Topical Article

SOCIETY FOR THE TEACHING OF PSYCHOLOGY

Teaching of Psychology
2016, Vol. 43(2) 99-107

Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0098628316636274
top.sagepub.com
SAGE

Team-Based Learning Improves Course Outcomes in Introductory Psychology

Lisa L. Travis[1], Nathan W. Hudson[2], Genevieve M. Henricks-Lepp[3], Whitney S. Street[1], and Jennifer Weidenbenner[3]

Abstract

This study investigated the influence of team-based learning (TBL) methods on exam performance and student satisfaction in an introductory psychology class. Fifteen instructors teaching 29 sections (with a combined enrollment of approximately 1,130 students) were randomly assigned to use TBL for 7 of 12 major topics or to use lecture. All students took the same midterm and final exams and completed midsemester and end-of-semester satisfaction surveys. Multilevel logistic models revealed that across both exams, students in the TBL sections performed significantly better on items that tested content covered in the TBL modules. In terms of the overall course satisfaction, there was no difference between the students taught via TBL versus lecture. These findings suggest that TBL is more effective than lecture in contributing to learning among introductory psychology students—without negatively impacting course satisfaction.

Keywords

team-based learning, introductory psychology, assessment

Team-based learning (TBL) is an increasingly popular type of cooperative learning in which students are assigned to teams that remain constant throughout the semester and work together to solve complex application problems based on course material (Michaelson, 2004). Many studies have reported positive outcomes associated with the use of TBL (for reviews see Fatmi, Hartling, Hillier, Campbell, & Oswald, 2013; Sisk, 2011). However, many of these studies have suffered from methodological limitations—such as lacking appropriate comparison groups or randomized designs—which have limited their ability to provide clear evidence that TBL *causes* increased academic performance. Furthermore, very few studies have tested TBL with introductory-level undergraduate students (cf. Carmichael, 2009). To address these limitations, the present paper describes the results of a randomized experiment evaluating the efficacy of TBL methods—as compared with more traditional, lecture-based methods—in improving learning outcomes and student satisfaction in a large introductory psychology course.

What is TBL?

TBL is a form of small group learning that has two distinctive features (Fink, 2004). The first involves the formation of teams that remain constant throughout the semester, with the rationale that permanent teams foster commitment among teammates and allow students to learn how to effectively interact with one another. The second distinctive feature of TBL is that content is presented in modules that include (1) preclass preparation, (2) in-class quizzes, and (3) application exercises. Specifically, students prepare for each module by studying assigned material before coming to class. In class, the module begins with a challenging multiple-choice quiz taken first individually and immediately afterward as a team. Team members discuss each question until they identify the correct answer and are given immediate feedback. Subsequently, students complete application exercises that require the use of course concepts to solve meaningful problems. These applications are designed with the goal of stimulating in-depth analysis, discussion, and critical thinking. Finally, teams report their solutions simultaneously, with a successful application exercise culminating in instructor-guided debate between groups about the merits of competing answers.

Previous Research on TBL

TBL and Learning Outcomes

Although many studies have found positive associations between the use of TBL and academic outcomes across a variety of disciplines (e.g., Fatmi et al., 2013; Sisk, 2011), many of

[1] Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL, USA
[2] Department of Psychology, Michigan State University, East Lansing, MI, USA
[3] Department of Educational Psychology, University of Illinois at Urbana-Champaign, Champaign, IL USA

Corresponding Author:
Lisa L. Travis, Department of Psychology, University of Illinois at Urbana-Champaign, 603 E. Daniel St., Champaign, IL, 61820, USA.
Email: lltravis@illinois.edu

these studies lacked strong experimental designs. Specifically, researchers tend to use three designs to evaluate TBL's effectiveness: within-person experiments, cohort comparisons, and between-person experiments. With respect to the first, several within-person experiments have suggested that, within a cohort of students, performance on topics taught via TBL exceeds performance on topics taught using other methods (e.g., Koles, Stolfi, Borges, Nelson, & Parmelee, 2010). One limitation of these types of within-person designs, however, is that course content is confounded with method of instruction; and consequently, it is difficult to rule out the possibility that the content taught via TBL was simply intrinsically easier.

Second, other studies have attempted to evaluate the efficacy of TBL by comparing cohorts of students taught via traditional curriculum to later cohorts taught using TBL. Several such studies have reported that TBL cohorts performed better on class exams or on externally administered exams such as medical board exams (Levine et al., 2004; McInerney & Fink, 2003; Vasan, DeFouw, & Compton, 2011; Zgheib, Simaan, & Sabra, 2010). Although impressively consistent in providing support for TBL, cohort comparisons may reveal differences for a variety of reasons, making it desirable to bolster these findings with experimental comparisons of TBL versus other methods within a single cohort.

Finally, to the best of our knowledge, only three studies have directly compared the performance of students taught via TBL to their peers *in the same cohort* who were taught *the same material* via lecture versus other methods. Two of these studies were conducted within medical schools (Thomas & Bowen, 2011; Zingone et al., 2010). Both studies reported superior performance on measures of learning for groups taught via TBL methods. In the study most similar to ours, Carmichael (2009) compared learning in two 200-student sections of an introductory biology class. One section was taught solely by lecture, whereas in the other section, TBL was used in conjunction with lecture. Students in the TBL section outperformed students in the other section on three of four major exams and also showed advantages on a separate measure of their ability to interpret data.

To summarize, there is a fair amount of support for the effectiveness of TBL in improving learning. However, only a few studies have employed randomized designs, and even fewer have investigated TBL in the context of introductory-level undergraduate courses. Our study was designed to help fill these gaps by employing a between-persons true experimental design in an introductory psychology class.

TBL and Student Satisfaction

The evidence linking TBL to improved student satisfaction is less compelling than the evidence for its association with improved academic performance. Although many studies report that students are enthusiastic about TBL, many of these suffer from methodological limitations such as lack of a comparison group or exclusive use of cohort comparisons (Abdelkhalek, Hussein, Gibbst, & Handy, 2010; Haberyan, 2007; McInerney & Fink, 2003; Vasan, DeFouw, & Compton, 2009, Vasan et al., 2011; Zgheib et al., 2010). Moreover, not all studies investigating student satisfaction have produced positive results; several studies employing focus groups have found that during interviews, students express preference for lectures over TBL (Bick et al., 2009; Hunt, Haidet, Coverdale, & Richards, 2003).

One recent meta-analysis of the effectiveness of TBL in medical education (Fatmi et al., 2013) included 7 studies that featured a direct comparison of student satisfaction with TBL to a control group, and only one of these studies found that satisfaction was higher among students taught via TBL (Levine et al., 2004). Thus, overall, the literature presents somewhat mixed findings with regard to student satisfaction with TBL. Consequently, we did not have strong predictions about how course satisfaction would differ among introductory psychology students taught via TBL versus more traditional instruction methods.

Overview of the Present Study

In the present study, we randomly assigned 14 different introductory psychology instructors either to incorporate a standardized TBL curriculum into their course or to teach using traditional, lecture-based methods. Performance on exams (which were the same across instruction method), as well as responses to student satisfaction surveys, was compared across teaching methods.

Method

Participants

Data were collected from approximately 1,130 undergraduate students who were enrolled in an introductory psychology course in fall 2013. Students were enrolled in 29 different sections, taught by 15 different graduate student instructors. Graduate instructors were all advanced students in psychology or related fields. Ten of the 15 were teaching the course for the first time.

Procedure

Overview. Fourteen of the 15 introductory psychology instructors in fall 2013 were randomly assigned to either (1) incorporate a standardized TBL curriculum into their course or (2) use traditional lecture.[1] For instructors assigned to implement the TBL curriculum, 12 of the 42 total class sessions—spread across 7 chapters—were dedicated to completing a series of standardized TBL modules that were developed by the authors of this article. Instructors assigned to the traditional lecture condition were not allowed to implement any of the TBL quizzes or activities into their classrooms and instead provided a primarily lecture-based course.[2] All students completed the same midterm and final exams and the same course satisfaction survey.

Prior to the start of the semester, instructors completed a 2-hr training session designed to introduce them to TBL. Throughout the semester, prior to each module, instructors met as a group to review the upcoming material and discuss strategies for improving the effectiveness of our implementation of TBL. At the beginning of the semester, students in TBL sections were randomly assigned, using a computer program, into teams of 5–7 students

TBL modules. Each TBL module consisted of four components: (1) out-of-class preparation, which involved reading approximately 10 pages of the textbook; (2) an in-class individual quiz; (3) the same quiz retaken as a team, with immediate feedback provided by Immediate Feedback Assessment Technique scratch cards; and (4) team application exercises, which typically involved applying course concepts to solve problems or answer questions about scenarios that could occur in everyday life (see the Appendix for sample team application exercise questions). Between-group discussion and instructor debriefing followed each application exercise.

Dependent Measures

Midterm exam. All students (n = 1,115) completed the same midterm exam during Week 7 of the semester (students who took a different makeup exam were excluded from the analyses). The midterm exam consisted of 60 multiple-choice questions, covering a total of 6 chapters—3 of which contained TBL modules. Students' responses to each individual question on the exam were collected and scored as correct or incorrect. Prior to performing any analyses, two independent judges rated whether each question covered topics that were explicitly featured in the team activities. Any discrepancies between the judges were resolved by discussing the items in question. A total of 18 (30%) items were rated as pertaining to topics covered during TBL modules.

Final exam. All students (n = 1,126) completed the noncumulative final exam during Week 16 of the semester. The final exam consisted of 90 multiple-choice questions. A total of seven chapters were covered on the final exam, four of which contained TBL modules. A total of 24 (27%) items were rated as pertaining to topics explicitly covered during team activities.

Course satisfaction. A 13-item course satisfaction survey was created for the purposes of this study. Students anonymously completed the survey twice—once during Weeks 12–13 of the semester (after nine TBL modules had been completed) and again during Weeks 15–16 (after an additional two TBL modules had been completed). The survey contained some basic demographic information (gender, college, self-reported midterm grade) as well as four satisfaction subscales that are described below. All items in each of the four subscales were rated on a response scale from *strongly disagree* (1) to *strongly agree* (5). The items were averaged into subscale composites, and the four subscale composites were averaged together with equal weighting (α =.74) to form an overall satisfaction composite. The four subscales were as follows:

Enjoyment: Students used a 5-item scale to rate how much they enjoyed the course. Sample items include, "I enjoyed this class" and "If I had the opportunity, I would take another class in psychology." Items were averaged together (α = .88).

Motivation: Students rated their motivation to achieve in the class using a 3-item subscale. The items were, "I am motivated to work hard in this class," "I attend this class regularly," and "I read the course textbook more than I would in other introductory courses." Items were averaged together (α = .54).

Perceived learning: Students used a 3-item scale to rate their perceived learning in the course. Sample items include, "I have learned a lot in this class" and "This class increased my ability to think critically." Items were averaged together (α = .69).

Favorableness toward instruction method: Students rated how favorably they felt toward the instruction method employed in their section using a 2-item scale. The items were "The methods of instruction used in this class helped me learn" and "I would prefer if the teacher used different instruction methods in this class" (reversed). The items were averaged together (r = .61).

Student perceptions of TBL. For students in TBL sections only, we administered a 12-item questionnaire measuring their perceptions of TBL. All items were measured on a scale from *strongly disagree* (1) to *strongly agree* (5). Two items explicitly asked participants to compare TBL to lecture. A principal components analysis with varimax rotation revealed that the remaining 10 items clustered into two factors: TBL-positivity and TBL-involvement.

Preference for TBL over lecture. Two questions asked students to rate how favorably they viewed lectures directly in comparison to TBL (e.g., "I enjoyed regular classes with lectures more than Team Exercise Days;" "I would prefer a typical lecture to Team Exercise Days"). Both items were reversed such that higher scores represented preferences for TBL over lecture and were averaged together (r = .76).

TBL-positivity. Six items measured students' generalized positivity toward TBL. The items were "I found myself less interested in the subject because of Team Exercise Days" (reversed), "My motivation to learn increased because of Team Exercises," "Team Exercises were not useful to my learning" (reversed), "The group discussions allowed me to correct my mistakes and improve my understanding of concepts," "Team Exercise Days helped me prepare for course examinations," and "Redoing quizzes with my team helped me learn more." Items were averaged to form a composite (α = .83).

TBL-involvement. Four items asked participants to rate how involved they and their team members were in team activities. The items were, "I prepared for Team Exercise Days," "My team worked well together," "Most of my team contributed meaningfully during Team Exercise Days," and "I contributed meaningfully during Team Exercise Days." Items were averaged to form a composite ($\alpha = .68$).

Results

We analyzed students' probability of selecting correct answers on the exams using multilevel logistic models (MLLMs). Specifically—first for each exam and subsequently for both exams collapsed together—the following parameters of the MLLM were estimated:

$$\ln\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right) = b_0 + b_1(\text{TBL Curriculum})_k + b_2(\text{Item TBL})_i + b_3(\text{TBL Curriculum})_k\,(\text{Item TBL})_i + U_{jk} + U_k + \varepsilon_{ijk}.$$

In these analyses, the log-transformed odds of students correctly answering individual questions on the exams were modeled as a function of (1) whether the student was taught using the TBL curriculum or lecture [b_1], (2) whether the individual item pertained to topics explicitly covered in TBL activities [b_2], (3) the interaction between instruction-method and item type [b_3], (4) a random intercept for the student [nested within instructors] to control for within-student dependencies in answering questions correctly [U_{jk}], and (5) a random intercept for each instructor to control for instructor effects [U_k].[3,4] Although not depicted for simplicity, we also controlled for whether the students were enrolled in an honors section in all analyses.[5]

Midterm Exam

We first tested for a main effect of instruction method on overall midterm performance. This was accomplished by dropping the b_2 and b_3 parameters from the above model. Overall, students taught via TBL performed significantly better overall on the midterm exam, compared with their peers taught using traditional lecture, odds ratio (OR) = 1.18, 95% confidence interval (CI) [1.04, 1.34].[6] Translated into model-predicted probabilities, students taught using TBL had a 73.1% (95% CI [71.4, 74.7]) probability of correctly answering questions on the exam, whereas students taught via lecture had a 69.7% (95% CI [67.8, 71.6]) probability of answering questions correctly.

We next examined whether the effect of TBL was stronger for items that were specifically covered by the team activities. As shown in Table 1, there was a significant interaction between the TBL intervention and item type (OR = 1.35, 95% CI [1.25, 1.46]) such that TBL students performed especially well on items covered by the application exercises (simple OR = 1.48; 95% CI [1.29, 1.70]). As can be seen in Figure 1, students taught via TBL were predicted to correctly answer 80.5% (95% CI [79.0, 82.0]) of items that were explicitly covered by application exercises, whereas their peers taught via lecture were predicted to correctly answer only 73.7% (95% CI [71.7, 75.5])—a difference of 6.8 percentage points. However, with respect to items that were not explicitly covered by application exercise, performance was not statistically significantly higher among students taught via TBL (70.0%, 95% CI [68.1, 71.7]), as compared with those taught via lectures (68.0%, 95% CI [66.0, 70.0]; simple OR = 1.09; 95% CI [0.97, 1.23]).

Table 1. Multilevel Logistic Regression Coefficients Predicting Probability of Answering Midterm Exam Questions Correctly.

				OR 95% CI	
	b	*SE*	OR	LB	UB
TBL intervention[a]	.09	0.06	1.09	0.97	1.23
Item TBL	.27	0.03	1.31	1.25	1.39
TBL intervention × Item TBL	.31	0.04	1.35	1.25	1.46

Note. OR = odds ratio; CI = confidence interval; LB = lower bound of 95% CI; UB = upper bound of 95% CI; TBL = Team-Based Learning; *SE* = standard error. 95% CIs that do not contain 1.00 are significant, $p < .05$.
[a]Because of how the model is specified, this is the simple effect of the TBL intervention on items *not* covered by the TBL curriculum.

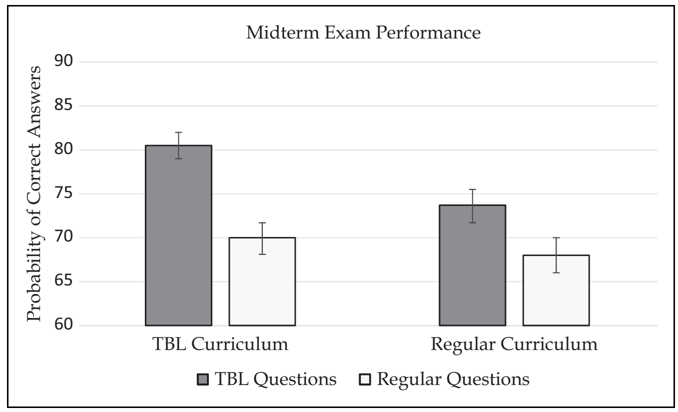

Figure 1. Model-predicted probability of answering midterm exam items correctly as a function of instruction method and item type, with 95% confidence intervals depicted.

Final Exam

There was no main effect of instruction type on overall performance on the final exam (OR = 1.15, 95% CI [0.94, 1.41]). However, as can be seen in Table 2 and Figure 2, replicating the midterm results, there was a significant interaction between the TBL intervention and item-type (OR = 1.16, 95% CI [1.09, 1.25]) such that for items that pertained to topics covered by the application exercises, TBL students performed statistically significantly better (model-predicted probability: 80.8%, 95% CI [78.5, 82.9]) than students taught using lectures (76.5%, 95% CI [73.7, 79.1]; simple OR = 1.29, 95% CI [1.05, 1.59]).

Table 2. Multilevel Logistic Regression Coefficients Predicting Probability of Answering Final Exam Questions Correctly.

				OR 95% CI	
	b	*SE*	OR	LB	UB
TBL Intervention[a]	.10	0.10	1.11	0.91	1.36
Item TBL	.10	0.02	1.10	1.05	1.16
TBL Intervention × Item TBL	.15	0.04	1.16	1.09	1.25

Note. OR = odds ratio; CI = confidence interval; LB = lower bound of 95% CI; UB = upper bound of 95% CI; TBL = Team-Based Learning; *SE* = standard error. 95% CIs that do not contain 1.00 are significant, $p < .05$.
[a]Because of how the model is specified, this is the simple effect of the TBL intervention on items *not* covered by the TBL curriculum.

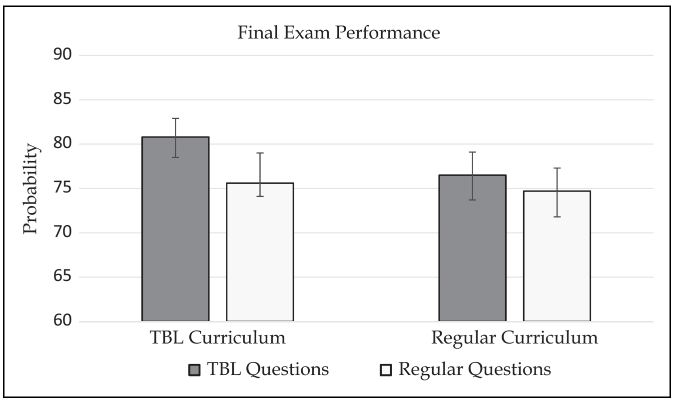

Figure 2. Model-predicted probability of answering final exam items correctly as a function of instruction method and item type, with 95% confidence intervals depicted.

However, for items that were not explicitly covered by the team activities, there was no difference in performance depending on whether students were taught via TBL (75.6%, 95% CI [74.1, 79.0]) or lecture (74.7%, 95% CI [71.8, 77.3]; simple OR = 1.11, 95% CI [0.91, 1.36]).

Combined Analysis Across Both Exams

We also examined students' total performance on both the midterm and the final exam collapsed together (150 items total). In terms of a main effect, students exposed to the TBL intervention had significantly higher odds of correctly answering questions, OR = 1.18, 95% CI [1.01, 1.38]. Moreover, students taught using TBL performed especially well on items that were specifically covered by the team activities. As can be seen in Table 3, there was a significant interaction between the TBL intervention and item-type (OR = 1.24, 95% CI [1.18, 1.31]) such that students taught via TBL performed especially well on exam items that were explicitly covered during team activities (simple OR = 1.39, 95% CI [1.17, 1.64]). However, students taught using TBL did not have statistically significantly higher odds of correctly answering questions that were not covered by the TBL curriculum (simple OR = 1.12, 95% CI [0.94, 1.32]).

Table 3. Multilevel logistic regression coefficients predicting probability of answering questions correctly across the midterm and final exams.

				OR 95% CI	
	b	*SE*	OR	LB	UB
TBL Intervention[a]	0.11	0.09	1.12	0.94	1.32
Item TBL	0.17	0.02	1.18	1.14	1.22
TBL Intervention × Item TBL	0.22	0.03	1.24	1.18	1.31

Note. OR = odds ratio; CI = confidence interval; LB = lower bound of 95% CI; UB = upper bound of 95% CI; TBL = Team-Based Learning; *SE* = standard error. 95% CIs that do not contain 1.00 are significant, $p < .05$.
[a]Because of how the model is specified, this is the simple effect of the TBL intervention on items *not* covered by the TBL curriculum.

Taken together, these findings suggest that, across both the midterm and final exams, students taught using TBL performed significantly better than their peers taught using traditional lectures. This effect seems to be primarily driven by TBL students having a higher probability of correctly answering questions, which pertained to material that was explicitly covered by the TBL curriculum.

Course Satisfaction

For our final series of analyses, we examined whether students' self-reported course satisfaction varied as a function of instruction method (see Table 4 for descriptive statistics and intercorrelations for all satisfaction variables). To do so, we used multilevel models to predict the satisfaction composite at Time 1 and Time 2 as a function of instruction method, controlling for (1) instructor effects (with a random intercept) and (2) whether the student was enrolled in an honors section (as a fixed effect). These models revealed that there were no differences between students taught with TBL versus lectures in the satisfaction composite at Time 1 ($\beta = .03$, 95% CI [−0.18, 0.25]) or Time 2 ($\beta = .06$, 95% CI [−0.17, 0.29]).[7] There were also no differences between the students taught via TBL versus lectures with respect to any of the individual variables that comprised the satisfaction composite (enjoyment, motivation, perceived learning, and favorability toward instruction method).

Previous research has suggested that demographics, such as gender or scholastic performance, predict students' satisfaction with TBL (Espey, 2010; Reinig, Horowitz, & Whittenburg, 2011; Vasan et al., 2009). In terms of main effects, collapsing across instruction type, there were no differences in course satisfaction by gender ($\beta = -.04$, 95% CI [−0.11, 0.03]). In contrast, students enrolled in the college of liberal arts and sciences (LAS; which contains the psychology major) were slightly more satisfied with the course than were students in any other college ($\beta = .06$, 95% CI [0.01, 0.10]). Similarly, students who self-reported higher midterm exam scores rated their satisfaction with the course higher than did students with lower self-reported midterm grades ($\beta = .21$, 95% CI [0.16, 0.25]). Despite these main effects, neither gender, college, nor self-reported midterm grade moderated instruction type in

Table 4. Course Satisfaction Descriptive Statistics and Correlations.

	M	*SD*	Correlations 1	2	3	4	5	6	7
1. TBL	0.52	0.50	–						
2. Male	0.41	0.49	−.02	–					
3. Self-report midterm grade	3.49	1.14	**.13**	−.02	–				
4. Enrolled in LAS	0.43	0.50	.01	.00	.01	–			
5. Course Satisfaction	3.68	0.63	.02	−.03	**.23**	**.09**	–		
6. Preference for TBL vs. Lecture[a]	2.63	1.17	–	.01	**−.07**	.01	**−.07**	–	
7. TBL-Specific Positivity[a]	3.41	0.77	–	.04	.02	**.09**	**.38**	**.59**	–
8. TBL-Specific Involvement[a]	4.11	0.64	–	**−.08**	**.07**	**.08**	**.33**	**.26**	**.47**

Note. TBL = Team-Based Learning.
[a]These variables were measured only for students in the TBL sections; correlations in boldface are significant, $p \leq .05$.

predicting any satisfaction measure, all $|\beta|s \leq .04$, 95% CIs ranged from [−0.05, 0.04] to [−0.01, 0.08]. This indicates that there were no differences in course satisfaction for students in TBL versus lecture-based sections, irrespective of gender, college of enrollment (LAS vs. other), or self-reported midterm exam score.

Student perceptions of TBL. Finally, students in TBL sections (but not those in lecture sections) were given a brief survey measuring their perceptions and impressions of team activities. As can be seen in Table 4, students, on average, reported positive impressions of TBL, $M = 3.41$ (95% CI [3.37, 3.45]; the scalar midpoint of 3.00 represents neutrality toward TBL with higher numbers indicating favorability toward TBL). In contrast, when directly asked whether they preferred TBL to lectures, most students reported a preference for lecture, $M = 2.63$ (95% CI [2.57, 2.69]; the scalar midpoint of 3.00 indicates neutrality for lecture vs. TBL with higher numbers indicating preferences for TBL).

Discussion

The present study compared the relative effectiveness of TBL to lecture-based instruction in enhancing students' exam performance and course satisfaction in a large introductory psychology class. The structure of our course—in which several instructors taught separate sections, but all students took the same exams—afforded the opportunity to use a stronger design than is typically found in studies of TBL. We were able to randomly assign instructors to use TBL versus traditional lectures and to examine the effects of TBL while statistically controlling for instructor effects.

Does TBL Improve Exam Performance?

We hypothesized that TBL would promote exam performance, based on the well-established positive associations between small group learning and academic outcomes (e.g., Springer, Stanne, & Donovan, 1999), as well as existing empirical studies demonstrating that TBL in particular may be an effective form of small group learning (e.g., Carmichael, 2009; Koles et al., 2010). Consistent with expectations, we found that students taught via TBL performed moderately better across both the midterm and the final exams, as compared with students taught via lectures. Notably, the performance gains were specific to content that was explicitly covered by TBL activities, suggesting that these differences were not attributable to preexisting differences in ability between the two groups of students. These findings support the idea that TBL, when implemented effectively, can increase students' academic performance.

It is important to note that TBL had only a moderate impact in increasing students' exam performance, with students taught via TBL outperforming their lecture-taught peers by 5–7 percentage points on exam questions pertaining to TBL topics. While modest, this effect size is practically significant, because it is large enough to be represented as an increase in letter grade. In other words, this effect size is one that academic systems tend to acknowledge as meaningful.

Why might TBL have increased students' exam performance? Although we did not test specific mechanisms through which TBL might promote learning, various scholars have argued that small group learning—and TBL in particular—can catalyze learning through a conjunction of multiple different processes (Slavin, 2013). In the present study, both motivational and cognitive processes are likely to have been influential. The structure of the TBL method required that students study and come prepared to 12 class sessions and rewarded them for doing so. Thus, students taught via TBL may have studied more overall, or in a more distributed fashion, than did students exposed only to lecture. Their motivation to prepare for team activities may have been further enhanced by feelings of accountability to their teammates. As many have suggested (e.g., Slavin, 2013; Springer et al., 1999), we believe that working on challenging problems with peers is a context that is especially conducive to learning.

Does TBL Increase Students' Course Satisfaction?

We found no differences in overall course satisfaction between students who received TBL instruction and those taught using traditional lectures. In general, students expressed positive attitudes toward team exercises. Nevertheless, when directly asked whether they preferred lecture or TBL, most students in TBL sections indicated a preference for lecture.

These results align with Fatmi, Hartling, Hillier, Campbell, and Oswald' (2013) meta-analysis suggesting that course satisfaction does not differ as a function of whether students are taught via TBL versus traditional lectures. In addition, our findings may provide a framework for reconciling some of the mixed findings in the existing literature regarding the impact of TBL on course satisfaction. Consistent with previous research (e.g., Abdelkhalek et al., 2010; Haberyan, 2007; Zgheib et al., 2010), we found that students did, in fact, report high levels of satisfaction with TBL. However, this finding in isolation does not imply that TBL *improves* course satisfaction, as compared with traditional lecture. Indeed, in line with previous research using focus groups (Bick et al, 2009; Hunt et al., 2003), when directly asked whether they preferred lecture or TBL, students in our TBL sections reported preferences for lecture. It may be that the greater workload associated with TBL days affected this preference. Although these findings may initially seem contradictory, in reality they are not. It is possible that, as in our data, students report *liking* TBL but liking lectures *better* (cf. Carmichael, 2009).

Ultimately, however, it is important to emphasize that in direct, between-group comparisons, students were not, on average, *more* satisfied with a course based primarily on lecture than with one that regularly incorporated TBL. As such, TBL students' assessments of whether they *would have been* more satisfied with a course based entirely on lecture appear to have been inaccurate. This may partially reflect the fact that students are poor at assessing how much they learn and benefit from different instructional and study techniques (Bjork, Dunlosky, & Kornell, 2013; McCabe, 2011).

Limitations and Future Directions

Several limitations of our study should be considered. One limitation is that we did not measure or statistically control for time that instructors spent covering particular topics. With respect to TBL instructors, the standardized curriculum ensured some degree of consistency across TBL sections in amount of instructional time devoted to specific topics. However, for non-TBL sections, instructors were free to choose how much time to devote to those topics, almost certainly resulting in more variance in time spent on these topics in non-TBL sections. Moreover, TBL instructors may have spent more time covering TBL-related concepts than did lecture-based instructors. Better exam performance on topics explicitly covered in team activities may simply have reflected more class time spent covering those topics.

That being said, given that class time is constant, to the extent that TBL instructors spent more time covering topics in the TBL modules, they necessarily would have been required to spend less time covering other topics. If class time devoted to topics were the primary factor driving exam performance, we would expect students in TBL sections to perform *worse* on non-TBL items. This trade-off, however, was not observed on either exam. Nevertheless, we cannot rule out the possibility that instructional time spent is one factor contributing to our pattern of results.

A second limitation of our study is that we did not explore the specific mechanisms by which TBL might increase academic performance. For example, quizzing, preparation for quizzing, or group discussion may have contributed disproportionately to increased exam performance. Proponents of TBL argue that the TBL method combines these elements into a particularly effective package—an intriguing possibility to be addressed in future research.

Third, our instructors were relatively inexperienced. TBL might have different effects with more experienced instructors. It is possible that with more experienced instructors, the effects would either be reduced or increased. More experienced instructors might result in more effective teaching regardless of method. Alternatively, more experienced instructors might be better able to effectively implement TBL, increasing its efficacy. Future research should disentangle these possibilities.

Finally, it is possible that experimenter or placebo effects may partially explain our findings. That is, instructors or students may have expected that TBL would boost academic performance, and these expectations per se may have subtly influenced their behavior in ways that contributed to our pattern of findings.

Conclusion

A growing body of research findings suggest that active methods are superior to lecture in enhancing learning outcomes (e.g., Freeman et al., 2014; Slavin, 2013). In addition, creative ideas for alternatives to lecture abound in the educational literature (e.g., Barkley, Cross, & Major, 2005; Sweet & Michaelson, 2012). Our study adds further experimental evidence that, at least for introductory psychology students, active participation—and TBL in particular—is more effective than lecture in supporting student learning. Additionally, although our students reported a preference for lectures over TBL, the use of TBL did not negatively affect course satisfaction. We believe that these results should be encouraging for instructors who are considering adopting TBL for introductory-level undergraduate classes.

Appendix

Sample TBL Team Exercise

Below you will find 3 vignettes about fictional scientists testing some of their theories/hypotheses. For each vignette, your job is to decide whether the idea (theory/hypothesis) being tested is scientific or not. Subsequently, you will need to decide whether the provided evidence supports, fails to support, proves, or disproves the theory/hypothesis. You must be prepared to justify your answers to the class.

The quality of adult romantic relationships is determined by childhood experiences with mom. John believes that children's experiences with their mothers affect their adult romantic relationships. He interviews hundreds of adults about their relationships with their romantic partners and with their mothers. He finds absolutely no relationship between the quality of people's current romantic relationships and the quality of their relationships with their moms.

1. Is the theory/hypothesis scientific?
   a. Yes
   b. No
2. Does the evidence described support, fail to support, prove, or disprove the theory/hypothesis?
   a. Support
   b. Fail to support
   c. Prove
   d. Disprove
   e. None of the above

Artificial sugars help people to lose weight. Cynthiana believes that artificial sugars, like saccharin, help people to lose weight. She takes a couple hundred obese rats and feeds half gloop flavored with saccharin, and the other half gloop flavored with sugar. She finds that neither group of rats loses any weight over time.

3. Is the theory/hypothesis scientific?
   a. Yes
   b. No
4. Does the evidence described support, fail to support, prove, or disprove the theory/hypothesis?
   a. Support
   b. Fail to support
   c. Prove
   d. Disprove
   e. None of the above

Acknowledgments

We thank Chris Burns for his valuable advice regarding implementation of team-based learning, for providing a training session for our instructors, and for his helpful comments on an earlier draft of the manuscript.

Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by the University of Illinois, Champaign-Urbana Liberal Arts and Sciences Top 10 Initiative.

Notes

1. One instructor (teaching only 1 section with 18 students) decided that she did not want to use the TBL materials prior to the randomization procedure. Consequently, the remaining 14 instructors (who taught 2 sections and approximately 96 students each) were randomly assigned to teach using TBL versus traditional lectures. Data from the 15th instructor's students ($n = 18$) were included in all analyses. However, a fixed effect was included in all models to control for any unique impact of this instructor.
2. When lecturing, all instructors employed daily, nonstandardized I-Clicker questions. Instructors were free to incorporate demos, videos, or other methods as they deemed appropriate.
3. The TBL Curriculum and Item TBL variables were both dummy-coded such that a score of "1" indicated that the curriculum/item was TBL whereas a "0" meant the curriculum/item was *not* TBL.
4. Collapsing across the midterm and final exam, the intraclass correlation (ICC) for instructors was .003. The ICC for students was .06.
5. We also controlled for the 15th instructor—who opted out of the randomization procedure—as a fixed effect. Similarly, one instructor in the control group used a dramatically different "quiz bowl" style of lecturing. We modeled any unique impact of this instructor as an additional fixed effect.
6. An odds ratio is a standardized effect size that can be approximately converted into more familiar metrics, such as Cohen's d, using the formula $d \approx \ln(\text{OR})/1.81$ (Chinn, 2000).
7. All predictor and criterion variables were standardized across the full sample prior to being entered in the multilevel models.

References

Abdelkhalek, N., Hussein, A., Gibbst, T., & Handy, H. (2010). Using team-based learning to prepare medical students for future problem-based learning. *Medical Teacher*, *32*, 123–129.

Barkley, E. F., Cross, K. P., & Major, C. H. (2005). *Collaborative learning techniques: A handbook for college faculty*. San Francisco, CA: John Wiley.

Bick, R. J., Oakes, J. L., Actork, J. K., Cleary, L. J., Felleman, D. J., Ownby, A. R., ... Seifert, W. E. (2009). Integrative teaching: Problem-solving and integration of basic science concepts into clinical scenarios using team-based learning. *Journal of the International Association of Medical Science Education*, *19*, 26–34.

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

Carmichael, J. (2009). Team-based learning enhances performance in introductory biology. *Journal of College Science Teaching*, *38*, 54–61.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*, 3127–3131.

Espey, M. (2010). Valuing teams: What influences students' attitudes? *North American Colleges and Teachers of Agriculture Journal*, *54*, 31–40.

Fatmi, M., Hartling, L., Hillier, T., Campbell, S., & Oswald, A. E. (2013). The effectiveness of team-based learning outcomes in health professions education: BEME Guide No. 30 *Medical Teacher*, *35*, e1608–e1624.

Fink, L. D. (2004). Beyond small groups: Harnessing the extraordinary power of learning teams. In L. K. Michaelson, A. B. Knight

& L. D. Fink (Eds.), *Team-based learning: A transformative use of small groups in college teaching* (pp. 3–26). Sterling, VA: Stylus.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoraofor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8410–8415.

Haberyan, A. (2007). Team-based learning in an Industrial/Organizational psychology course. *North American Journal of Psychology*, *9*, 143–152.

Hunt, D., Haidet, P., Coverdale, J., & Richards, B. (2003). The effect of using team-based learning in an evidence-based medicine course for medical students. *Teaching and Learning in Medicine*, *15*, 131–139.

Koles, P. G., Stolfi, A., Borges, N. J., Nelson, S., & Parmelee, D. X. (2010). The impact of team-based learning on medical students' academic performance. *Academic Medicine*, *85*, 1739–1745.

Levine, R. A., O'Boyle, M., Haidet, P., Lynn, D. J., Stone, M. M., Wolf, D. V., & Paniagua, F. A. (2004). Transforming a clinical clerkship with team learning. *Teaching and Learning in Medicine*, *16*, 270–275.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory and Cognition*, *39*, 462–476.

McInerney, M. J., & Fink, L. D. (2003). Team-based learning enhances long-term retention and critical thinking in an undergraduate microbial physiology course. *Microbiology Education*, *4*, 3–12.

Michaelson, L. K. (2004). Getting started with team-based learning. In L. K. Michaelson, A. B. Knight & L. D. Fink (Eds.), *Team-based learning: A transformative use of small groups in college teaching* (pp. 27–50). Sterling, VA: Stylus.

Reinig, B. A., Horowitz, I., & Whittenburg, G. E. (2011). The effect of team-based learning on student attitudes and satisfaction. *Decision Sciences: Journal of Innovative Education*, *9*, 27–47.

Sisk, R. J. (2011). Team-based learning: Systematic research review. *Journal of Nursing Education*, *50*, 665–669.

Slavin, R. E. (2013). Cooperative learning and achievement: Theory and research. In I. B. Weiner (Series Ed.) W. M. Reynolds & G. J. Miller (Vol. Eds.) *Handbook of psychology, Vol 7 educational psychology* (2nd ed., pp. 179–198). New York, NY: John Wiley.

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*, 21–51.

Sweet, M., & Michaelson, L. K. (Eds.). (2012). *Team-based learning in the social sciences and humanities*. Sterling, VA: Stylus.

Thomas, P. A., & Bowen, C. W. (2011). A controlled trial of team-based learning in an ambulatory medicine clerkship for medical students. *Teaching and Learning in Medicine: An International Journal*, *23*, 31–36.

Vasan, N. S., DeFouw, D. O., & Compton, S. (2009). A survey of students' perceptions of team-based learning in anatomy curriculum: Favorable views unrelated to grades. *Anatomical Sciences Education*, *2*, 150–155.

Vasan, N. S., DeFouw, D. O., & Compton, S. (2011). Team-based learning in anatomy: An efficient, effective and economical strategy. *Anatomical Sciences Education*, *4*, 333–339.

Zgheib, N. K., Simaan, J. A., & Sabra, R. (2010). Using team-based learning to teach pharmacology to second-year medical students improves performance. *Medical Teacher*, *32*, 130–135.

Zingone, M. M., Fransk, A. S., Guirguis, A. B., George, C. M., Howard-Thomson, A., & Heidel, R.E. (2010). Comparing team-based and mixed active-learning methods in an ambulatory care elective course. *American Journal of Pharmaceutical Education*, *74*, 1–7.