

9-2019

Components of Fluency-Based Instruction in the College Classroom

Jennifer Kourassanis Velasquez

The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

 Part of the [Applied Behavior Analysis Commons](#), [Curriculum and Instruction Commons](#), and the [Higher Education Commons](#)

Recommended Citation

Kourassanis Velasquez, Jennifer, "Components of Fluency-Based Instruction in the College Classroom" (2019). *CUNY Academic Works*.

https://academicworks.cuny.edu/gc_etds/3452

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.

COMPONENTS OF FLUENCY-BASED INSTRUCTION IN THE COLLEGE CLASSROOM

by

JENNIFER KOURASSANIS-VELASQUEZ

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements of the degree of Doctor of Philosophy, The City University of New York

2019

© 2019

JENNIFER KOURASSANIS-VELASQUEZ

All Rights Reserved

Components of Fluency-Based Instruction in the College Classroom

by

Jennifer Kourassanis-Velasquez

This manuscript has been read and accepted for the Graduate Faculty in Psychology to satisfy the dissertation requirement for the degree of Doctor of Philosophy

Date

Dr. Emily A. Jones
Chair of Examining Committee

Date

Dr. Richard Bodnar
Executive Officer

Dr. Nancy Hemmes

Dr. Daniel Fienup

Dr. Barbara Bucklin

Dr. Carl Binder
Supervisory Committee

ABSTRACT

Components of Fluency-Based Instruction in the College Classroom

by

Jennifer Kourassanis-Velasquez

Advisor: Emily A. Jones

The current research regarding the use of fluency-based instruction (FBI) to teach academic skills suggests the addition of FBI to traditional instruction produces better learning outcomes than traditional instruction alone. However, there is a lack of comparative research of the addition of FBI to traditional instruction vs. traditional instruction alone on student performance outcomes with college students. The present study was composed of two experiments to examine the effects of the addition of a component of FBI using a modified SAFMEDS (Say All Fast Minute Every Day Shuffled) strategy to traditional instruction within the course's existing curriculum on quiz and exam scores for both introductory and advanced level undergraduate students across small and large class sizes. The findings are mixed, but generally suggest that the addition of components of FBI may produce better student performance outcomes than the traditional instruction alone. The majority of the students reported that they preferred using SAFMEDS to learn concepts and would use it to learn concepts in other courses.

Keywords: fluency-based instruction, college students, SAFMEDS, precision teaching, college instruction

ACKNOWLEDGEMENTS

I would like to express my gratitude to several people who contributed towards the completion of this project and to my success in graduate school. To my advisor and mentor, Dr. Emily A. Jones, you have guided me since my early days of undergraduate school and helped me achieve my goals. I am most grateful to you as I could always rely on your support through all the challenges I faced in graduate school. Thank you to Dr. Daniel Fienup who encouraged me to pursue graduate school and for your endless guidance, intuitive feedback, and support through all my projects. To Drs. Nancy Hemmes, Carl Binder, and Barbara Bucklin, thank you for your insightful feedback and support as committee members. I want to thank my colleagues, Sally Izquierdo and Ridda Sheikh, who collaborated and contributed to the success of this project. A special thanks to Theresa Fiani, who I could always rely on for her expertise and unwavering support. Lastly, an endless thank you to my husband, Robert Velasquez, who without, I would not of completed this journey. You remained my beacon throughout the years. You have always lifted me in ways that can not be described and found ways to keep things light no matter how heavy the task.

TABLE OF CONTENTS

Introduction	1
Experiment 1	6
Method	6
Participants	6
Setting and Materials.....	7
Experimental Design.....	7
Procedure.....	17
No-FBI.....	9
FBI.....	9
Dependent Variable and Interobserver Agreement.....	11
Procedural Integrity.....	12
Student Satisfaction Survey.....	12
Results and Discussion.....	12
Experiment 2	16
Method	17
Participants	17
Setting and Materials.....	17
Experimental Design.....	17
Procedure.....	18

No-FBI.....	18
FBI.....	18
Dependent Variable and Interobserver Agreement.....	19
Procedural Integrity.....	20
Student Satisfaction Survey.....	20
Results and Discussion.....	21
General Discussion	24
References	48

LIST OF TABLES

Table 1. Experiment 1 Course Content Reviewed per Quiz Across Conditions33

Table 2. Experiment 1 Student Satisfaction Survey34

Table 3. Experiment 2 Course Content Reviewed per Quiz Across Conditions35

Table 4. Experiment 2 Student Satisfaction Survey36

LIST OF FIGURES

Figure 1. Group mean quiz scores across FBI and no-FBI conditions in Experiment 1	37
Figure 2. Each student's mean quiz score for the FBI and no-FBI conditions in Experiment 1 ..	38
Figure 3. Each student's individual quiz score across quizzes for FBI and no-FBI conditions in Experiment 1	39
Figure 4. Group mean quiz scores across FBI and no-FBI conditions in Experiment 2	40
Figure 5. Each student's mean quiz score for the FBI and no-FBI conditions in Experiment 2 ..	41
Figure 6. Each student's individual quiz score across quizzes for FBI and no-FBI conditions in Experiment 2	42

LIST OF APPENDICES

Appendix A. Fluency Training Instructions	43
Appendix B. Instructor Procedural Integrity	47

Components of Fluency-Based Instruction in the College Classroom

With college enrollment on the rise, large instructional settings where there is high student attendance are more prevalent than smaller settings. These high attendance classes typically use lecture as the primary teaching strategy, which often leads to a passive learning environment. This type of learning environment may not be conducive to student learning and may result in poor student outcomes and low student experience satisfaction (Hanover Research, 2010). Dropout rates are still moderate, ranging from 20-40% even with the ever-growing emphasis on the necessity of college education (National Center for Education Statistics, 2018). This may suggest students are struggling to be successful with their academic careers. While there may be many obstacles to student success (e.g., financial issues or family and or work commitments), struggles in learning course material continue to be a prominent issue with college students today. This concern coupled with poor student outcomes and low retention rates suggests something needs to change with college instruction (Tuckman & Kennedy, 2011).

There are many different approaches to instruction that may improve college student performance outcomes, including active learning vs. traditional lecture formats (Freeman et al., 2014), small group instruction (Springer, Stanne, & Donovan, 1999), digital-based materials vs. paper-based materials (Beauregard & Basile, 2013), and Keller's personalized system of instruction to name a few (Kulik, Kulik, C., & Cohen, 1979). All of these approaches suggest to further improve student outcomes and, while these approaches use different strategies, they all seem to focus on increasing active learning in the classroom.

Another approach that may be promising to improve active learning is fluency-based instruction (FBI). FBI focuses on teaching both accurate and fluent responding. O.R. Lindsley was the first to apply this approach also referred to as "Precision Teaching" within education

(Binder & Watkins, 1990). During FBI, students are active learners who receive guidance about how to engage with instructional materials and they receive immediate performance feedback. Student performance then guides the instructor how to proceed with further instruction. When an instructor implements FBI, he/she incorporates timed practice drills into the existing curriculum and instructional materials. The instructor chooses how students will respond (referred to as a learning channel) during drills. Learning channels are defined by the topography of how the stimuli are presented and the expected student response. For example, if the target response is a see/say learning channel, the student would view materials (i.e., “see” the materials) that he/she would be expected to read and/or say something about (i.e., “say”). During timed practice drills, students are timed as they rapidly practice concepts, such as terms with corresponding definitions. The instructor sets a mastery criterion (referred to as frequency aims) for students to work towards achieving. Practice drill performance is recorded, and students are often expected to continue practice drills until the mastery criterion (frequency aim) is met. FBI incorporates all of these components (e.g., learning channels, timed practice drills, frequency aims, immediate performance feedback, performance data collection, and instructor use of student performance outcomes to guide further instruction) to improve student learning outcomes. However, there are select components of FBI that can be used with instruction as well.

A common component of FBI is using a see/say learning channel as done with SAFMEDS (Say All Fast Minute Every Day Shuffled). SAFMEDS consists of carefully designed flashcards that students use for practice drills during FBI (Graf & Auman, 2005; Meindl, Ivy, Miller, Neef, & Williamson, 2013). The flashcards present a term on one side with the corresponding definition on the other. Students are expected to read out loud the definition and state the corresponding term. A student flips the card over to determine if his/her answer is

correct/incorrect. This is done with a deck of flashcards in rapid succession while the student is being timed for a set duration (a “timing”). As the student flips through the cards in order, the student makes two piles (correct and incorrect) until either all the cards have been reviewed or time is up. The number of correct responses is tallied. Typically, this procedure is repeated a minimum of three consecutive times (shuffling the cards in a different order each time). The objective is for the student to increase the number correct and rate of responding with each successive timing. Practice drills typically occur daily and are repeated until a mastery criterion or some other termination criterion is met.

The proposed benefits of FBI are quicker rate of acquisition and greater retention of skills learned compared to traditional instruction alone. FBI focuses on response accuracy and speed (Binder, 1996). Using this combined measure of performance (fluency) during instruction is proposed to have benefits that exceed other traditional teaching methods such as those that focus predominately on accuracy (Binder, 1996; Lindsley, 1991; Quigley, Peterson, Frieder, & Peck, 2017). It is suggested that a learner must achieve accurate responding, followed by fluency in order to result in long-term retention of the skill (Kubina & Wolfe, 2005). Therefore, it would be logical that FBI is a potential solution to problems in college instruction; a solution that perhaps can easily and readily be accepted into existing curriculum.

FBI has been demonstrated to be an effective approach to improve a range of skills across a variety of populations (Beck 1979; Cohen, 2017; Lindsley, 1991; Kubina & Wolfe, 2005). However, only a few studies have examined FBI with college students (Bucklin, Dickinson, & Brethower, 2000; Cihon & Eshleman, 2012; Commons, Crone-Todd, & Chen, 2014; Fox & Ghezzi, 2003; Meindl et al., 2013). Of those few studies, Bucklin et al. (2000) compared the effects of fluency training vs. accuracy training (examining response rates, accuracy, and

retention). Students who were in the fluency training condition had higher response rates, better accuracy, and better retention rates as compared to those in the accuracy training condition. Cihon and Eshleman (2012) examined the effects of instructor vs. student created flashcards using a SAFMEDS approach on quiz scores with undergraduate and graduate students. The authors found that a combination of both instructor and student created flashcards improved student outcomes. Commons, Crone-Todd, and Chen (2014) examined SAFMEDS with a standard celeration chart to record student performance to teach increasingly more complex concepts relating to behavioral development. All participants demonstrated acquisition of the concepts taught in the workshop. Fox and Ghezzi (2003) compared the effects of FBI to other practice learning strategies using a computer-based application on exam scores with undergraduate students. Findings were inconclusive, but the authors believe the findings are useful to educators who look to improve learning strategies in the classroom. The study also highlighted some of the methodological challenges that researchers often face when examining this type of instruction in the college setting, such as determining the number of instructional items that should be selected, examining massed vs. distributed practice, incorporating the measurement of students' terminal response rates, and finding ways to measure for generalization and retention within the context of college instruction (Fox & Ghezzi, 2003). Lastly, Meindl et al. (2013) also used a SAFMEDS approach with a small sample of graduate students to examine generalization of rates of responding. Their findings indicated that using a multi-exemplar approach produced better generalization outcomes as compared to using the same stimuli during practice drills.

While most of the aforementioned studies contribute to demonstrating the positive effects of FBI instruction in improving student outcomes, none of them compare the addition of FBI to

traditional instruction vs. traditional instruction alone. Unfortunately, there are only a handful of studies that have examined this (e.g., Beverley, Hughes, & Hastings, 2009; & Isaacs, 1973). Therefore, it is still unclear if the use of FBI is a superior approach compared to traditional instruction alone in the college classroom. Both studies that compared the addition of FBI to traditional instruction alone revealed higher student performance outcomes in the FBI condition. In Isaacs (1973), hearing-impaired students in an introductory level psychology class (n=31) who received FBI scored better on their exams than students who did not receive FBI. In Beverley et al. (2009), students who fell at 50% or below their class standing in a large first year statistics class of 340 students (n=55) either received the addition of FBI or lecture only. Students who received the addition of FBI performed significantly better than those students who received traditional instruction only. Traditional instruction in both studies seemed to involve a lecture format with no active learning strategies, such as students being engaged in group activities. For the FBI conditions, the students in the Beverley et al. (2009) study engaged in three 1-minute practice drill timings daily. It was not clear in Isaacs (1973) the duration and frequency of the practice drills. For both studies, students engaged in practice drills until they met criterion or until the study ended. The Isaacs (1973) study did not incorporate lectures with the FBI group. However, for the Beverley et al. (2009) study, course content was presented at the same pace as the no-FBI group regardless of students' practice drill performances.

Implementing FBI in the college classroom may pose some practical challenges given the limited number of class meetings and the time constraints to adhere to a course curriculum. The ability for instructors to run practice drills with students directly, monitor student performance effectively, and ensure appropriate student use of FBI outside the classroom may also make

implementing FBI in the college classroom difficult. Such limitations could be a reason why very little research with FBI in the college classroom has been conducted.

The current study aimed to address logistical and practical concerns of FBI implementation while adhering to the standard college curriculum in a variety of instructional settings (advanced level in a small class setting vs. an introductory level in a large class setting). We examined whether the addition of components of FBI to traditional instruction improves student outcomes compared to traditional instruction alone. In a series of two experiments, we compared the effects of adding a modified SAFMEDS approach to more traditional instruction on student course performance. Other components of FBI used were a set mastery criterion and student self-recorded data during SAFMEDS practice sessions. Experiment 1 involved advanced level students in a small intensive class which already incorporated active learning strategies. In the first experiment, we alternated between the use of FBI and no-FBI across quiz content. Experiment 2 involved introductory level students in a typical large lecture hall class setting; we alternated between the use of FBI and no-FBI across exam content.

Experiment 1

Method

Participants

Forty-eight advanced level undergraduate psychology students (40 females and eight males) enrolled in a required writing-intensive experimental psychology course participated in this study. Half of the students were enrolled in a morning session class and the other half in an afternoon session class. Each session had a different instructor. All enrolled students participated, and none dropped out during the semester. The course met twice a week for 180 min each meeting across a 16-week semester.

Setting and Materials

Lectures and quizzes were held in a small classroom with approximately 24 desks and 24 individual computer stations. The instructors used a podium, computer, projector screen, and white board for lectures. In addition, students used personal computers, laptops, and smart devices. The textbook was *Research Methods* (9th ed., White & McBurney, 2013).

Instructors provided students all course content via Blackboard© (an online course management system), access to Quizlet© (an online student learning application), and instructions about how to implement FBI and use Quizlet© (see Appendix A). Six quizzes consisting of 10 multiple-choice questions (five questions per chapter) were administered throughout the semester as per the syllabus schedule.

Experimental Design

A within-groups ABABAB reversal design was employed. The instructor began the semester teaching concepts with the addition of SAFMEDS (a component of FBI) (A) and then switched to teaching concepts without (B). All students followed the same order of conditions alternating instruction with FBI and without FBI for every quiz. Six quizzes were administered throughout the semester with each quiz covering two chapters. Performance was evaluated with a quiz at the end of each of two chapters. Table 1 shows the content covered for each quiz and each condition. The content covered for Quiz 1 was taught with the addition of FBI, the content covered for Quiz 2 was taught without FBI. This alternated for the remaining content and quizzes for a total of six quizzes. For quizzes 1 and 3, all the questions on the quizzes covered content that had been covered in SAFMEDS. The content contained on Quiz 5 included many more terms and concepts than that for Quizzes 1 and 3, and 30% of the questions on Quiz 5 referred to content that had not been practiced in SAFMEDS.

Procedure

Both sections of the experimental psychology class were relatively small with only 24 students and already incorporated active learning strategies. The course covered 12 chapters relating to experimental design from the textbook (refer to Table 1). Each class meeting consisted of either a lecture reviewing a chapter in the textbook, followed by a laboratory component or a full laboratory for the entire class meeting. Laboratory classes involved either a group activity, independent work, or writing exercises. Some of the laboratory classes involved conducting literature reviews and drafting experimental papers. Group activities consisted of conducting experiments, peer review of lab papers, and a variety of active exercises practicing the material (e.g., writing operational definitions for target behaviors). A quiz was administered following every two chapters throughout the semester for a total of 6 quizzes. For example, on Monday, the instructor started a new chapter by presenting material from the textbook while also providing visual examples, such as videos. The students were expected to engage in class discussions throughout the lecture, and then the instructor asked the students to engage in a group activity to further practice the material reviewed that day. The following class, a new chapter was introduced along with other laboratory activities. A quiz was administered on the next class meeting.

Two exams were administered (midterm and final) during the term. During the class prior to exams, the instructor asked students to split into two teams for a trivia game. During the trivia game, the instructor asked students a series of questions relative to course content in preparation for the exam.

Instructors provided traditional instruction with the use of lecture format and group activities for all content and added FBI for content relating to Quizzes 1, 3, and 5 only (discussed shortly and see Appendix A).

No-FBI condition. The instructor presented course content in a lecture format using PowerPoint© and/or a white board to highlight main points of lectures. Following lectures for most class meetings, instructors directed students to engage in group activities, such as conducting experiments, proof reading each other's lab report drafts, and other activities related to concepts discussed. The instructor administered the quizzes on the class meeting following the last lecture for every two chapters that were reviewed. Typically, one chapter was reviewed per class meeting. The content of some chapters required two class meetings to review.

FBI condition. The same procedures were implemented for the FBI condition as the no-FBI condition with the exception of adding SAFMEDS. This resulted in less time for some of the activities in the no-FBI condition. For example, if the instructor typically provided students 1 hour for writing lab reports during class, less time was provided for writing on FBI class days to allow for FBI practice. FBI practice took approximately 25 minutes per class.

The first author, who also taught the morning section of the class, provided the FBI procedures to the other course instructor. Both instructors reviewed the procedures together for understanding. Instructors reviewed the syllabus and FBI procedures on the first day of class. Instructors asked students to prepare SAFMEDS materials in advance of the next class meeting for when FBI would be implemented.

FBI procedures consisted of log in instructions for Quizlet© and a data sheet. The first author selected 10 terms per chapter (two decks) for the FBI condition for a total of 20 terms per quiz. The instructors provided the terms and corresponding definitions to the students and

instructed them to make flashcards for the in class FBI procedure at home prior to the next class meeting. The same terms and corresponding definitions were also used for the Quizlet© in home FBI procedure. The instructor told students to begin with Deck 1 (terms from one chapter) and continue with that deck until meeting mastery criterion before moving on to Deck 2 (terms from second chapter) (mastery criterion to be discussed shortly).

Students engaged in SAFMEDS practice at the beginning of class followed by a lecture. During class, students used the paper flashcard SAFMEDS and then between classes students used Quizlet© for additional fluency practice.

For in class SAFMEDS, the instructors gave students 10 min at the beginning of each class to review the paper flashcards silently and independently. Following the review, students engaged in practice drills with a student partner. Each student completed three 1-minute timings while the student partner timed and observed his/her partner's performance and provided feedback if he/she placed a card in an incorrect pile. Students recorded the number correct for their own performance on their data sheet for each timing, circling the best performance out of the three timings. Students reversed roles once the first student completed all three timings.

For in home (Quizlet©) SAFMEDS, instructors asked students to join the computer-based online platform using the link provided. Quizlet© consisted of computer-based flashcards of the same terms and corresponding definitions as the paper flashcards students used during class. Students only completed Quizlet© drills on non-class days (including absences). If they were absent they were expected to complete Quizlet©. Students engaged in computer-based practice drills using the same procedures as with in class SAFMEDS, however, drills were completed independently.

Instructors were able to view Quizlet© activity levels for all students. To encourage use of Quizlet©, instructors informed students verbally and in writing that they monitored Quizlet© performance. Activity levels were periodically monitored throughout the semester. Instructors checked the day(s) students logged in and whether they completed the SAFMEDS exercise. However, instructors did not collect data on Quizlet© activity levels nor did they provide any consequences to students who did or did not complete the Quizlet© SAFMEDS as instructed.

Instructors directed students to complete practice drills daily. For class days, students were expected to complete SAFMEDS using the flashcards and for non-class days or if a student was absent, he/she was expected to complete SAFMEDS using Quizlet©. Students engaged in daily practice drills until either mastery criterion was met for all terms for both chapters or completion of the quiz for those chapters. The mastery criterion consisted of correct responding for all 10 terms for a chapter for at least two consecutive timings in the same day. Students replaced terms/definitions of one chapter with those of the second chapter if they met mastery and the quiz had not yet been completed. Students had an average of 7 days (two in class SAFMEDS and five Quizlet©) to meet the mastery criterion for both decks between quizzes.

Dependent Variable and Interobserver Agreement

The dependent variable was the percentage of correct responding for each of six quizzes. Each quiz consisted of 10 multiple-choice questions all related to the chosen SAFMEDS terms. The instructor calculated the percentage correct for each quiz. No score was provided for incomplete quizzes when a student was absent for data collection purposes, however, students received a 0% for grading purposes. Interobserver agreement (IOA) was scored for 30% of the quizzes by the first author and by a second independent grader. The graders independently

scored quizzes using a standardized answer key. IOA was calculated by dividing the number of agreements by the total number of questions and multiplying by 100. IOA was 100%.

Procedural Integrity

Instructors completed a self-report checklist of their accuracy following all procedures (see Appendix B) in order to promote procedural integrity. For the FBI condition, instructors checked off whether they: (1) provided FBI instructions and terms on Blackboard®; (2) provided instructions about how to access Quizlet® the first day of class; (3) gave students 10 minutes to independently review flashcard terms on FBI content related class days; (4) gave students time to complete practice drills in class on FBI class days. For the no-FBI condition, instructors check off whether they: (1) did not give students terms, instruct them to create flashcards, provide time for independent review of flashcards, or time to complete practice drills for no-FBI related concepts. Self-reported procedural integrity was 100%.

Student Satisfaction Survey

The instructor administered an anonymous 10-item student satisfaction survey by providing students with a survey monkey® link using the course Blackboard® (refer to Table 2). Students rated a series of statements on a Likert scale (1 = strongly disagree to 5 = strongly agree) about how important they felt academic success was and about their perceived effectiveness/usefulness of SAFMEDS. The survey was administered at the end of the semester prior to the final exam.

Results and Discussion

Several students missed at least one quiz. Little's Missing Completely at Random (MCAR) test indicated that data were missing at random, $\chi^2 = 48.85$, $df = 54$, $p = .673$. Therefore, we used multiple imputation ($n = 5$) to estimate missing values.

Figure 1 illustrates group mean performance for individual quizzes in the FBI and no-FBI conditions ($n = 48$). Visual analysis of the group means does not suggest a functional relation between the implementation of components of FBI and student quiz performance. However, a two-way repeated ANOVA revealed there was a significant main effect of FBI condition on quiz scores, with students performing better in the FBI condition ($M = 83.11$, $SEM = 1.11$) than the no-FBI condition ($M = 80.39$, $SEM = 1.43$), $F(1,47) = 5.38$, $p = .025$, $d = -.39$ (small effect). Performance for Quizzes 1 and 3 of the FBI condition was higher than Quizzes 4 and 6 in the no-FBI condition; performance on Quiz 3 in the FBI condition was higher than performance on Quiz 2 of the no-FBI condition, and performance on Quiz 5 of the FBI condition was lower than Quizzes 2, 4, and 6 of the no-FBI condition. Results revealed there was a significant main effect of the order of quizzes (time), with students performing significantly worse in the third set of quizzes for both FBI and no-FBI conditions ($M = 71.37$, $SEM = 1.69$) than the first and second sets of quizzes ($M = 87.75$, $SEM = 1.46$, $M = 86.14$, $SEM = 1.13$, respectively), $F(2,94) = 69.31$, $p = .000$, $\eta_p^2 = .596$ (large effect). Students performed significantly higher in the second set of quizzes than the third set of quizzes. Although students performed highest in the first set of quizzes there was no significant difference between the first and second set of quizzes.

In addition to two significant main effects, there was also a significant interaction. Students performed better on Quiz 3 in the FBI condition ($M = 95.79$, $SEM = 1.01$) than Quizzes 1 and 5 of the FBI condition ($M = 89.32$, $SEM = 1.24$, $M = 64.23$, $SD = 2.13$, respectively). In the no-FBI condition, students performed better on Quiz 2 ($M = 86.18$, $SEM = 2.11$) than Quizzes 4 and 6 ($M = 76.48$, $SEM = 2.00$, $M = 78.50$, $SEM = 2.06$, respectively), $F(2,94) = 54.43$, $p = .000$, $\eta_p^2 = .537$ (large effect).

Figure 2 illustrates mean quiz scores per condition (averaged across the three quizzes within condition) for each student. Overall, students' average performance was 82% across quizzes. An analysis of the mean quiz score difference between conditions revealed that 56% of the students had higher quiz scores in the FBI condition than in the no-FBI condition. For students who performed better in the FBI condition, mean quiz scores across quizzes were 3-21% higher than in the no-FBI condition. Figure 3 illustrates quiz scores for each of the six quizzes for each student. Performance illustrated is consistent with Figure 2. Across all quizzes 54% of the students scored the lowest on Quiz 5, which was also in the FBI condition.

An analysis of the mean quiz scores to determine if the students quiz scores improved by at least one letter grade between conditions was conducted. The criterion of an improvement in at least one letter grade was determined if the quiz score increased one whole grade classification (e.g., an increase in quiz score from B- to A- or B+ to A-). Improvements within the same grade classification such as an increase from a B- to B+ did not meet criterion as an improvement by at least one letter grade. Across all quizzes, the letter grade analysis revealed that 40% of the whole sample of students improved their quiz scores by at least one letter grade in the FBI condition compared to only 15% of students who improved their quiz scores by at least one letter grade in the no-FBI condition.

Table 2 shows student satisfaction with using SAFMEDS. Students indicated strong agreement with statements on the importance of achieving good academic performance and the positive effects of using SAFMEDS. Students also indicated a stronger preference for the use of flashcards over Quizlet©.

Group mean quiz scores were higher in the FBI condition compared to the no-FBI condition, however, a visual analysis of a functional relation is not clear. Although a functional relation is not suggested, the students did indicate they liked using SAFMEDS as a learning tool.

Although students reported satisfaction, the mixed findings on student performance outcomes still raise a number of questions about the effects of using components of FBI compared to traditional instruction alone on learning outcomes. There are several reasons why the visual analysis is not consistent with the statistical analysis. First, scores on Quiz 5 in the FBI condition were lower than all the scores across conditions. A contributing factor for the low performance may have been that the level of difficulty of the course content for each quiz was not controlled across quizzes. It may be that the content covered for Quiz 5 was more difficult compared to the content covered in the other quizzes. We controlled for the number of terms used across quizzes during SAFMEDS practice, limiting the number of terms to be used from the larger proportion of terms covered in Quiz 5. Thus, students practiced a smaller proportion of the terms for the content covered in Quiz 5 compared to the other quizzes. If FBI were used in a typical classroom setting outside of experimental conditions, a less controlled procedure would allow for the number of terms used during SAFMEDS practice to be proportionate to the amount of content covered in a given quiz and perhaps would help to improve performance.

Other factors which may have mitigated against demonstration of a functional relation include ceiling effects, lack of student procedural integrity to ensure SAFMEDS were only being used for the FBI condition, the limited number of practice opportunities, and the classroom/instructional setting. Individual and group mean quiz scores show that students performed rather well across quizzes overall suggesting a potential ceiling effect. The student population in this course tends to be further along in their academic career, which may indicate

the students had experience with psychology concepts as well as how to study and thus may explain the minimal distinction in quiz scores between conditions for many students. Another factor for the minimally observed differences between scores is the failure to measure student's use of SAFMEDS. It is probable that students used SAFMEDS or a similar approach for the no-FBI quiz content or perhaps the students did not need to use flashcards to learn new concepts, but rather learned how to study using some form of rote practice for the no-FBI condition. Mastery criterion may have also influenced performance. Since content for each quiz was covered over two class days before the instructor administered the quiz, this limited the number of opportunities for students to practice terms before each quiz. Allowing for more opportunities to practice may improve performance outcomes. Lastly, the small class and intensive instruction may be conducive to better student outcomes than larger classes that use solely passive learning strategies such as lecture. FBI is hypothesized to help improve student outcomes in instructional settings where there is a lack of active learning and minimal direct support and interaction from instructors. Both the lack of control over the use of SAFMEDS and type of instruction could result in higher performance outcomes for the no-FBI condition in this type of class than in a larger class setting.

Experiment 2

To address some of the factors that may have affected demonstrating a functional relation between quiz performance and components of FBI in Experiment 1, we conducted Experiment 2 in a large introductory level passive learning environment. This type of class was selected to reduce the potential of ceiling effects since students were more likely to be just beginning their academic careers, thus having less experience with psychology concepts and study strategies. The class was also larger in a passive learning environment with the instructor teaching with the

use of traditional lecture and no other student and/or instructor engagement. Based on students' feedback on the social validity questionnaire from Experiment 1 about their preference of SAFMEDS stimuli presentation, we removed Quizlet© and instead used only paper flashcards. The mastery criterion was also increased from 10 correct responses across two consecutive timings in the same day to 10 correct responses across two consecutive timings across two consecutive days to hopefully further improve student outcomes in the FBI condition.

Method

Participants

Seventy-five undergraduate psychology students (49 females and 26 males) enrolled in an introductory level human motivation psychology course participated in this study. All enrolled students participated, and none dropped out during the semester. The course met twice a week in the afternoons for 75 min each meeting across a 16-week semester.

Setting and Materials

Lectures and exams were held in a large lecture hall with approximately 100 desks. The instructors used a podium, computer, projector screen, and white board for lectures. In addition, students used personal computers, laptops, and smart devices. The textbook was *Motivation: Theory, Research, and Application* (6th ed., Petri & Govern, 2013). Students completed exams on scantron forms.

The instructor provided students with all course content via Blackboard© and instructions on how to implement FBI as with Experiment 1. Four exams consisting of approximately 40 multiple-choice questions and five short answer questions were administered throughout the semester as per the syllabus schedule.

Experimental Design

A within-groups ABAB reversal design was employed. The instructor began the semester teaching concepts without FBI (A) followed by an exam and then switched to teaching concepts with SAFMEDS (a component of FBI) (B) followed by an exam; etc. All students followed the same order of conditions alternating instruction without FBI and with FBI for content covered for each exam. The content covered for Exam 1 was taught without the addition of FBI, the content covered for Exam 2 was taught with FBI. This alternated for the remaining content and exams for a total of four exams.

Procedure

Similar procedures as with Experiment 1 were employed with slight variations. Teaching in this class relied solely on lecture; there were no group activities. The course covered 16 chapters relating to human motivation from the textbook (refer to Table 3). An exam was administered following approximately every four chapters, for a total of four exams. The instructor provided traditional instruction in the form of lectures for all content. She added FBI for content relating to Exams 2 and 4 only. The instructor administered the exams on the class meeting following the last lecture for approximately every four chapters that were reviewed. Typically, one chapter was reviewed per every two to three class meetings.

No-FBI condition. The instructor presented course content in a lecture format using PowerPoint© and/or the white board to highlight main points of lectures.

FBI condition. The same procedures were implemented for the FBI condition as the no-FBI condition with the exception of adding SAFMEDS. The instructor provided the first author with an electronic copy of the exams and syllabus. The first author selected the terms to be used for SAFMEDS for Exams 2 and 4 (88% and 81% of the exam content, respectively, was covered by terms in SAFMEDS). A total of 20 terms (10 terms per deck) were selected for each exam

(the number of terms per chapter varied). A deck consisted of 10 terms/definitions chosen from various chapters covered under each exam. Students were permitted to work with one deck at a time. The first author provided the FBI procedures to the course instructor and they reviewed them together for understanding. The instructor reviewed the FBI procedures with the class prior to teaching content using SAFMEDS. The instructor informed students to use SAFMEDS for content relating to Exams 2 and 4 only.

Students implemented SAFMEDS with the use of flashcards (unlike Experiment 1, Quizlet© was not used). As with Experiment 1, instructors directed students to complete practice drills daily including on non-class days and when a student was absent.

The only other variation from Experiment 1 was the mastery criterion. Mastery criterion consisted of correct responding for all 10 terms in a deck for at least two consecutive timings across two consecutive days. Students replaced terms/definitions of Deck 1 with those of Deck 2 if they met mastery before the exam was administered. Students had an average of 16 days (an average of five class meetings) to complete mastery criterion for both decks prior to exam day.

Dependent Variable and Interobserver Agreement

The dependent variable was the percentage of correct responding on each exam. Each exam consisted of approximately 40 multiple-choice and five short answer questions related to the chosen SAFMEDS terms.

The instructor calculated the percentage correct for each exam. The instructor calculated the students' final grades by averaging the three highest exam scores (dropping the lowest graded exam). If students were satisfied with their average of the first three exams, they had the option to opt out of the final exam. No score was provided for incomplete exams for data collection purposes, however, students received a 0% for grading purposes. Interobserver agreement (IOA)

was scored for 30% of the exams by the course instructor and the first author. The graders independently scored exams using a standardized answer key. IOA was calculated by dividing the number of agreements by the total number of questions and multiplying by 100. IOA was 98%

Procedural Integrity

The instructor completed a self-report checklist of her accuracy following all procedures in order to promote procedural integrity. For the FBI conditions, the instructor checked off whether she: (1) provided FBI instructions and terms on Blackboard©; (2) gave students 10 minutes to independently review flashcard terms on FBI content related class days; (3) gave students time to complete practice drills in class on FBI class days. For the no-FBI condition, the instructor checked off whether she: (1) did not give students terms, instruct them to create flashcards, provide time for independent review of flashcards, or time to complete practice drills for no-FBI related concepts. Self-reported procedural integrity was 100%.

Student Satisfaction Survey

The instructor administered an anonymous 10-item student satisfaction survey by providing students with a survey monkey® link using Blackboard© (refer to Table 4). Students rated a series of statements on a Likert scale (1 = strongly disagree to 5 = strongly agree) about how important they felt academic success was and about their perceived effectiveness/usefulness of SAFMEDS. The survey was administered at the end of the semester prior to the final exam. On the day of the final exam, students were asked to self-report on the exam if they used SAFMEDS for no-FBI related exam content.

Results and Discussion

Several students missed at least one exam. Little's Missing Completely at Random (MCAR) test was significant and indicated that data were not missing at random, $\chi^2 = 31.09$, $df = 14$, $p = .005$. Therefore, a multiple imputation to estimate missing values was not used. Data were excluded for 33 students who missed/opted out of any exams ($n = 42$ included in analysis out of 75 students). Of those excluded, seven students missed Exam 1, two missed Exam 2, four missed Exam 3, and 20 opted out of Exam 4.

Figure 4 illustrates group mean performance for each exam in the no-FBI and FBI conditions. Visual analysis of the group means does not suggest a functional relation between the implementation of components of FBI and student outcomes. However, a two-way repeated ANOVA revealed there was a significant main effect of FBI condition on exam scores, with students performing better in the FBI condition ($M = 89.94$, $SEM = .73$) than the no-FBI condition ($M = 84.04$, $SEM = .95$), $F(1,41) = 53.94$, $p = .000$, $d = 1.03$ (large effect). Performance for Exams 1 and 3 in the no-FBI condition was lower than Exams 2 and 4 in the FBI condition.

Results revealed there was a significant main effect of the order of exams (time), with students performing significantly better in the second set of exams for both FBI and no-FBI conditions ($M = 90.97$, $SEM = .70$) than the first set of exams ($M = 83.01$, $SEM = 1.09$), $F(1,41) = 56.49$, $p = .000$, $\eta_p^2 = .579$ (large effect).

In addition to two significant main effects, there was also a significant interaction. Students performed better on Exam 3 in the no-FBI condition ($M = 91.07$, $SEM = .93$) than Exam 1 ($M = 77.02$, $SEM = 1.52$). In the FBI condition, students performed better on Exam 2 (M

= 90.87, $SEM = .853$) than Exam 4 ($M = 89.00$, $SEM = .853$), $F(2,94) = 59.42$, $p = .000$, $\eta_p^2 = .592$ (large effect).

Figure 5 illustrates mean exam scores per condition (averaged across the two exams within condition) for each student. An analysis of the mean exam score difference between conditions revealed that 88% of the students had higher exams scores in the FBI condition than in the no-FBI condition. For students who performed better in the FBI condition, an increase in mean exam scores were 1- 23% higher than in the no-FBI condition. Figure 6 illustrates individual exam scores for each of the four exams for each student. Analysis is consistent with Figure 5, but now we are able to see a visual analysis of the individual exam scores, which does not suggest a clear functional relation between the implementation of FBI and student outcomes, however, for some individual students a functional relation is evident. Performance differences are clear between Exams 1 and 2, but less evident with Exams 3 and 4. Fifty percent of the students scored the highest on Exam 4 of the FBI condition and 79% of the students scored the lowest on Exam 1 of the no-FBI condition across all exams.

An analysis of the mean exam scores to determine if the students' exam scores improved by at least one letter grade between conditions was conducted. The criterion of an improvement in at least one letter grade was determined if the exam score increased from one whole grade classification (e.g., an increased in exam score from B- to A- or B+ to A-). Improvements within the same grade classification such as an increase from a B- to B+ did not meet criterion as an improvement by at least one letter grade. Across all exams, 55% of the whole student sample improved their exam scores by at least one letter grade in the FBI condition compared to only 2% of the students who improved their exam scores by at least one letter grade in the no-FBI condition.

Table 4 shows student satisfaction with SAFMEDS instruction. Overall, student indicated strong agreement with statements on the importance of academic success and perceived SAFMEDS to be useful and effective in improving their learning outcomes. Seventeen students reported they used SAFMEDS for no-FBI related content. However, only 25 students responded, which may reflect a biased sample.

Group mean exam scores were higher in the FBI condition compared to the no-FBI condition. Similar to Experiment 1, there is a difference in conclusions based upon tests of statistical significance and traditional visual analysis of single subject design data. There looks to be an effect of components of FBI based on the statistical analysis, but it appears to be largely driven by the first exam. Since some students reported using SAFMEDS for non-FBI related concepts, it is possible that students performed better than they would have for Exam 3 in the no-FBI condition, which may support why a visual analysis of group means comparing Exams 2 and 4 in the FBI condition to Exam 3 is inconsistent with the statistical analyses.

Students who performed well in the first three exams, opted out of the final exam and could not be included in the analyses. This resulted in a select sample of poorer performing students, which may have contributed to the significant difference of Exam 1 compared to the other exams. It is feasible that students are more apt to improve performance on exams across time (practice effects) and thus could be a factor in the higher performance observed across the remaining exams. Practice effects may also explain why a functional relation between components of FBI and student outcomes for the group means is not evident. Overall, the use of components of FBI was observed to have similar effects as in Experiment 1 with a larger class size with introductory level students. These results further suggest that the addition of

components of FBI may potentially improve student performance outcomes regardless of class size (where there is little control of student procedural integrity) or student learner level.

General Discussion

In a series of two experiments, we examined the effects of components of FBI on college student performance in a small advanced level class and a larger introductory lecture class. Findings are inconclusive, but may suggest there are some positive effects of using components of FBI that warrant further investigation.

Distinctions Between Experiments

This is one of the few studies to report the effects of using components of FBI with the general college population, with both introductory or advanced level students in classes utilizing different types of instructional styles (passive learning with lecture format only and active learning with lecture format). Had they been more robust, the generality of these findings might suggest that components of FBI can help to improve even the most advanced learner already receiving adequate instructional support, but can also help to improve student outcomes for introductory level students in settings where students receive minimal instructional support.

A smaller effect size was observed with the two smaller classes' whose data were combined compared to the larger class indicating the possibility of a ceiling effect perhaps because of the advanced level of the students and the use of active learning in the class instruction. A larger effect size may have been observed with the larger class since typically introductory level students have less college education experience and more room for improvement as compared to advanced level students. There were slight procedural differences between Experiment 1 and Experiment 2, such as the removal of Quizlet© and the use of only paper flashcards in Experiment 2. Although a direct comparison of stimulus presentation

parameters was not examined in the present study, the existing literature suggests differences in performance outcomes may depend upon the presentation of the practice stimuli (digital vs. paper flashcards) (Dizon & Tang, 2017). Yet, consistent findings of which method produces better outcomes appears to remain unanswered and should be further explored.

We also increased mastery criterion from responding correctly for all 10 terms for at least two consecutive timings in the same day to correct responding for all 10 terms for at least two consecutive timings across two consecutive days. There were no quizzes in Experiment 2, rather only exams. This increased the number of practice drill opportunities between exams, allowing students more opportunity to achieve mastery criterion prior to the exams. One of the practical issues instructors faced was the limited number of practice opportunities to adhere to the course timeline. Thus, finding a feasible strategy to allow students more practice opportunities could improve learning outcomes.

Feasibility

A handful of studies suggest the addition of FBI to traditional instruction may improve learning outcomes with college students even when faced with logistical challenges such as the infrequent class meetings and lack of teacher/student interactions (Cihon & Eshleman, 2012; Commons, Crone-Todd, & Chen, 2014; Fox & Ghezzi, 2003; Meindl et al., 2013). Despite limitations (e.g., being restricted to the number of practice opportunities and SAFMEDS cards provided per quiz or exam content), we did develop a practical approach to incorporate components of FBI into college instruction. We used a modified version of SAFMEDS and did not require that students meet mastery criterion before taking the quizzes or exams. Students self-recorded their data so instructors did not monitor students' progress. Unlike typical FBI procedures, instructors did not use FBI practice drill outcomes to dictate the pace of instruction.

It remains unclear whether mastery criterion is even necessary or is it simply some number of practice opportunities that affects performance outcomes. Future research should examine these parameters specific to college instruction.

Other studies of FBI in college instruction typically use student managers or have additional class meetings to assist with FBI implementation and improve procedural integrity (Beverley et al., 2009; & Isaacs, 1973). While the present study did not provide additional support for FBI implementation, the addition of components of FBI appeared to still improve student outcomes compared to traditional instruction alone in a setting with minimal control of student procedural integrity. These findings address some of the practical concerns with FBI in college classrooms, such as implementing FBI without additional resources or the necessity to revise current curriculum.

Findings and Future Direction

While statistically significant findings were revealed for both experiments, a visual analysis indicated contradictory findings as the typical pattern of a reversal design was not observed for either experiment. The statistically significant differences in Experiment 2 were largely driven by the first exam. There were also a handful of students whose performance patterns did suggest a functional relation between FBI and student performance outcomes. This may suggest that FBI is worthy of further exploration.

Other than the statically significant differences being driven by the first exam in Experiment 2, there are several other potential reasons for the contradictory findings that if addressed may provide more conclusive evidence to the effects of components of FBI. First, it is necessary to control for the level of difficulty of the course material covered on each quiz/exam across conditions that may influence student performance. In addition, the number of SAFMEDS

terms practiced for each quiz/exam was held constant, but may not reflect the number of terms relevant for that section of course material. The number of terms practiced should reflect the course content to allow for students to adequately demonstrate acquisition with FBI. Second, many students missed quizzes/exams. In fact, in Experiment 2 more than half of the students were excluded from the analyses due to missing an exam or opting out of the final exam if they were satisfied with their performance on the first three exams. The sample for Experiment 2 may have been skewed to lower performing students, presumably because the higher performing students opted out of Exam 4. Future experiments need to replicate Experiment 2 without the opt out option. Attendance was also an issue for both experiments, a factor that increased the number of missed exams and missing data. There should be an attendance contingency in place to encourage attendance, which would decrease missing data and ensure a representative sample of students. Then if there are data missing at random it could be corrected with the use of MCAR as with Experiment 1. This would allow one to examine a larger student sample across a broader range of performance levels to determine the effects of FBI. Third, the use of a within-subject groups design posed some potential issues. There may have been order and practice effects. Typically, students improve in quiz and exam performances as the semester progresses having been exposed to the testing conditions on multiple occasions. Splitting the student sample in half and reverse counterbalancing the order of conditions or using a between-subject groups design may limit order and practice effects. There may also have been carryover effects. One of the proposed benefits of FBI is that training a skill to fluency will not only impact the acquisition of initial skills, but will increase the rate of learning for more complex skills as well (Binder, 1996; Johnson & Layng, 1992). Students may have acquired the concepts quicker for material that did not receive the addition of SAFMEDS simply from their use of SAFMEDS with other concepts.

This could further explain the lack of a functional relation observed with a reversal design. Again, using a different design may help to eliminate this confound. Lastly, we could not control for students' using SAFMEDS for concepts they were not instructed to do so. The instructor for Experiment 2 surveyed students to identify how many of them used SAFMEDS for no-FBI content. Several students self-reported that they did use SAFMEDS for no-FBI content, however, the self-report was limited as only a small proportion of students responded. A between-subject design would eliminate this issue as well. While from an experimental control perspective the use of FBI across conditions could be problematic, it is an overall positive issue to have, as it suggests FBI is a strategy students find helpful and will use to achieve academic success across courses, which essentially is the overarching aim of this area of research.

There were other limitations in this study that if addressed would help to improve the validity and reliability of FBI research findings. First, instructors did not have the resources to obtain student procedural integrity. The students were asked to practice SAFMEDS at home and this was not something that could be monitored effectively. A potential solution to monitor in class student procedural integrity would be to have students rate each other's procedural integrity, but this would not be without its challenges. Second, the first author obtained instructor procedural integrity, but only through instructor self-report, which is not the most objective and reliable approach. The lack of procedural integrity measures seems to be a prevalent limitation in the FBI with college student literature and certainly needs to be addressed. Meindl et al.'s (2013) study was one of the few studies of FBI with college student to measure both IOA and procedural integrity. The authors used an independent observer to assess both measures. Thus, it is possible to address this limitation.

As many of the limitations of this study suggest, further examination of the most appropriate experimental design and method of data analysis for FBI in college instruction is needed. Perhaps a single subject design is not the most accurate representation of student performance outcomes for several of the reasons discussed including issues such as carryover effects.

In this study, instructors integrated SAFMEDS into college instruction. Introducing this one component of FBI seemed feasible. But there are other aspects of FBI that could be incorporated into the college classroom, though they may present more significant feasibility and logistical obstacles. First, frequency aims (fluent responding) should be set for SAFMEDS practice drills. The instructor selects a set number of words per minute (the aim) that students should work towards achieving during practice drills. The frequency aim is often set based upon performance of those who are knowledgeable in the targeted concepts. To determine the appropriate frequency aim, experts in the concepts being taught conduct a few repeated timings to obtain the range of count per minute performance that becomes the frequency aim for students.

Second, we did not reinforce student's individual performance, but it is a critical part of FBI. Instructors should reinforce each student's personal best performance during practice drills. Instructors can provide verbal praise such as making a class announcement for students who performed their personal best that day or perhaps provide incremental bonus points.

Third, In the current study we limited practice to 10 terms in a deck per practice drill opportunity. This may have limited the maximum number of terms a student could increase his/her fluency to and thus possibly resulting in a measurement ceiling. Students should be given more opportunities to respond during practice drills than they can complete during a timing. For

example, rather than only use the 10 terms, the instructor can provide all the terms for the semester in the deck to be used during practice drills.

Fourth, In Experiment 1, we used card stock paper flashcards for in-class practice drills and a computer-based program for in-home practice drills where students had to read out loud both the terms and definitions. Sometimes technology, the mechanics of handling the type of materials used, and the selected learning channel could actually slow down the students' rate of responding. We also did not control for the number of words of the terms and definitions used. Too many words used during SAFMEDS could also affect rate of responding. Instructors should keep these factors into consideration when developing materials and procedures. Improved technology/materials or alterations to the way students are expected to respond may help. For example, removing the requirement to read the term out loud and rather have the student sight read the term and then say out loud the corresponding definition may help minimize affecting the student's rate of responding.

Lastly, we had students self-record their own performance, but did not use their data to examine SAFMEDS performance against quiz or exam outcomes. Instructors should either collect student self-recorded data and/or record each student's count per minute data and examine if student performance during practice drills correlates with test scores. One practical way to incorporate this is for instructors to collect pre and post-test count per minute measures. The instructor could sit one-on-one with each student and conduct the practice drills during the pre and post-test measures. Another suggestion would be for the instructor to periodically (can be done systematically or at random) run drills with each student throughout the semester.

Future studies should also further explore with whom FBI is effective or not with college instruction. There are limited studies with the college student population and even less that have

examined an FBI approach across a variety of college curricula. The majority of FBI studies in college instruction focus on psychology concepts, as in this study. Future studies should examine the effects of FBI in other disciplines as done with younger students (Beck 1979; Cohen, 2017; Lindsley, 1991; Kubina & Wolfe, 2005). There are also several other parameters of FBI that could be examined: mastery criterion, timing duration, practice drill frequency, class meeting frequency, stimuli presentation, a stronger more reliable measure of instructor procedural integrity, and how to obtain student procedural integrity.

The implications of these findings could help to improve student enrollment retention, as well as, retention of concepts learned that go beyond one semester or one class. If students can maintain what is learned in one class, this can improve performance outcomes across other classes. This may also improve maintaining the knowledge obtained in college beyond graduation and help to generalize those skills learned to the workplace. Student enrollment retention and generalization of acquired skills are important variables to examine in future research.

Conclusion

This is one of the few studies to examine components of FBI in the college classroom using a within group reversal design. Both experiments revealed mixed outcomes. While we observed statistically significant findings and modest effect sizes, a visual analysis of performance across conditions did not indicate a clear functional relation between the use of FBI and improved student outcomes. While these findings are not conclusive, they do offer exploratory guidance and suggest that the addition of components FBI to traditional college instruction may improve student performance as compared to traditional instruction alone.

This study provides a practical modified approach to adding components of FBI to college teaching that allows for instructors to easily incorporate into their existing curriculum. Having introduced different forms of SAFMEDS stimuli, class sizes, level of instructional support, and student learning level, this further illustrates the generality of the use of FBI and its positive learning outcomes with the general college population.

This study contributes to the literature examining FBI with college students and raises many new questions that warrant further examination (Beverley et al., 2009; & Isaacs, 1973). The authors hope that these findings will encourage more robust examination of FBI. Future studies should continue to explore the efficacy and the various parameters and components of FBI (e.g., the number of practice opportunities) in college instruction, while also addressing some of the limitations and practical issues faced with implementing FBI in a college setting. Incorporating all the components of FBI in to college instruction may be difficult, but continued research may suggest how to do so and what components are necessary and sufficient to result in improved student performance outcomes. Examination of multiple components of FBI and the use of a between subject design may produce better outcomes. If we are able to demonstrate the positive effects of FBI without disrupting the existing curriculum, it is possible that institutions will begin to adopt FBI into their teaching practices.

Table 1

Experiment 1 Course Content Reviewed per Quiz Across Conditions

Quiz	Content	Condition
1	Psychology and Science Ethics	FBI
2	Writing in Psychology Variables	No-FBI
3	Tables and Graphs Statistics	FBI
4	Validity Control	No-FBI
5	Single Factor Designs	FBI
6	Observational, Archival, Case Studies, and Surveys Factorial Designs	No-FBI

Table 2

Experiment 1 Student Satisfaction Survey

	Question	Distribution of Ratings 1-2-3-4-5	<i>M</i>
1	I feel it is important to achieve good academic performance	0-0-2-10-28	4.65
2	I feel I need strategies to improve academic performance	0-2-4-19-15	4.17
3	I feel it is important to learn strategies to improve academic performance	0-0-1-13-26	4.63
4	I enjoyed learning to use FBI to learn new concepts	1-0-3-25-11	4.13
5	FBI helped me learn new concepts	0-2-3-23-12	4.13
6	FBI strategies were easy to follow	1-1-6-14-18	4.17
7	I would use FBI for other classes	1-2-8-17-12	3.92
8	I feel FBI improved my academic performance for this class	0-2-2-21-15	4.22
9	How well did you follow the FBI?	*1-8-24-7	2.92
10	Rate your preference for flashcards versus Quizlet© for FBI	*5-5-8-13-9	3.40

Note. Students rated agreement for each item on a scale from 1 (strongly disagree) to 5 (strongly agree). Items marked with (*) used a different scale. Q9 used a scale from 1 (poorly) to 4 (perfectly) and Q10 used a scale from 1 (prefer Quizlet©) to 5 (prefer flashcards).

Table 3

Experiment 2 Course Content Reviewed per Exam Across Conditions

Exam	Content	Condition
1	Human Motivation Overview Evolutionary Psychology Genetic Contributions to Motivated Behavior Learning	No-FBI
2	Aggression Eating	FBI
3	Behavioral Economics Delay Discounting Sensory Stimulation Drug Addition Sexual Motivation and Mate Selection	No-FBI
4	Cognitive Aspects Cognitive Dissonance Conformity and Obedience Attribution Approaches Competence and Control	FBI

Table 4

Experiment 2 Student Satisfaction Survey

	Question	Distribution of Ratings 1-2-3-4-5	<i>M</i>
1	I feel it is important to achieve good academic performance	0-0-1-4-20	4.76
2	I feel I need strategies to improve academic performance	0-0-3-13-9	4.24
3	I feel it is important to learn strategies to improve academic performance	0-0-0-12-13	4.52
4	I enjoyed learning to use FBI to learn new concepts	1-1-1-11-11	4.20
5	FBI helped me learn new concepts	1-1-1-14-8	4.08
6	FBI strategies were easy to follow	1-0-3-10-11	4.20
7	I would use FBI for other classes	1-0-4-10-10	4.12
8	I feel FBI improved my academic performance for this class	1-1-2-11-10	4.12
9	How well did you follow the FBI?	*1-3-11-10	3.20
10	I used FBI for my other classes	*11-11-2-1	1.72

Note. Students rated agreement for each item on a scale from 1 (strongly disagree) to 5 (strongly agree). Items marked with (*) used a different scale. Q9 used a scale from 1 (poorly) to 4 (perfectly) and Q10 used a scale from 1 (I did not use it all all) to 4 (I used it all the time).

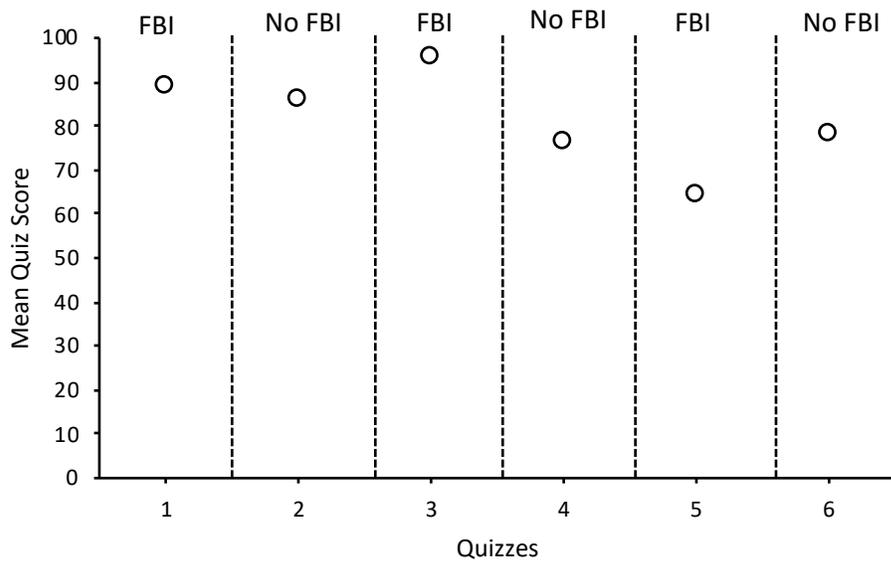


Figure 1. Group mean quiz scores across FBI and no-FBI conditions in Experiment 1.

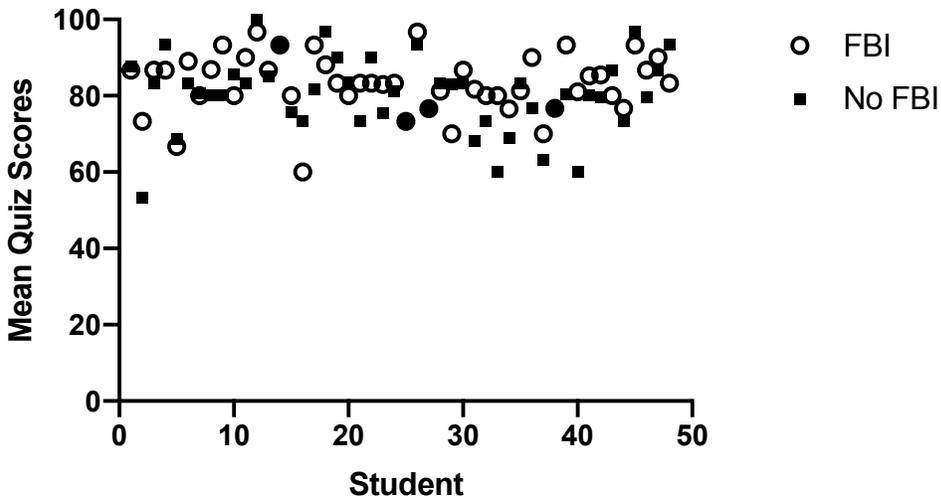


Figure 2. Each student's mean quiz score for the FBI and no-FBI conditions in Experiment 1.

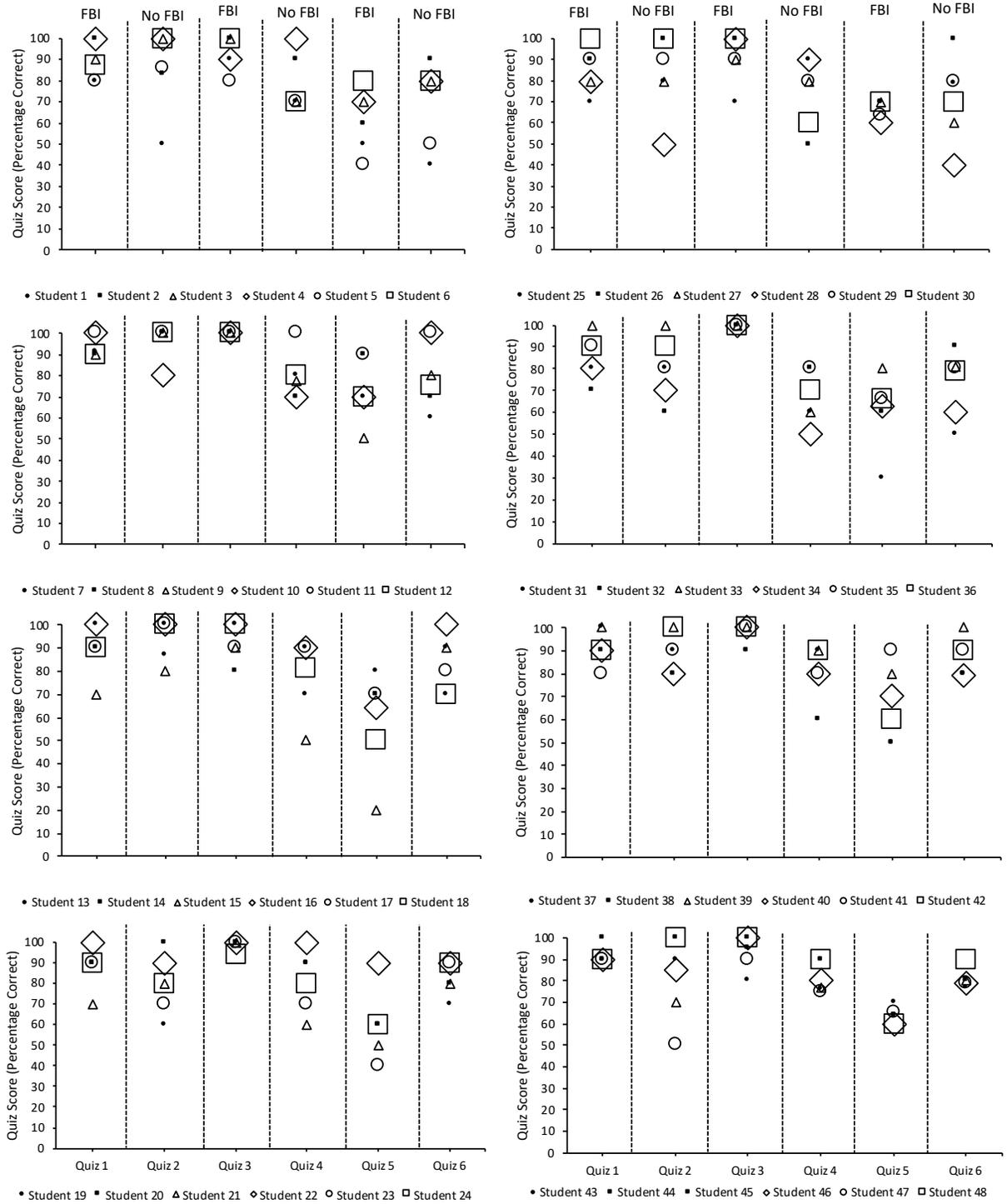


Figure 3. Each student's individual quiz score across quizzes for FBI and no-FBI conditions in Experiment 1.

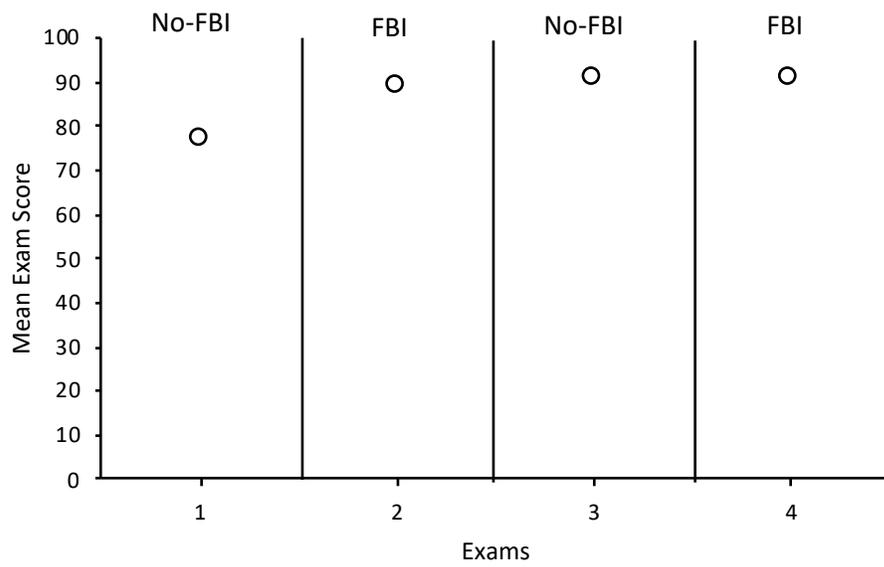


Figure 4. Group mean exam scores across FBI and no-FBI conditions in Experiment 2.

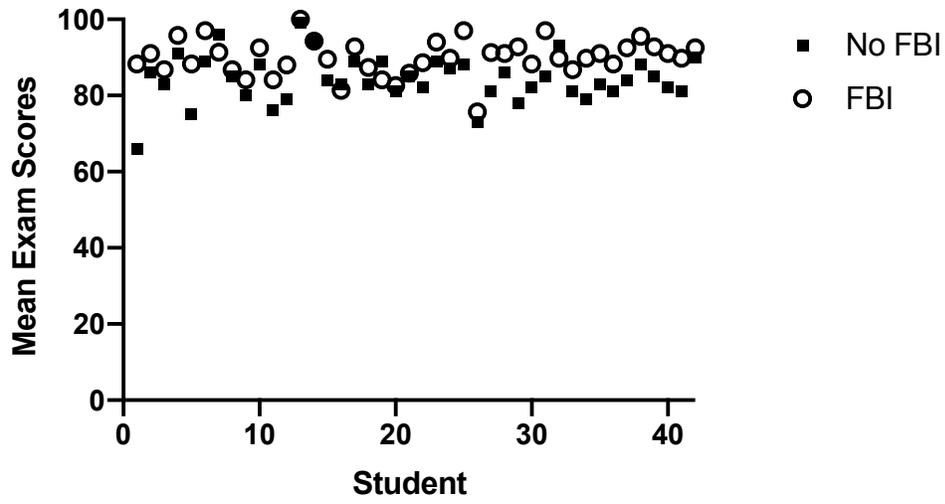


Figure 5. Each student's mean exam score for the no-FBI and FBI conditions in Experiment 2.

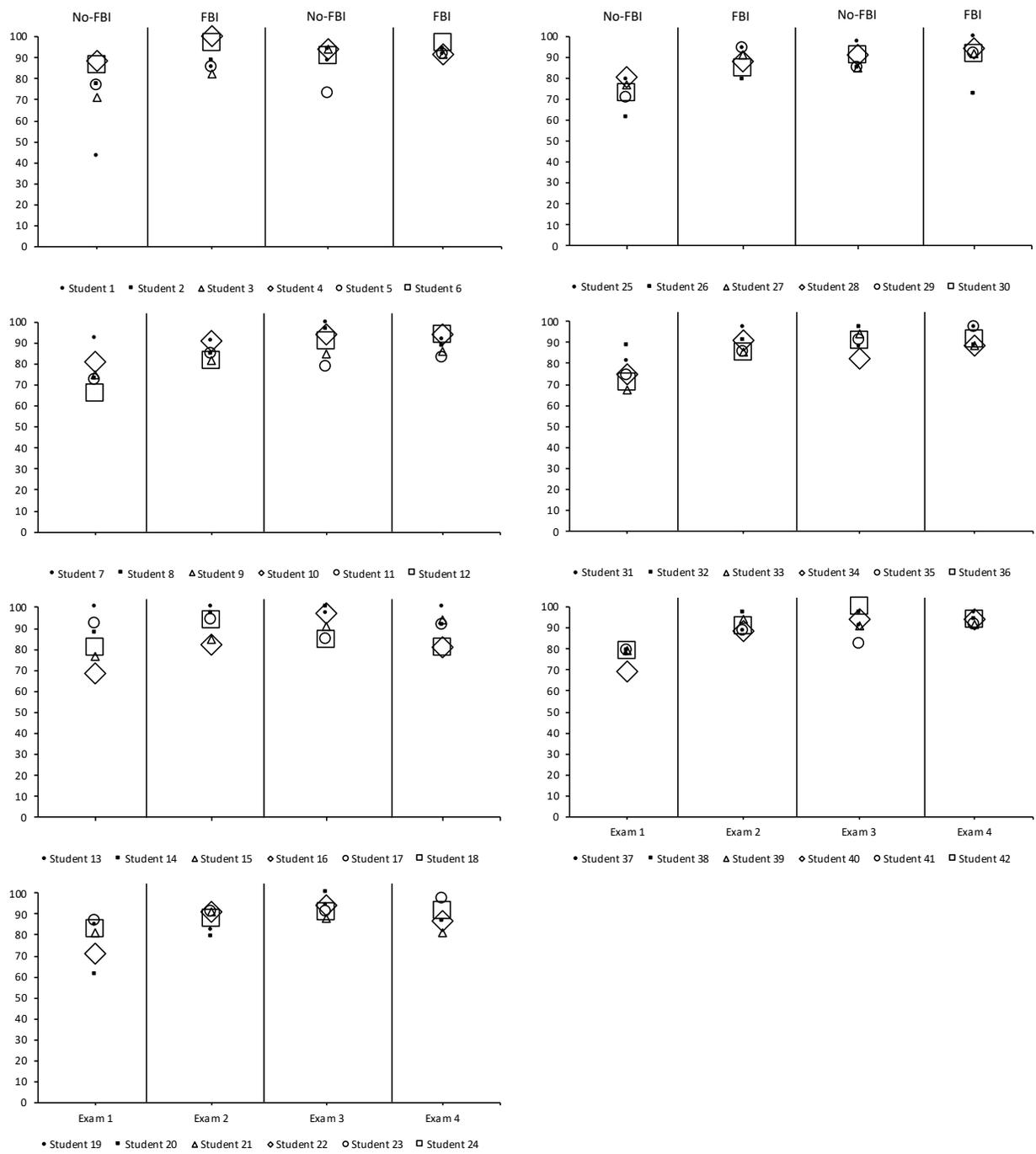


Figure 6. Each student's individual exam score across exams for FBI and no-FBI conditions in Experiment 2.

Appendix A

Fluency Training Instructions

Fluency-based instruction will be used for content relating to quizzes 1, 3, and 5. Students will make index cards with terms on one side and the corresponding definitions on the other from the list provided. An additional application using Quizlet will be used for in home practice drills.

1. Flashcards

- Make flashcards for Deck 1. Complete Deck 2 once criterion is met for Deck 1. **Keep them separate.**

2. Review

- Review the cards at the beginning of each class for 10 minutes silently and independently.

3. In Class Practice Drills

The idea in the practice drill is to state as many correct terms as possible in 1 minute. Your aim should be to increase the number of correct responses for each subsequent timing.

- Following the review, sit across from a partner.
- Randomly mix up your flashcards and hold them with the terms facing your partner.
- Your partner will time you for 1 minute and provide feedback on any errors made.
 - You will make two separate piles as you go through the terms (correct vs incorrect). If an error is made, place the card in the incorrect pile.
- Once your partner states “go” and starts the timer, read **out loud** the definition side of the card and state the corresponding term.
- Flip the card over quickly to confirm your answer is correct and then place the card in its respective incorrect or correct pile.
- Partners if any errors are made that are not self-corrected, be sure to comment that it was an error and confirm card is placed in its respective pile.
- Partners will state “Time” when 1 minute is up.
- Any remaining cards should be set aside.
- Review instructions for data collection.
- Quickly review the cards in the error pile.
- Shuffle all the cards together again.
- Repeat this drill for a total of **three** times.
- Now switch roles with your partner and repeat the procedure.

4. Data collection (on your data sheet)

- Print two data sheets. One for home data collection and one for in class data collection. Circle on data sheet “home” or “class”
- Identify which Deck you are practicing.
- Record your total of correct responses.
- Circle the best performance (number correct) out of the 3 timings.

- Identify if criterion was met **Yes** or **No**.
- Attest to data's accuracy by initialing daily record.
- For Quizlet data, you must record your data using the home data sheet (**bring data sheet with you to class**).
 - Transpose home data to your class data sheet on the following class day.
 - Instructors will collect and store the class data sheet after every class.

5. Mastery Criterion

- Practice drills should be practiced once daily (in and outside of class) until mastery is met for all terms.
- For in home practice drills, complete the drills independently without a partner using Quizlet. Instructions are below.
- Once you have demonstrated mastery of Deck 1 (correct responding for all 10 terms in the deck for at least **2** consecutive timings **in the same day**), replace terms/definitions in your deck with Deck 2.
- Begin Deck 2 on the following day.

6. In Home Practice Drills using Quizlet

- Join the class by clicking on the link: <https://quizlet.com/join/aXgsQvNfd>
- You must set up Quizlet account using your QC email
- Search your class by:
 - Username: **INSTRUCTORJKV**
 - Class name: **JKV213W SP18**
- Additional instructions on how to access your class (If needed): <https://quizlet.com/help/2444142/finding-your-teachers-class>
- Select the Quiz assigned and begin with Deck 1. Move on to Deck 2 **ONLY** when mastery criterion has been met.
- Choose “flashcards” from the menu icons
 - Go to “Options”
 - Select “terms” from the “answer with” dropdown box
- Follow the same procedure as with the flashcards. Set and start your timer for 1 minute
- Say the definition and corresponding term out loud.
 - Click on the card to confirm your answer (term)
 - Select the right arrow to move on to the next term
 - At the end of the timing, record the number correct
 - Click on the “shuffle” icon and repeat the practice drill for a total of three times.
 - Record your best performance and transpose that information on your data sheet upon return to your next class.

You should be aware that your instructor can track and monitor performance

Data Sheet

Student Name: _____
one)

Home/Class (**Circle**

By initialing the daily log I am attesting that I completed the practice drills and the data submitted are true and accurate

Date	Circle Deck	Number Correct			Criterion Met	Initials
Chapters 1 & 3						
2/1	Deck 1/Deck 2				Y/N	
2/2	Deck 1/Deck 2				Y/N	
2/3	Deck 1/Deck 2				Y/N	
2/4	Deck 1/Deck 2				Y/N	
2/5	Deck 1/Deck 2				Y/N	
2/6	Deck 1/Deck 2				Y/N	
2/7	Deck 1/Deck 2				Y/N	
Chapters 14 & 15						
3/13	Deck 1/Deck 2				Y/N	
3/14	Deck 1/Deck 2				Y/N	
3/15	Deck 1/Deck 2				Y/N	

3/16	Deck 1/Deck 2				Y/N	
3/17	Deck 1/Deck 2				Y/N	
CONTINUE ON PAGE 2						
Date	Chapters Studied	Number Correct			Criterion Met (Y/N)	Initials
3/18	Deck 1/Deck 2				Y/N	
3/19	Deck 1/Deck 2				Y/N	
Chapters 10 & 12						
4/17	Deck 1/Deck 2				Y/N	
4/18	Deck 1/Deck 2				Y/N	
4/19	Deck 1/Deck 2				Y/N	
4/20	Deck 1/Deck 2				Y/N	
4/21	Deck 1/Deck 2				Y/N	
4/22	Deck 1/Deck 2				Y/N	
4/23	Deck 1/Deck 2				Y/N	

**** TRANSPOSE DATA FROM QUIZLET AT YOUR NEXT CLASS and SUBMIT YOUR DATA LOG TO YOUR INSTRUCTOR AT THE END OF CLASS***

Appendix B

Instructor (213w): _____

Date: _____

Procedure	Yes	No	Comments
Instructor provided FBI instructions and terms on Blackboard			
Instructions and how to access Quizlet were reviewed the first day of class			
Students were given 10 minutes to review independently on FBI content related class days			
Students were instructed to complete practice drills on FBI class days			
FBI was not implemented for no-FBI related concepts			
Percentage of intervention implemented correctly:			

References

- Beauregard, A., & Basile, K. (2013). Comparing the efficacy of digital flashcards versus paper flashcards to improve receptive and productive L2 vocabulary. *The EUROCALL Review*, 25(1), 3–15. <https://doi.org/10.4995/eurocall.2017.6964>
- Beck, R. (1979). *Report for the Office of Education Joint Dissemination Review Panel*. Great Falls, Montana: Precision Teaching Project.
- Beverley, M.J., Hughes, C., & Hastings, R.P. (2009). What's the probability of that? Using SAFMEDS to increase undergraduate success with statistical concepts. *European Journal of Behavioral Analysis*, 10(2), 235-247.
- Binder, C. (1996). Behavior fluency: evolution of a new paradigm. *The Behavior Analyst*, 19(2), 163-197.
- Binder, C., & Watkins, C.L. (1990). Precision teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly*, 3(4), 74-96.
- Bucklin, B., Dickinson, A., & Brethower, D. (2000). A Comparison of the Effects of Fluency Training and Accuracy Training on Application and Retention. *Performance Improvement Quarterly*, 13(3), 140-163.
- Cihon, T.M., Sturtz., A.M., & Eshleman, J. (2012). The effects of instructor-provided or student created flashcards with weekly, one-minute timings on unit quiz scores in introduction to applied behavior analysis courses. *European Journal of Behavior Analysis*, 13(1), 45-57.
- Cohen, M. (2017). Fluency-based instruction and its effects on learning. *EC Psychology and Psychiatry*, 2(3), 80-81.
- Commons, M.L., Crone-Todd, D., & Chen, S.J. (2014). Using SAFMEDS and direct instruction

- to teach the model of hierarchical complexity. *The Behavior Analyst Today*, 14(1–2), 31–45. <http://doi.org/10.1037/h0101284>
- Fox, E.J., & Ghezzi, P.M. (2003). Effects of computer-based fluency-based instruction on concept formation. *Journal of Behavioral Education*, 12(1), 1–21.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410 LP-8415. <https://doi.org/10.1073/pnas.1319030111>
- Dizon, G., & Tang, D. (2017). Comparing the efficacy of digital flashcards versus paper flashcards to improve receptive and productive L2 vocabulary. *The Eurocall Review*, 25(1), 1-14.
- Graf, S. A. & Auman, J. (2005). *SAFMEDS: a tool to build fluency*. Graf Implements.
- Hanover Research. (2010). Strategies for Teaching Large Undergraduate Classes. Retrieved from https://www.baruch.cuny.edu/facultyhandbook/documents/LargeClasses_StrategiesforTeaching.pdf
- Isaacs, M. (1973). Precision teaching of the deaf. *American Annals of the Deaf*, 118(6), 686-690.
- Johnson, K.R., & Layng, T.J. (1992). Breaking the structuralist barrier: literacy and numeracy with fluency. *American Psychologist*, 47(11), 1475-1490.
- Kubina, Jr., R.M., & Wolfe, P. (2005) Potential Applications of Behavioral Fluency for Students With Autism. *Exceptionality*, 13(1), 35-44. https://doi.org/10.1207/s15327035ex1301_5
- Kulik, J. A., Kulik, C. C., & Cohen, P. A. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. *American Psychologist*, 34(4), 307-318.
- Lindsley, O.R. (1991). Precision teaching's unique legacy from B.F. Skinner. *Journal of*

Behavioral Education, 1(2), 253-266.

Meindl, J.N., Ivy, J.W., Miller, N., Neef, N.A., & Williamson, R.L. (2013). An Examination of Stimulus Control in Fluency-Based Strategies: SAFMEDS and Generalization. *Journal of Behavioral Education*, 22(3), 229–252. <http://doi.org/10.1007/s10864-013-9172-6>

Petri, H., & Govern, J. (2013). *Motivation: Theory, Research, and Application* (6th ed.). Belmont, CA: Cengage Learning.

Quigley, S.P., Peterson, S.M., Frieder, J.E., & Peck, K.M. (2017). A Review of SAFMEDS: Evidence for Procedures, Outcomes and Directions for Future Research. *The Behavior Analyst*, 4(1), 283-301. <http://doi.org/10.1007/s40614-017-0087-8>

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis. *Review of Educational Research*, 69(1), 21–51.
<https://doi.org/10.3102/00346543069001021>

Tuckman, B.W., & Kennedy, G.J. (2011). Teaching Learning Strategies to Increase Success of First-Term College Students. *The Journal of Experimental Education*, 79(4), 478–504.

U.S. Department of Education, National Center for Education Statistics. (2018). *The Condition of Education 2018* (NCES 2018-144). Article retrieved from
<https://nces.ed.gov/fastfacts/display.asp?id=40>

White, T.L., & McBurney, D.H. (2013). *Research Methods* (9th ed.). US: Wadsworth Cengage Learning