Evaluating the Efficacy, Preference, and Cultural Responsiveness of Student-Generated Content in an Undergraduate Behavioral Course

Cynthia E. Nava¹ · Tara A. Fahmie¹ · Sandy Jin¹ · Priya Kumar¹

Published online: 2 May 2019
© Association for Behavior Analysis International 2019

Abstract

Increasing diversity in the field of behavior analysis may begin with an evaluation of culturally responsive practices in the college classroom. This study leveraged the various backgrounds of students in a university nationally recognized for diversity to evaluate the effects of peer-generated course materials on student performance in an undergraduate behavior analysis course. First, graduate students created multimedia examples (videos, pictures) of the behavioral principles in their everyday lives. Next, we curated an online bank of these examples corresponding to 4 topics (respondent conditioning, reinforcement, antecedent control, extinction and punishment) taught in an undergraduate behavior analysis course. We used a multiple-probe and between-group design to evaluate the effects of these peer-generated materials as supplements to traditional instruction. Students showed evidence of concept acquisition on all topics. However, results showed that peer-generated examples, as supplements to textbook and lectures, did not enhance students' performance on knowledge assessments but were rated by students as more preferred, culturally responsive, and diverse than textbook examples.

Keywords College teaching · Concept learning · Culturally responsive teaching · University coursework

Undergraduate coursework can be the gateway to a career in behavior analysis, serving as a source of both information and inspiration. To maximize students' success in their college classrooms, many behavioral faculty have adopted evidence-based pedagogical tactics, including guided notes (e.g., Neef,

McCord, and Ferreri, 2006), frequent practice opportunities (e.g., Perrin, Miller, Haberlin, Ivy, Meindl, and Neef, 2011), response cards (e.g., Marmolejo, Wilder, and Bradley, 2004), points for attendance (Carroll & St. Peter, 2017), points for homework completion (e.g., Rehfeldt, Walker, Garcia, Lovett, & Filipiak, 2010), and stimulus equivalence (e.g., Walker, Rehfeldt, & Ninness, 2010), among others. These practices often focus on designing instruction and arranging contingencies to maximize observable student outcomes, like high test scores and attendance.

An additional aspect of critical importance to instructional design is the responsiveness of teaching practices to the learning of students of diverse backgrounds (e.g., Ladson-Billings, 1995). Although the field of behavior analysis is well poised to evaluate the effects of prior history (e.g., cultural upbringing) on student outcomes, few behavioral studies exist on this topic. Instead, a burgeoning area of research on “culturally responsive teaching” (CRT) has occurred outside of the field of behavior analysis. According to Gay (2002), CRT is defined as

Research Highlights

- Designing instruction for an increasingly diverse student body is of great relevance to the cultivation of diversity in the field of behavior analysis.
- Current technologies easily enable students to generate and share in-the-moment examples of the principles of behavior analysis in their everyday lives.
- Peer-generated content can supplement traditional college textbooks to increase the cultural responsiveness of teaching practices.
- Students performed similarly on tests of knowledge whether or not they had access to peer-generated examples as supplements to textbook chapters and lectures.
- Students rated peer-generated examples of behavioral concepts as more preferred, diverse, and culturally responsive compared to traditional textbook examples.

✉ Tara A. Fahmie
tara.fahmie@csun.edu

¹ Psychology Department, California State University, Northridge, 18111 Nordhoff St., Northridge, CA 91330-8255, USA

using the cultural characteristics, experiences, and perspectives of ethnically diverse students as conduits for teaching them more effectively. It is based on the

assumption that when academic knowledge and skills are situated within the lived experiences and frames of reference of students, they are more personally meaningful, have higher interest appeal, and are learned more easily and thoroughly. (p. 106)

To advance our science, improve our services, and become a more equitable field, behavior analysts must recruit and retain practitioners and scientists of diverse backgrounds (e.g., Fong, Catagnus, Brodhead, Quigley, & Field, 2016). The application of CRT practices to behavior-analytic college courses may be an important first step.

One potential strategy toward this end is to engage students of various backgrounds in the development of materials used to teach. For example, most textbooks and lectures of the behavioral principles include a collection of “everyday” examples of broad concepts, such reinforcement, extinction, and stimulus control (e.g., Miller & Weaver, 1976). Although the course instructor or textbook author may find it easy to generate examples of these concepts, students’ examples of the principles in their everyday lives may represent a more diverse range of applications. Interactive digital technologies and online course management systems make the collection and distribution of student-generated content increasingly efficient. Cultivating a bank of examples of the foundational principles operating on the lives people of diverse backgrounds may improve what some refer to as content integration, or the “extent to which teachers use examples and content from a variety of cultures and groups to illustrate key concepts, principles, generalizations, and theories in their subject area discipline” (Banks, 2006, p. 8).

Unfortunately, no experimental research has been conducted on the effect of student-generated content on the acquisition of foundational behavior-analytic concepts in undergraduate courses. Thus, the current study sought to evaluate the general benefit, if any, added by providing undergraduate students with peer-generated examples from a diverse student body. We assessed students’ concept acquisition, preference for peer-generated examples in comparison to traditional textbook examples, and the extent to which students found peer-generated examples to be culturally responsive.

Method

Setting and Participants

This study took place in an undergraduate capstone course in behavior analysis in a 4-year public university in the San Fernando Valley of California, serving nearly 40,000 students of diverse backgrounds. The course included a 75-min lecture (attended simultaneously by all registered students) and two subsequent 50-min break-out seminars (i.e., each attended by

half of the students) that immediately followed the lecture. At the time of this study, the university was recognized as both a Hispanic Serving Institution and an Asian American, Native American, and Pacific Islander Serving Institution. It also was ranked 16th in the country in awarding bachelor’s degrees to underrepresented minority students and enrolled the largest number of deaf and hard-of-hearing students of any mainstream university in the United States.

The participants of the study were 45 undergraduate students enrolled in the capstone course in behavior analysis who accessed the “peer-generated” content during efficacy and preference evaluations. The peer-generated content was created by 42 graduate students in the 2nd year of a master’s in applied behavior analysis program who were not evaluated in the present study. That is, graduate students created the course material (peer-generated examples) but were not participants in the current study, whereas undergraduate students viewed the course material and were participants in the study. Demographic information on undergraduate psychology majors and graduate students in the behavior analysis program at the time of the study can be found in Table 1. The second author was the instructor of each of these courses during the study. All data were obtained as part of regular and ongoing course evaluation, and all procedures were approved by the university’s institutional review board.

Materials and Independent Variables

The course instructor assigned readings of book chapters and research articles, suggested completion of reading guides, delivered PowerPoint lectures, and administered a one-question reading quiz during each lecture.

Peer-generated examples A semester prior to the launch of the undergraduate course, the second author delivered an optional extra-credit assignment to graduate students enrolled in a developmental psychopathology course. The

Table 1. Percentage of Students by Gender and Ethnicity

		Graduate (Content Creators)	Undergraduate (Participants)
Gender	Male	24	27
	Female	76	73
Ethnicity	African American	0	5
	American Indian	0	.2
	Latino	36	57
	Pacific Islander	0	.1
	Asian American	13	7
	White	30	22
	Multirace/Other	15	3
	Unknown	6	4
	International	0	1

assignment required students to use a mobile application called dscout (<https://dscout.com/>) to collect examples of the behavioral principles in their everyday lives. The mobile application sent out “missions” to the students’ cellular phones, which prompted them to collect in-the-moment photos, videos, and audio of a range of learning principles (e.g., positive reinforcement, stimulus control). After a piece of media was collected by the student, the mobile application prompted the student to caption the example using lay terms and then describe the example using behavioral jargon (see examples in [Appendix A](#)). The instructor checked these entries weekly and provided feedback to the students as necessary. A 2% bump in the student’s final course grade was obtained for each student who submitted at least 10 examples across the 9-week course. In addition, a test exemption was provided to five students who provided the best examples, which were ranked by both quality and quantity. Over 800 examples were obtained from this extra-credit assignment. The graduate students signed media releases if they approved the sharing of their examples for educational purposes; only those examples with releases were used in the study.

The graduate-student examples were collected and added to an online course learning management system (LMS) called Canvas, which would later be accessed by the undergraduate student participants. The LMS page displayed the topic name (e.g., “Respondent conditioning in real life”; see [Appendix A](#)) in a stylized button. When clicked, the button opened up the display of examples, including the image (or still-frame with play button, in the case of a video example) and layperson caption. Each image could be clicked to further access the full description of the concept with behavioral jargon (see [Appendix A](#)).

Peer-generated examples combined with flowchart Halfway through the efficacy evaluation, we introduced a flowchart of the behavioral concept that participants could use while reviewing the examples. These flowcharts were displayed immediately under the full description of the concept and included a series of questions that the student could answer about the example to arrive at the correct behavioral concept (see example in [Appendix A](#)).

Dependent Variables and Data Collection

Knowledge assessments The undergraduate students (hereafter referred to as “participants”) completed a knowledge assessment on four behavioral topics (i.e., respondent conditioning, reinforcement, antecedent control, extinction and punishment) at six separate points in the semester: once prior to any instruction (baseline probe), once immediately following instruction on each of the four topics (postteaching probe for the topic taught, additional

baseline probes for topics not yet taught, maintenance probes for topics previously taught), and once 8 weeks following instruction on the final topic (follow-up probe). These knowledge assessments were constructed to measure *concept identification*, defined as selecting the correct concept when given an example (e.g., Shimamura & Malott, 1994).

Knowledge assessments included five multiple-choice questions per topic in which the stem of each question was an example of a concept sampled from five typical introductory textbooks that were not assigned in the course (Cooper, Heron, & Heward, 2006; Malott & Shane, 2015; Miller, 2005; Miltenberger, 2011). We attempted to equate the difficulty of the stems across questions and topic areas by balancing vocabulary and length of the question. Each question contained the minimum but necessary amount of information regarding each concept. For instance, question stems on the topic of reinforcement included a description of the behavior, the stimulus change immediately following the behavior, and the effect on the future probability of the behavior under similar circumstances. Each question had three answer alternatives in which the participants had to (a) discriminate between a correct concept and an incorrect concept and (b) discriminate between an example and a nonexample (Rodriguez, 2005). We programmed for two possible errors for each topic: *overgeneralization*, or identifying nonexamples as examples, and *undergeneralization*, or identifying examples as nonexamples (e.g., Shimamura & Malott, 1994). The first and fourth authors independently scored the assessments and entered them into a spreadsheet. Any discrepancy was resolved by the second author and updated in the spreadsheet.

The participants did not receive individual feedback on their performance on these assessments, and the assessment scores did not affect their grade in the course; however, participants were eligible for a lottery drawing for gift cards to on-campus dining if the entire group improved their average assessment score compared to their average on the previous assessments. This group contingency was met all but once.

Preference assessment A preference assessment was conducted following the knowledge assessment on the fourth topic (extinction and punishment), following the lecture-based teaching of three new topics (shaping, chaining, and schedules of reinforcement), and prior to a standard course test of these three new topics. Participants’ preference was assessed via a concurrent-chains arrangement with nine choice points (DeLeon, Bullock, & Catania, 2013) on a Canvas LMS page. Initial-link selections were made among three options: textbook example, peer example, and peer example plus flowchart. Our textbook examples were sampled from one of the introductory texts used to design the knowledge assessment and included both text

and any accompanying image (e.g., diagram, cartoon, graph). One third of textbook examples included an image. Our peer examples were sampled from graduate-student submissions.

Selection of an option resulted in access to an example in the format selected. The first three selections were forced choice (i.e., only one option was presented) to ensure that participants sampled each of the three options; the final nine selections were free choice (i.e., all three options were presented), with three selections for each of the three new topics. The display of the three options on the computer screen was counterbalanced across selections.

Social validity questionnaire Immediately following the completion of the preference assessment, participants were asked to complete a questionnaire (see Table 2 and Figure 4) in which they answered a series of questions and rated on a Likert-type scale the extent to which they found the peer-generated and textbook examples relatable to their everyday lives, diverse, and culturally sensitive.

Experimental Design

We used a combination of a group design and multiple-probe design. Participants were able to select one of two seminar sections upon registration into the course. This process was not randomized, as they could elect to register for either of the course sections until that section was full. Whereas all

students received traditional teaching on the four topics during the lecture they attended simultaneously, the independent variable (access to peer-generated examples) was staggered across the seminar sections in an alternating fashion.

Following the 75-min lecture on each topic, the participants received 5 min of question-and-answer time with the instructor, 15 min to view the peer-generated examples on a personal or borrowed electronic device, and 30 min to complete the knowledge assessment during the seminar. The knowledge assessment was conducted prior to access to the peer-generated examples in one seminar and following access to the peer-generated examples in the other seminar. The seminar who received the peer-generated examples prior to the assessment was alternated for each topic. This arrangement ensured that (a) all participants received a similar overall quality of instruction prior to the seminar, (b) the peer-generated examples were not withheld entirely from one group of participants, and (c) the independent variable was staggered.

Thus, the efficacy of the independent variable was evaluated across the two seminars, in an alternating fashion per topic, by comparing the performance of the seminar that had access to the peer-generated examples and the seminar that did not have access to the peer-generated examples; it was evaluated within a seminar by probing the performance on the knowledge assessment via staggering of the topics taught with and without prior exposure to the examples.

Table 2. Answers on the Social Validity Survey

Question	Answer Options				
	Number (Percentage) of Student Selections				
Which of the following most accurately reflects your ability to relate examples of behavioral principles to your own personal life?	I relate best to textbook examples. 1 (3%)	I relate best to peer-generated examples. 23 (79%)	I relate equally to textbook and peer-generated examples. 5 (17%)	I do not relate to textbook examples and I do not relate to peer-generated examples. 0 (0%)	
Which of the following is most helpful for learning new concepts when you study on your own?	Textbook alone 0 (0%)	Textbook + Study guides 3 (10%)	Textbook + Study guides + Peer-generated examples 4 (14%)	Textbook + Study guides + Peer-generated examples + Flowcharts 22 (76%)	None of the above are helpful. 0 (0%)
Which of the following (check all that apply) do you like to see incorporated into lectures while learning new concepts?	Text 16 (55%)	Pictures 24 (83%)	Videos 26 (90%)	Captions 14 (48%)	Charts/ diagrams 24 (83%)
If given the choice, which would you prefer?	A few peer-generated examples incorporated into lectures on each topic 7 (24%)	Access to all peer-generated examples for review on your own time 2 (7%)	Both of the above 20 (69%)	No access to peer-generated examples 0 (0%)

Fig. 1 Mean percentage correct on knowledge assessments in Seminar 1 (black circles) and Seminar 2 (white circles). The shaded area of the line graphs depicts performance immediately following teaching (postteaching probe); the shaded area of the bar graphs depicts this

same performance as a percentage change from baseline. The “IV” indicates the seminar that received the peer-generated examples (independent variable) prior to their assessment. BL = baseline

Fig. 2 Error patterns demonstrated during baseline and follow-up assessments. Each data point depicts the number of errors of a single participant, and bars represent group means

Fig. 3 Number of participants selecting each of the three preference options across nine choice points

Results

Figure 1 depicts the results of the knowledge assessments across topics. Prior to any teaching, participants scored 53%–60% correct on Topic 1 (respondent conditioning), 55%–65%

correct on Topic 2 (reinforcement), 45%–54% correct on Topic 3 (antecedent control), and 49%–75% correct on Topic 4 (punishment and extinction). Following teaching (i.e., postteaching probe), each seminar showed improvements in assessment scores compared to their average baseline score. Postteaching scores on Topic 1 were 62%–78% correct, on Topic 2 were 78%–80% correct, and on Topic 3 were 73%–80% correct. Postteaching scores on Topic 4, punishment and extinction, were similar to baseline scores, at 64%–73% correct; baseline scores on Topic 4 were higher than baseline scores on all other topics. The bar graph on the right panel of Figure 1 depicts the differential effects of access to peer-generated examples on assessment scores. Access to peer-generated examples prior to the assessment (i.e., the independent variable) showed a potential negative effect (Topic 1) or no effect (Topic 2) on assessment scores. We assumed that simple exposure to peer-generated examples was insufficient and added a flowchart to the subsequent two topics to encourage guided practice of critical thinking. Although Topic 3 showed a potential positive effect of peer-generated examples plus the flowchart, this effect was not replicated with Topic 4. Thus, overall, we did not demonstrate a main effect of the independent variable. We also analyzed the extent to which the independent variable influenced group performance when students scoring 90% or higher on baseline measures were removed from this analysis; similar effects were obtained with our secondary analysis, in that the independent variable did not produce a differential effect on the students' acquisition of concepts.

Maintenance and follow-up probes of the assessments showed scores similar to postteaching probes for all but Topic 1 (respondent conditioning). A further analysis of error patterns in Figure 2 shows that errors of undergeneralization generally decreased, whereas errors of overgeneralization persisted.

Figure 3 depicts preference assessment data at the group level. For all choice points except one, peer-generated examples plus flowcharts were selected most frequently, and textbook examples were selected least frequently. Due to a data extraction error, we were unable to report on all selections for

Fig. 4 Scores on the social validity survey

the third choice point of the chaining topic; thus, complete data for only eight choice points are shown in Figure 3. A review of individual cumulative graphs (Appendix B) shows that 12 participants showed mixed preferences, with either selections that alternated across all options or selections that alternated across two options; 7 participants selected peer-generated examples substantially more than the other two options; and 21 participants selected the peer-generated examples plus flowcharts substantially more than the other two options. No participant selected textbook examples substantially more than the other two options.

Social validity survey results are displayed in Table 2 and Figure 4. A larger percentage of participants reported relating better to the peer-generated examples (79%) compared to textbook examples (3%). The majority of participants (76%) also reported that the textbook combined with study guides, peer-generated examples, and flowcharts was the most helpful for learning new concepts. On two survey questions, participants indicated that they liked the pictures and videos from the peer-generated examples and would prefer that the peer-generated examples were incorporated into lectures, as well as available for review on their own time. Overall, participants rated peer-generated examples as more “relatable to your personal life,” “diverse,” and “culturally sensitive” compared to textbook examples when given a Likert-type scale (Figure 4). Several answers to the open-ended questions on the survey revealed that students liked the “visuals” provided in the peer-generated examples and liked seeing examples that were similar to what they do daily (e.g., being stuck in traffic).

Discussion

This study used student-generated examples of behavioral principles as a way to introduce diverse and culturally representative content into an undergraduate course in behavior analysis. The examples were created by advanced graduate students and were delivered to undergraduate peers, some of whom were learning about behavior analysis for the first time. The examples, as supplements to traditional lectures and readings, did not affect the students’ concept acquisition on tests of knowledge. However, the peer-generated examples were preferred by students when assessed against traditional textbook examples; the peer-generated examples also were rated as superior on a social validity questionnaire administered to students. These latter effects may have been achieved by the enriched format (multimedia: video, picture, and text); the depiction of diverse students engaging in diverse

behaviors across diverse contexts; the depiction of nontraditional social norms, family structures, and relationships; the depiction of behavior-environment relations experienced on a daily basis by university students; or some combination of these factors that were captured in the peer-generated examples.

Although the peer-generated examples did not enhance concept acquisition in our study, we feel that their differential effect on preference and social validity were particularly meaningful in light of recent calls to increase diversity and equity in behavior analysis (e.g., Fong et al., 2016; Fong & Tanaka, 2013). That is, when students perceive our instructional practices as culturally sensitive and responsive, we may be more likely to recruit and retain students of diverse backgrounds into our field. On the other hand, lack of cultural responsiveness in a classroom may lead to disengagement or discontent of underrepresented students (e.g., Garcia & Chun, 2016; Griner and Stewart, 2013; Sampson & Garrison-Wade, 2011). Educators acknowledge that student success is a combination of specialized skill acquisition and, equally important, the motivation to continue to engage with course content outside of the classroom, in future courses, in graduate study, and in career decisions. Thus, an intervention like ours, which requires minimal effort on the part of instructors but that enhances the *value* of course content for students of diverse backgrounds, should be considered a priority of curricular decision-making (Brown, 2007). Unfortunately, few studies exist on the efficacy, acceptability, and use of instructional strategies designed explicitly to promote diversity in higher education (e.g., Bottiani, Larson, Debnam, Bischoff, & Bradshaw, 2017). We feel that this is a ripe area for behavior analysts to contribute their experimental methodologies.

However, there are several limitations of the study worth noting. First, the university system for course registration prohibited a random assignment of students across seminars. Given the differences in performance across seminars in our baseline measures, it is likely that our groups differed in ways that impacted the integrity of our efficacy analysis. Second, not only did our baseline scores on concept identification greatly exceed chance responding (33%), but also a notable increasing trend was shown during the extended baseline for Topic 4. This may have impacted our ability to show meaningful differences across seminars. It is possible that these high baseline scores were achieved by students in this upper-division course because they had prior exposure to behavior analysis in a lower-division course. Although prior exposure to behavior analysis was not a prerequisite for this course, 30 participants had taken a lower-division course in behavior analysis

previously, 3 were concurrently enrolled in a lower-division course in behavior analysis, and 11 were employed as behavior interventionists during the study. Future research should evaluate the effect of this intervention with students who are completely naive to the behavioral concepts and learning them for the first time.

Third, our use of the peer-generated examples (and flowcharts, when applicable) as an independent variable was restricted to a 15-min, in-class viewing on a personal device. We structured the delivery of intervention this way to standardize the “dose” of intervention across students, seminars, and topics, as well as to complete the analysis without disrupting the general course schedule. However, active responding is known to be a crucial part of concept acquisition (e.g., Miller & Weaver, 1976). A fourth limitation was the students’ continued errors on questions that probed for overgeneralization, or the selection of a nonexample as an example. We should not be surprised by this outcome, however, given that our intervention included examples only, with no exposure to nonexamples or the boundary conditions of the concepts we taught (e.g., McIlvane, 2013). These limitations highlight the importance of considering previous work in our field in the realm of instructional design (e.g., Bernstein & Chase, 2013). For example, our intervention may have proven more efficacious if we required active responding (e.g., homework assignments that included completion of quiz questions or flowchart questions), incorporated both examples and nonexamples, and required the participants to generate new examples with instructor feedback.

The results of our study produced several avenues for future research. Access to peer-generated examples may have influenced student engagement or perception of the behavioral sciences in ways that were not captured by our dependent measures. Future research should assess the effect of peer-generated examples on *concept exemplification*, defined as

creating a correct example when given the concept. In addition, we encourage researchers to create operational definitions for, and to evaluate objective measures (e.g., content analyses; Gullicks, Pearson, Child, and Schwab, 2005) of, the items captured by our Likert-type scale: diversity, relevance, and cultural sensitivity of peer-generated and textbook examples. For example, we received several submissions showing grandparents as caregivers; it would be interesting to analyze how often extended family members are mentioned as caregivers in traditional behavior-analytic texts. Our omission of operational definitions and objective measures made it difficult to pinpoint the active components of our intervention. Finally, we recommend conducting a longitudinal study to track the effect of this intervention, in combination with other culturally responsive instructional practices, on students’ commitment to behavior analysis as a profession.

Acknowledgements We thank the California State University, Northridge, Master of Science in Applied Behavior Analysis students for their vast contributions to the completion of this study. We thank Emily Tierman, Jasmine Poetry, Elizabeth Hernandez, and Carissa Basile for their help with this study.

Funding This research was supported in part by NIGMS BUILD (Building Infrastructure Leading to Diversity) grants TL4GM118977 and RL5GM118975.

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Appendix A

Fig. 5 Stylized button on Canvas site for “Antecedent Control” (Topic 3)

"Complained about the sun in my eyes so he stuck this on the visor to shut me up." (Reinforcement)

"Kids' soccer game" (Stimulus control)

"Looking out at the parkour gym prior to class." (Respondent conditioning)

"Virgin Mary Shrine" (Cultural selection)

"Michael is vacuuming our dusty floors" (Reinforcement)

Fig. 6 A collection of peer-generated examples, shared with permission by Uilani Elmer, Alejandra Chiu, Michael Aragon, and Neomi Garzon. Example of the behavioral jargon description of the top left example, above: I complained (highly aversive auditory stimulus) very much

about the sun in my face one day until he stuck this piece of cardboard on the visor. Then my complaining suddenly stopped. Now when the sun is bright and I'm in the car, he sticks this up and avoids my complaining (negative reinforcement)

Fig. 7 Example of the flowchart to accompany the topic “Extinction and Punishment” (Topic 4)

Appendix B

Fig. 8 Cumulative Graphs of Each Participants' Preference Assessment

References

- Banks, J. A. (2006). *Cultural diversity and education: Foundations, curriculum and teaching*. Boston: Pearson Education.
- Bernstein, D., & Chase, P. N. (2013). Contributions of behavior analysis to higher education. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 523–543). Washington, DC: American Psychological Association. <https://doi.org/10.1037/13938-021>.
- Bottiani, J. H., Larson, K. E., Debnam, K. J., Bischoff, C. M., & Bradshaw, C. P. (2017). Promoting educators' use of culturally responsive practices: A systematic review of inservice interventions. *Journal of Teacher Education*, 69(4), 67–385. <https://doi.org/10.1177/0022487117722553>.
- Brown, M. R. (2007). Educating all students: Creating culturally responsive teachers, classrooms, and schools. *Intervention in School and Clinic*, 43(1), 57–62. <https://doi.org/10.1177/10534512070430010801>.
- Carroll, R. A., & St. Peter, C. C. (2017). Proportion of available points predicts student attendance in college courses. *The Psychological Record*, 67(1), 61–69. <https://doi.org/10.1007/s40732-016-0207-y>.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2006). *Applied behavior analysis* (2nd ed.). Upper Saddle River: Prentice Hall.
- DeLeon, I. G., Bullock, C. E., & Catania, A. C. (2013). Arranging reinforcement contingencies in applied settings: Fundamentals and implications of recent basic and applied research. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 47–75). Washington, DC: American Psychological Association. <https://doi.org/10.1037/13938-021>.
- Fong, E. H., & Tanaka, S. (2013). Multicultural alliance of behavior analysis standards for cultural competence in behavior analysis. *International Journal of Behavioral Consultation and Therapy*, 8, 17–19. <https://doi.org/10.1037/h0100970>.
- Fong, E. H., Catagnus, R. M., Brodhead, M. T., Quigley, S., & Field, S. (2016). Developing the cultural awareness skills of behavior analysts. *Behavior Analysis in Practice*, 9, 84–94. <https://doi.org/10.1007/s40617-016-0111-6>.
- Garcia, C., & Chun, H. (2016). Culturally responsive teaching and teacher expectations for Latino middle school students. *Journal of Latina/o Psychology*, 4(3), 173. <https://doi.org/10.1037/lat0000061>.
- Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education*, 53(2), 106–116. <https://doi.org/10.1177/0022487102053002003>.
- Griner, A. C., & Stewart, M. L. (2013). Addressing the achievement gap and disproportionality through the use of culturally responsive teaching practices. *Urban Education*, 48(4), 585–621. <https://doi.org/10.1177/0042085912456847>.
- Gullicks, K. A., Pearson, J. C., Child, J. T., & Schwab, C. R. (2005). Diversity and power in public speaking textbooks. *Communication Quarterly*, 53(2), 247–258. <https://doi.org/10.1080/01463370500089870>.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3), 465–491. <https://doi.org/10.3102/00028312032003465>.
- Malott, R., & Shane, J. T. (2015). *Principles of behavior*. New York: Psychology Press.
- Marmolejo, E. K., Wilder, D. A., & Bradley, L. (2004). A preliminary analysis of the effects of response cards on student performance and participation in an upper division university course. *Journal of Applied Behavior Analysis*, 37, 405–410. <https://doi.org/10.1901/jaba.2004.37-405>.
- McIlvane, W. J. (2013). Simple and complex discrimination learning. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (pp. 129–163). Washington, DC: American Psychological Association. <https://doi.org/10.1037/13938-021>.
- Miller, L. K. (2005). *Principles of everyday behavior analysis*. Monterey: Brooks/Cole Publishing.
- Miller, L. K., & Weaver, F. H. (1976). A behavioral technology for producing concept formation in university students. *Journal of Applied Behavior Analysis*, 9, 289–300. <https://doi.org/10.1901/jaba.1976.9-289>.
- Miltenberger, R. G. (2011). *Behavior modification: Principles and procedures*. Belmont, CA: Wadsworth, Cengage Learning.
- Neef, N. A., McCord, B. E., & Ferreri, S. J. (2006). Effects of guided notes versus completed notes during lectures on college students' quiz performance. *Journal of Applied Behavior Analysis*, 39, 123–130. <https://doi.org/10.1901/jaba.2006.94-04>.
- Perrin, C. J., Miller, N., Haberlin, A. T., Ivy, J. W., Meindl, J. N., & Neef, N. A. (2011). Measuring and reducing college students' procrastination. *Journal of Applied Behavior Analysis*, 44, 463–474. <https://doi.org/10.1901/jaba.2011.44-463>.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: a meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3–13. <https://doi.org/10.1111/j.1745-3992.2005.00006.x>.
- Rehfeldt, R. A., Walker, B., Garcia, Y., Lovett, S., & Filipiak, S. (2010). A point contingency for homework submission in the graduate school classroom. *Journal of Applied Behavior Analysis*, 43, 499–502. <https://doi.org/10.1901/jaba.2010.43-499>.
- Sampson, D., & Garrison-Wade, D. F. (2011). Cultural vibrancy: Exploring the preferences of African American children toward culturally relevant and non-culturally relevant lessons. *The Urban Review*, 43(2), 279–309. <https://doi.org/10.1007/s11256-010-0170-x>.
- Shimamura, S., & Malott, R. W. (1994). An analysis of concept learning: Simple conceptual control and definition-based conceptual control. *The Analysis of Verbal Behavior*, 12, 67–78. <https://doi.org/10.1007/BF03392898>.
- Walker, B. D., Rehfeldt, R. A., & Ninness, C. (2010). Using the stimulus equivalence paradigm to teach course material in an undergraduate rehabilitation course. *Journal of Applied Behavior Analysis*, 43, 615–633. <https://doi.org/10.1901/jaba.2010.43-615>.