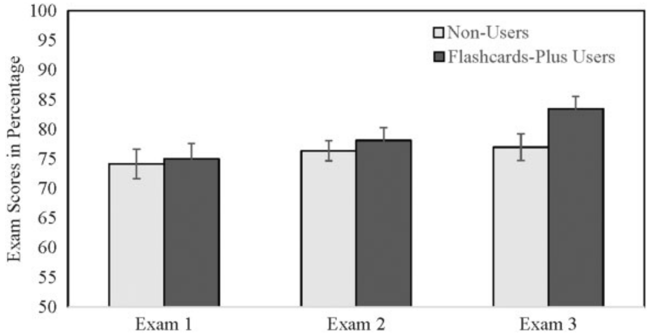

Figure 2. Mean (and standard errors) exam scores on Exams 1 and 3 from Flashcards-Plus users and non-users in Study 1b.

($n = 20$) and seniors ($n = 7$). The mean age was 20.9 years ($SD = 4.34$, range = 18–43) and the mean self-reported GPA was 3.18 ($SD = 0.55$).

Materials and Procedure. Study 1b used the same material and followed the same procedure as Study 1a. In both sections, all students received active instruction about the *FP* method immediately after Exam 2, approximately 7 weeks before Exam 3. At the end of the semester, students completed a survey that asked whether they used the *FP* study method, the frequency of their use, their perception of its helpfulness, whether they have used flashcards to study before, and an open-ended question about their experience with the *FP* method. Students received extra credit for completing the survey regardless of their *FP* method use.

Results and Discussion. We found no significant differences between the exam scores of sections 1 and 2: Exam 1, $t(60) = 1.51$, $p = 0.14$, Exam 2, $t(60) = 1.01$, $p = 0.28$, and Exam 3, $t(60) = 1.82$, $p = 0.06$. Furthermore, the reported GPA did not differ significantly across two sections, $t(60) = 1.84$, $p = 0.07$; thus we combined the data from Sections 1 and 2 in the main analyses. After combining both sections, 36 students reported using the *FP* studying strategy to study for the third exam, and 26 students reported that they did not use this method. GPA between *FP* users and non-users did not differ, $t(60) = 0.81$, $p = 0.42$.

To test the effectiveness of the *FP* strategy, we conducted a mixed-design ANOVA[2] with Exam as the repeated-measure variable and *FP* as the between-subject variable. The results revealed a main effect of Exam, $F(2, 120) = 10.80$, $p < 0.001$, $\eta_p^2 = 0.15$, as exam performance increased over time in general. There was no main effect of the use of *FP*, $F(1, 60) = 0.91$, $p = 0.35$, but similar to Study 1a, there was a significant interaction effect between Exam and *FP*, $F(2, 120) = 3.09$, $p = 0.049$, $\eta_p^2 = 0.05$ (Figure 2). Exam performance of *FP* increased significantly over time, $F(2, 70) = 20.18$, $p < 0.001$, $\eta_p^2 = 0.37$. Conversely, the three exam scores were not significantly different for non-users, $F(2, 50) = 0.75$, $p = 0.48$.

Twenty students reported that they have used flashcards to study in the past, and after excluding these 20 students, the difference on Exam 3 was still significant between *FP* users and non-users, $t(40) = 2.18$ $p = 0.04$. Among students who completed the *FP* strategy, the hours they used the *FP* strategy ($M = 4.96$, $SD = 2.95$, range = 1–10 hours) were not associated with the grade on their exam $r(36) = 0.02$, $p = 0.90$. As shown in Table 3, students generally had a positive perspective on this studying strategy, citing some of the main characteristics such as application and repetition.

Table 3. Students' Perception of the Helpfulness of the Flashcards-Plus System (Study 1b)

The method helped in remembering a topic or concept during the exam and after due to relating it to real life situations.
It was helpful to simply take the time to review each concept and better understand each one rather than reading the definition.
It was helpful to write the textbook definition and then create my own. The writing aspect alone helped.
It made me think about how the term applies in real life.
It helped because it forced you to think about each individual term and how it fits with the context.
Repetition helped me memorize things.
I was forced to think about terms not only in my own words, but also as an application. It made remembering a lot easier!
It helped because it allowed me to learn a term in multiple ways (definition, my own example, etc.). But this method was very tedious and took a long time.
It helped me relate the terms back to my life. I had a clearer picture of the terms afterwards.
I think it really helped me in learning definitions and how to apply them. By putting the definition in my own words and giving an example, I was forced to truly understand what I was studying which made the exams so much easier!

The results of Studies 1a and 1b demonstrated that *FP* helped students increase exam scores. Although these two studies provide us with rich data in real classrooms, classroom interventions are often difficult to examine, and this intervention was no exception. All students were provided the instructions and, as a result of this, a limitation of the two studies presented here was that the students' self-selected to use the *FP*. Moreover, students were offered extra credit, which may have inadvertently led to Hawthorne or novelty effects. To address these questions, Study 2 introduced randomization to examine the effects of exposure to the study strategy.

Study 2

In this study, we randomly selected class sections to receive the *FP* instruction. We hypothesized that students who received the *FP* instruction would perform better than students who did not receive the instruction and who studied as they normally did.

Method

Participants. Participants included students enrolled in six different introductory psychology courses at a Midwestern regional state university ($n = 434$). Different instructors (none of them authors of the current research) taught different sections of the same course. We collected student demographic information via a self-report measure given on the last day of the course. The population was predominantly women (62.4%), and first year ($n = 244$; 56.2%), but also contained sophomores ($n = 87$), juniors ($n = 40$), and seniors ($n = 21$). The majority of the students were Caucasian/White (74%), however other ethnicities were also represented: African-American/Black (9.7%); Asian (3.2%); Hispanic (1.8%); biracial (2.1%); Middle Eastern (1.2%); Native American (0.5); and others (1.2%).

Materials and Procedure. During the semester, the second author visited each classroom to provide a short lecture regarding improving studying skills (e.g., distributed rehearsal, avoiding distractions, and situation dependent memory). We randomly assigned three ($n = 231$ students) of the six sections to receive the introduction of the *FP* strategy. In the three experimental sections, the students received the same active instruction regarding the *FP* as in Study 1. The control sections did not receive any instructions regarding studying with flashcards. None of the students were offered extra credit for employing any study strategy; they were simply informed that using any of these strategies could result in positive outcomes. This was to control for potential Hawthorne or good participant biases.

Additionally, the experimental sections received the *FP* lecture at different times in the semester. One section received the instructions after Exam 1 ($n = 79$), one received the instructions near midterm immediately following Exam 3 ($n = 81$), and one section received the instructions prior to the Final Exam ($n = 67$). The way that we delivered the intervention to three classes at different times is based on the time-series experiments, particularly multiple baseline studies (Biglan, Ary, & Wagenaar, 2000). This design is useful in fields such as community intervention research and education (Kratochwill, 1978), because if the intervention is introduced at the same point of time for many groups, change in the dependent variable could be due to numerous other factors that co-occur with the change in the independent variable. For example, a particular exam that occurs before or after the intervention can be particularly easy or difficult. By introducing the intervention at a different point in time, we can avoid this limitation and be more confident in the results of change in exam performance.

At the end of the semester, the author who gave the study strategy lecture as well as a graduate student returned to each class. Students were given a hard copy of an informed consent form. It was explained that the effectiveness of multiple teaching strategies was being examined and that access to exam grades and the contents of a brief survey would be helpful in the investigation. Students were informed that completing the survey would indicate consent. Finally, students were told that the information would be entered by the graduate student, and the instructor would never see who consented and who did not.

Following the consent process, students completed a brief survey that asked whether they used the *FP* studying method and their frequency of studying. In total, 39 students did not complete the survey which indicated non-consent, absence, or potentially dropped enrollment. The procedure was approved by the institutional review board.

Results

Effects on Exam Scores. As each of the experimental sections received the *FP* lecture at different times and the number of total exams varied (either four or five) across sections, exam scores were recorded so we could compare how well students performed before and after the lecture that introduced the *FP* method. The mean exam scores (in percentage) and the SDs were listed in Table 4. Thus, exam scores prior to and following the lecture were averaged. Students in the control courses were subject to the same coding as the experimental course who received the lecture prior to Exam 2, so the scores of Exam 1 were used as pre-scores and the average of Exam 2 through the final exam were used as post-scores. As before, we used arcsine transformation for analyses. The exam scores were then submitted to a 2 (Condition: Experimental vs. Control) $\times$ 2 (Exposure: Pre- vs. Post- *FP* lecture) mixed-model ANOVA.

The results indicate there was a main effect of exposure, $F(1, 431) = 38.53$, $p < 0.001$, $\eta_p^2 = 0.08$, as students' exam scores increased over time. This main effect was qualified by

Table 4. Means (and Standard Deviations (SDs)) of Exam Scores in Percentage Separated by Sections in Study 2

	Exam 1 Mean (SD)	*Exam 2 Mean (SD)*	*Exam 3 Mean (SD)*	*Exam 4 Mean (SD)*	*Exam 5 Mean (SD)*
Control 1	0.64 (0.16)	0.66 (0.13)	0.72 (0.14)	0.74 (0.13)	–
Control 2	0.70 (0.12)	0.71 (0.14)	0.72 (0.13)	0.72 (0.11)	–
Control 3	0.74 (0.12)	0.72 (0.11)	0.70 (0.13)	0.75 (0.14)	0.75 (0.14)
Experiment 1	0.66 (0.12)	0.73 (0.13)	0.73 (0.13)	0.77 (0.12)	0.74 (0.14)
Experiment 2	0.67 (0.17)	0.69 (0.13)	0.78 (0.13)	0.78 (0.14)	0.76 (0.13)
Experiment 3	0.72 (0.12)	0.72 (0.12)	0.74 (0.14)	0.73 (0.12)	0.72 (0.11)

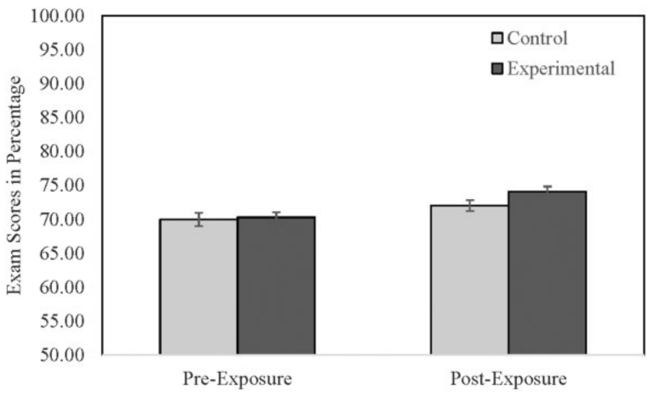

Figure 3. Mean (and standard errors) averaged exam scores in percentage pre- and post-exposure from control and experimental groups in Study 2.

an interaction between condition and exposure, $F(1, 491) = 4.89$, $p < 0.05$, $\eta_p^2 = 0.01$, as the experimental group scored higher than the control group after the exposure to the *FP* strategy (Figure 3). Finally, there was no main effect of condition, $F(1, 431) = 0.82$, $p = 0.37$, indicating that the experimental and control groups were not different other than the *FP* strategy.

Flashcard-Plus Use. Because students were introduced to the *FP* method randomly without any extra credit for using it, we also examined whether or not students used the new system. According to self-report, 48.3% of students reported they used flashcards of any format to study during the semester. Slightly over half ($n = 207$) reported that they did not use flashcards to study, and the remaining students used them sometimes ($n = 107$) or all the time ($n = 86$).

The experimental sections contained 231 students who were exposed to the *FP* studying strategy. A binary logistic regression using exposure to predict whether or not students will report using flashcards in general was significant, $\chi^2(1, n = 438) = 18.19$, $p < 0.001$, *Nagelkerke* $R^2 = 0.06$. Specifically, individuals who were exposed to the new study strategy were twice as likely to use flashcards at some point in the semester ($B = 0.87$, $Wald = 17.76$,

$p < 0.001$, odds ratio $(OR) = 2.38$); 37.3% of the students reported using flashcards in the control courses, whereas 58.3% of the students reported using flashcards in the experimental courses, $\chi^2(1, n = 400) = 18.04$, $p < 0.001$.

Additionally, students were asked to describe the type of flashcard system they used., Over one-third of the experimental students ($n = 114$) who responded to the question described using the *FP* strategy (38.6%; $n = 44$). For example, one student responded "concept on front, and book's definition, my definition, and example on back." Thus, a second binary logistic regression was conducted, using exposure condition to predict whether or not students will use the *FP* system. The results indicated the exposure was significant, $\chi^2(1, n = 438) = 32.68$, $p < 0.001$, *Nagelkerke* $R^2 = 0.18$. Specifically, individuals who were exposed to the new study strategy were 10 times as likely to use the *FP* study strategy ($B = 2.37$, $Wald = 19.45$, $p < 0.001$, $OR = 10.72$); only 3.5% ($n = 4$) of the students reported using a *FP* technique in the control courses, whereas 28% ($n = 46$) of the students in the experimental courses reported using the *FP* system, $\chi^2(1, n = 278) = 27.46$, $p < 0.001$.

Discussion. The advantage of Study 2 was that it experimentally manipulated the instruction of the *FP* system by randomly assigning exposure over six introductory psychology courses. Our analysis indicates that individuals who are exposed to the *FP* system were more likely to use flashcards to study in general than those who were not exposed to it, and they were much more likely to use the new *FP* system than those who were not exposed to it. This duplicates past studies indicating many students may not know how to study for college level exams, but are eager to learn how to do so (Bjork, Dunlosky, & Kornell, 2013).

Our analyses also indicate that introducing students to the system has a positive influence on exam scores and study habits. Overall, students exposed to the *FP* strategy fared better on exams than those who were not, and their exam grades increased at a greater rate. This may hint that students are actively searching for ways to improve their grades and/or study habits.

It is possible that being exposed to the *FP* system draws students' attention to the types of studying that is appropriate for college level questions. We conducted a brief exploratory analysis of the types of studying the students reported to determine if students mentioned higher levels of learning (e.g., application, and creating examples). We found a significant shift in student thinking regardless of whether they used the new flashcard strategy, $\chi^2(n = 298) = 17.29$, $p < 0.001$. That is, the control condition explicitly mentioned higher levels of learning approximately 19% of the time, whereas the experimental conditions mentioned higher levels 81% of the time. This may indicate that although the participants did not explicitly mention using the *FP* strategy in the survey, they may have inadvertently changed the level at which they study for exams.

General Discussion

The present studies report an initial attempt to examine the effectiveness of the enhanced flashcards studying strategy developed by Appleby (2013) in real classrooms. Study 1 demonstrated that students who used *FP* improved their exam performance more than students who did not use this strategy. Although we did not randomly assign *FP* method use to specific sections because we wanted to assess how many students would decide to use it when given the option, we demonstrated that *FP* users and non-users in the current study were not different in many demographic characteristics including their GPA and general motivation. Importantly, students who used the *FP* strategy were more likely to create examples of important concepts and engage in practice testing than students who did not use the strategy. *FP* users' reports of higher rates of these two studying methods is consistent with the idea that this strategy promotes application of concepts through the creation of original examples and subsequent practice testing of these concepts.

In Study 2, we experimentally manipulated *FP* use by introducing the techniques to three of six classes, and found that exposed students scored higher than those in the control group. Although the effect size for actual users was relatively small, we believe the results of these studies can provide encouragement for classroom teachers who would like to foster higher-order thinking in their students. Additionally, we believe that because our study is high in ecological validity (i.e., it was conducted in the actual classroom and did not force students to engage in any technique), some of these effect sizes may have been hindered. Future studies that examine this strategy in a laboratory setting may find increased effect sizes due to the maximization of the independent variable. We cautiously infer that the use of the *FP* method can contribute to increases in exam scores, but urge additional pedagogical research on this topic.

Previous research indicates that when students create their flashcards in a traditional manner, they tend to focus on the materials that they have already learned by checking their knowledge (Karpicke et al., 2009). Our findings add to previous work on flashcards by encouraging students to go beyond simple repetition of a definition to deeper processing. By paraphrasing the textbook definition and then generating an application, *FP* enables students to ascend Bloom's taxonomy for deeper understanding, increased application, and potentially longer-term retention.

While the results of the current studies provide possible benefits of the *FP* strategy in real classroom settings, there were limitations usually associated with research conducted in this environment. First, we did not directly test traditional flashcard use against *FP* in a detailed manner because previous research has demonstrated that overly focusing on key terms, as fostered by traditional flashcard use, may hurt student grades (Gurung, 2004). Consequently, we did not include a comparison group forced to engage in such an activity. Instead, we opted for a traditional control group (no-treatment) and allowed students to engage in their normal studying behaviors, believing this to be the more appropriate ethical option.

Another limitation is the possible effect of extra credit; in Studies 1a and 1b, students received extra credit to complete the *FP* system, which may have increased the motivation to try the new method. We attempted to avoid this issue in Study 2 by randomly assigning specific classes to be exposed with the *FP* system without any incentives to change their study habits, and, as a result, the number of students who completed the *FP* system (at least explicitly) was smaller than those in Studies 1a and 1b. Part of this problem may be because students are often naturally resistant to changes in their study habits (Langley, 1993). Students may hear about new study strategies, and may even be hopeful and excited to use them. However, as the exam gets closer, anxiety, fear, procrastination, or perhaps stubbornness may influence students to revert back to old habits (Dembo & Seli, 2004). This would have a deleterious effect on the number of students who were willing to try or continue the strategy. In fact, in the self-report analysis students explicitly described re-reading notes, textbooks, or PowerPoint slides; the proportions were higher in the control group (46%) than the experimental group (28%). Future studies could be designed to provide a means of giving both conditions an incentive to study—that would also simultaneously avoid Hawthorne effects—in order to increase the number of students willing to try the new system.

Additional work on establishing the efficacy of even basic flashcard use is important, and future studies should directly test the effect of the traditional, student-created flashcard and the *FP* method. It would also help us better understand the mechanisms of learning by closely examining the use of *FP*. Although our current data suggest that the number of hours spent using *FP* did not predict exam performance, it is plausible that retrospective questions asked at the end of the semester may not have captured the exact studying hours. Finally, future studies may also examine the differential effects of exposure to the *FP* strategy specifically as compared to a general study skills lecture.

In conclusion, the *FP* method presents a straightforward way for students to improve scores on multiple-choice exams. The results of the research reported here provide evidence for the use of flashcards, a finding that when shared with students may make it more likely they will use it, and it also provides a more active study strategy than traditional flashcard use. *FP* arms instructors faced with the repeated question of "How should I study?" with a clear-cut set of answers. Although this method can be improved, our data provide a starting point and strong foundation for the design of future similar interventions in the classroom.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. We also conducted a 2 (Exam: Exam 1 vs. Exam 3) × 2 (*FP*: Users vs. Non-users) mixed-model analysis of covariance, using students' reported grade point average as a covariance, with Exam as a within-subject variable and *FP* as a between-subject variable. The interaction effect between Exam and *FP* remained significant, $F(1, 68) = 7.312$, $p = 0.009$, $\eta_p^2 = 0.10$.
2. We also conducted a 3 (Exam: Exam 1 vs. Exam 2 vs. Exam 3) × 2 (*FP*: Users vs. Non-users) mixed-model analysis of covariance, using students' reported GPA as a covariance, with Exam as a within-subject variable and *FP* as a between-subject variable. The interaction effect between Exam and *FP* remained significant, $F(3, 118) = 7.312$, $p = 0.009$, $\eta_p^2 = 0.10$.

References

Anderson, L. W. & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Boston, MA: Allyn & Bacon.

Appleby, D. C. (2008). A cognitive taxonomy of multiple-choice questions. In L. T. Benjamin (Ed.), *Favorite activities for the teaching of psychology* (pp. 119–123). Washington, DC: American Psychological Association.

Appleby, D. C. (2013). A flashcard strategy to help students prepare for three types of multiple-choice questions commonly found on introductory psychology tests. Society for the Teaching of Psychology's Office of Teaching Resources. Retrieved from http://www.teachpsych.org/Resources/Documents/otrp/resources/appleby13flashcard.pdf

Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching*. New York, NY: The Guilford Press.

Balch, W. R. (1998). Practice versus review exams and final exam performance. *Teaching of Psychology*, *25*(3), 181–185.

Bartolewski, B. L., & Gurung, R. A. R. (2014). Comparing the relationship of learning techniques and exam score. *Scholarship of Teaching and Learning in Psychology*, *1*(3), 219–228.

Biggs, J. B. (2003). *Teaching for quality learning at university*. Buckingham, UK: Open University Press.

Biglan, A., Ary, D., & Wagenaar, A. C. (2000). The value of interrupted time-series experiments for community intervention research. *Prevention Science: The Official Journal of the Society for Prevention Research*, *1*(1), 31–49.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*(37), 417–444.

Brown, P. C., Roediger, H. L., & McDaniel, M. A. (2014). *Make it stick: The science of successful learning*. Boston: MA: Harvard University Press.

Craik, F. I. M. (1979). Human memory. *Annual Review of Psychology*, *30*(1), 93–102.

Craik, F. I. M. (2002). Levels of processing: Past, present… and future? *Memory*, *10*(5–6), 305–318.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

Dembo, M. H., & Seli, H. P. (2004). Students' resistance to change in learning strategies courses. *Journal of Developmental Education*, *27*(3), 2–11.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58.

Entwistle, N. (2009). *Teaching for understanding at university: Deep approaches and distinctive ways of thinking*. New York, NY: Palgrave Macmillan.

Golding, J. M., Wasarhaley, N. E., & Fletcher, B. (2012). The use of flashcards in an introduction to psychology class. *Teaching of Psychology*, *39*(3), 199–202.

Gronlund, N. E. (1998). *Assessment of student achievement*. 6th edition. Boston, MA: Allyn & Bacon.

Gurung, R. A. R. (2004). Pedagogical aids: Learning enhancers or dangerous detours? *Teaching of Psychology*, *31*(3), 164–166.

Gurung, R. A. R., & McCann, L. (2012). Study techniques and teaching. In B. M. Schwartz, & R. A. R. Gurung (Eds.), *Evidence-based teaching in higher education* (pp. 96–116). Washington, DC: American Psychological Association.

Hackathorn, J., Cornell, K., Garczynski, A. M., Solomon, E. D., Blankmeyer, K., & Tennial, R. E. (2012). Examining exam reviews: A comparison of exam scores and students' attitudes. *Journal of the Scholarship of Teaching and Learning*, *12*(3), 78–87.

Hartwig, M., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement?. *Psychonomic Bulletin and Review*, *19*(1), 126–134.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York, NY: Routledge.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772–775.

Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory*, *17*(4), 471–479.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. Ross (Ed.), *The psychology of learning and motivation* (p. 14). New York, NY: Elsevier.

Karpicke, J. D., & Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials?. *Memory & Cognition*, *38*(1), 116–124.

Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology*, *55*(1), 5–11.
Kratochwill, T. R. (Ed.). (1978). *Single subject research: Strategies for evaluation change*. New York, NY: Academic Press.
Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 989–998.
Langley, D. (1993). Teaching new motor patterns—Overcoming student resistance toward change. *Journal of Physical Education, Recreation & Dance*, *64*(1), 27–32.
Paller, K. A., Kutas, M., & Mayes, A. R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Psychophysiology*, *67*(4), 360–371.
Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass.
Pressley, M., McDaniel, M. A., Turnure, J. E., Wood, E., & Ahmad, M. (1987). Generation and precision of elaboration: Effects on intentional and incidental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 291–300.
Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688.
Tagg, J. (2003). *The learning paradigm college*. Boston, MA: Anker.
Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Wissman, K., Rawson, L., & Pyc, M. (2012). How and when do students use flashcards? *Memory*, *20*(6), 568–579.

Author biographies

Sawa Senzaki is an assistant professor in Human Development and Psychology at the University of Wisconsin-Green Bay. She enjoys teaching courses related to cultural diversity in psychology and human development, as well as infancy and early childhood development. Her research focuses on cross-cultural examinations of social and cognitive development and cultural transmission from parents to children.

Jana Hackathorn is an associate professor in the Department of Psychology at Murray State University. She received her PhD in Experimental Social Psychology from Saint Louis University in 2011. She has two main research interests: close relationships; and the scholarship of teaching and learning. Her main social psychological research interests include the study of romantic relationship processes. Her pedagogical research predominantly focuses on teaching effectiveness, especially in the ''dreaded'' (e.g., statistics) courses.

Drew C. Appleby received his BA from Simpson College in 1969 and his PhD from Iowa State University in 1972. During his four-decade career, he chaired the Marian University Psychology Department, was the Director of Undergraduate Studies in the Indiana University–Purdue University Indianapolis (IUPUI) Psychology Department, and served as the Associate Dean of the IUPUI Honors College. He used the results of his research on teaching, learning, advising, and mentoring to create strategies that enable college students to adapt to their educational environment, acquire academic competence, set realistic goals, and achieve their career aspirations. He published over 200 books and articles (including *The Savvy Psychology Major*); made over 600 conference and other professional presentations (including 25 invited keynote addresses); received 44 institutional, regional, and national awards for teaching, advising, mentoring, and service; and was honored for his contributions to the science and profession of psychology by being named a Fellow of the American Psychological Association, the Midwestern Psychological Association, and as the 30th distinguished member of Psi Chi. His work with IUPUI's varsity athletes led him to be named "My Favorite Professor" by 71 student-athletes. Over 300 of his students have earned graduate degrees in a wide variety of academic and professional fields, and he was designated as a mentor by 777 IUPUI psychology majors, 222 of whom indicated he was their most influential mentor by selecting the following sentence to describe his impact: "This professor influenced the whole course of my life and his effect on me has been invaluable." Dr. Appleby retired from IUPUI with the rank of Professor Emeritus in 2011.

Regan A. R. Gurung is Ben J. and Joyce Professor of Psychology and Human Development at the University of Wisconsin-Green Bay. An active scholar of teaching and learning, his work focuses on how students study and strengthening the methodology of the Scholarship of Teaching and Learning.