Team-Based Learning in the Social Studies: Replication Across Grade Levels

Jeanne Wanzek, Florida State University,
Shawn Kent, Florida State University,
Sharon Vaughn, University of Texas at Austin
Elizabeth Swanson, University of Texas at Austin
Greg Roberts, University of Texas at Austin
Martha Haynes, Florida State University

Background / Context

In their recent review of the literature to identify the signature teaching pedagogies in social studies classrooms, Beck and Eno (2012) reported two distinct instructional models, mainstream signature pedagogy and emerging signature pedagogy. Mainstream signature pedagogy, included instructional methods where the primary responsibility of teachers was to transmit knowledge and historical information to students. This pedagogy has prevailed historically and has included lectures, textbook reading, and memorization of facts. Emerging signature pedagogy was defined as a more student-centered process of inquiry that included discussion, debate, role play, or project/problem-based learning. Beck and Eno pointed out that even though teachers may receive training in these processes of inquiry they may not be equipped to transfer this instruction to the K-12 classroom setting. In addition, the authors noted very few studies considering the learning outcomes of students participating in instructional methods under the emerging pedagogy (Beck & Eno, 2012).

For decades, researchers have argued that given the rich and complex nature of the facts, concepts, and artifacts in the social studies, greater depth of critical thinking and understanding can be attained via instruction that encourages and promotes discussion, perspective taking and solving problems (Barton & Levstik, 2003; Okolo, Ferretti, & MacArthur, 2007; Paxton, 1999; Scruggs & Mastropieri, 2003). However, to date, based on observations, as well as teacher and student reports of instructional practices, there appears to be a scarcity of these emerging pedagogies, particularly in the secondary grades. For example, a survey of instructional methods revealed secondary teachers identified lecture as the most frequently implemented instructional technique as well as the most effective (Bolinger & Warren, 2007). In contrast, discussion/debate and role playing were reported as the least frequently implemented instructional activities. Observational research of secondary social studies classrooms has also reported the infrequency of discussion, with 82% of observations noting no discussion of content (Swanson et al., in review).

Collaborative student discourse can be defined as “reflective discussions among students about academic content” (Nussbaum, 2008, p.348). During these discussions, students have input and contribute substantively to the dialogue through sustained reciprocal interactions with peers and/or teachers. Theoretically, this collaborative discourse-based instruction is rooted in Vygotsky’s sociocultural framework of learning (Vygotsky, 1978; Vygotsky & Kozulin, 1986).
Vygotsky’s theory emphasized language as critical to cognitive development and posited that meaning was derived via social interaction and communication with others. Studies specifically examining the impact of collaborative discourse on students’ content learning and acquisition are limited (Nussbaum, 2008), particularly in the area of social studies and for students in the middle and secondary grades.

**Purpose / Objective**

The purpose of this presentation is to present two studies that examine the efficacy of a collaborative, discourse-based instructional framework, team-based learning (TBL; Michaeelsen & Sweet, 2011), on the content acquisition of eighth and eleventh grade social studies students. Specifically, we randomly assigned social studies teachers’ class periods to TBL or typical instruction. TBL was then implemented in the treatment classes for three instructional units (nine weeks of instruction). During the instructional period, we sought to examine teachers’ instructional practices in both conditions. Following the instructional units we examined students’ content learning on several content knowledge measures.

**Setting and Research Design**

Study 1 was conducted in three schools in two school districts in the Southeastern part of the United States. A total of seven, 11th grade social studies teachers and their 26 United States History classes participated. Class sections were randomly assigned to treatment (TBL) or comparison (typical practice) conditions blocking on teacher. In total, 15 classes were randomly assigned to treatment and 11 classes were randomly assigned to comparison.

Six, eighth grade United States History teachers and their students in four middle schools in two school districts in the Southeastern section of the United States participated in Study 2. Using the same randomization procedures as study 1, 13 classes were randomly assigned to treatment and 11 classes were randomly assigned to comparison.

**Population / Participants / Subjects**

In Study 1, all seven teachers (four male) were certified and held bachelor’s degrees; six of the teachers also held graduate degrees. Teaching experience ranged from three to 34 years (M = 14.71 years, median = 14 years). A total of 463 students in the teachers’ 26 class sections consented for participation in the study. The total sample included 193 (42%) and 270 (58%) females with approximately 47% White, 45% African American, 4% Hispanic, 2% Asian, and 3% multiracial. Thirty-eight percent of the sample was enrolled in free or reduced lunch programs. Six percent of the sample was identified with a disability (i.e. speech impaired, language impaired, specific learning disability, deaf/hard of hearing).

In Study 2, all six teachers (five male) were certified and held bachelor’s degrees; three of the teachers also held advanced degrees. Teaching experience ranged from one to 29 years (M = 10.8 years, median = 7.5 years) with over 90% of their experience at the eighth grade level. The 358 students in the teachers’ classes who consented to participate in the study included 163 (46%) males and 195 (55%) females. Approximately 39% of the students were White, 45% African American,
7% Asian, 4% Hispanic, and 5% multiracial. Fifty-one percent of the sample was eligible for the free or reduced lunch programs. Ten percent of the sample was identified with a disability (i.e. speech impaired, language impaired, specific learning disability, other health impaired, autism, orthopedic impairment).

**Intervention / Program / Practice**

Teachers were asked to implement the TBL intervention in three consecutive content units (15-days each). TBL implementation consisted of strategically formed, permanent student teams, comprehension checks, targeted instruction, and knowledge application. During each 15-day unit/cycle, two short comprehension checks and one cumulative comprehension check were conducted and each was followed by targeted instruction. One knowledge application activity occurred in each unit as a culminating activity and was spread across part of two class periods.

**Student teams.** During comprehension checks and knowledge application activities, students worked in teams of three to five students. Teachers strategically assigned students to heterogeneous teams prior to the first unit, and students remained in the same team throughout all three units in order to build team cohesion.

**Comprehension checks.** Comprehension checks included individual and team checks of content understanding through multiple choice questions. First, each student individually completed questions covering the unit content taught to date. Following the individual check, students moved into their teams and took the same quiz again. The teams were asked to discuss each question, identify evidence from text or other class materials to support their answers, and arrive at a consensus as to the correct answer. Each team received immediate feedback about their answer through the use of a scratch-off card keyed to the comprehension check answers. Teachers facilitated discussion and use of evidence during the team comprehension checks.

**Targeted instruction.** Teachers used the individual and team comprehension checks to identify misunderstandings or content requiring reteaching or clarification. Targeted instruction occurred for approximately 5 min at the beginning of each class period following a short comprehension check and for approximately 10 min at the beginning of the next class period following the cumulative comprehension check.

**Knowledge application.** Knowledge application took place at the end of each unit and was designed for students to integrate content learned through discussion of authentic, multilayer, overarching questions or activities incorporating key content from the unit. During knowledge application each team discussed the various aspects of the question, identified evidence from handouts, presentations, class notes, or readings to support their points, and recorded key information. A final written activity was completed to address the question and integrate the key information from the discussion. Following completion of the written activity, each team reported their conclusions and rationale to the group.

**Data Collection and Analysis**
Students in each class were administered a battery of assessments by the research team prior to unit instruction. During implementation, two independent researchers were trained to conduct observations in each teacher’s treatment and comparison classes to determine the extent to which the treatment was implemented and to identify any evidence of TBL elements in the comparison condition. A 3-point Likert scale was used to rate implementation and quality for each component as well as overall quality and classroom management. Following the completion of the units, all students were assessed with the complete set of posttest measures within two weeks of the last unit day.

To estimate the main effect of treatment on student outcomes, we fit hierarchical linear models (MLwiN 2.23; Rasbash, Steele, Browne, & Goldstein, 2004) with students nested in classes and classes nested in teachers. We grand-mean centered pretest scores on level-1 of the model (Enders & Tofighi, 2007). The effects of treatment were modeled at level 2.

Findings / Results

In Study 1, the average fidelity of implementation rating for targeted instruction across teachers was 3.76 ($SD = 1.69$) with a quality of implementation mean at 3.44 ($SD = .86$). Targeted instruction is the only intervention component that was occasionally not observed in implementation. Targeted instruction was not implemented by teachers in 24% of the observations. The comprehension checks were implemented consistently with all or most of the required elements and procedures noted in each observation. The mean implementation rating for the comprehension checks was 4.94 ($SD = .24$) with a mean quality of implementation at 3.41 ($SD = 1.03$). The knowledge application was also implemented consistently, but proved to be the most challenging for teachers, with some required elements and procedures missing from most of the observations. Knowledge application also had the greatest variation in implementation score. The mean implementation rating for knowledge application was 3.22 ($SD = 1.33$) suggesting that, overall, a little more than half of the required elements and procedures were implemented during the three knowledge application implementations. The mean quality of implementation for the knowledge application activities was 2.71 ($SD = 1.25$). No TBL elements were noted in the comparison condition.

There was a significant main treatment effect on students’ content area knowledge ($\gamma_{01} = 1.92, SE = .78, p = .01$). Participants in the intervention outperformed comparison students by an average of 1.92 raw score points. The Hedges’ $g$ was .19. Average posttest differences on the Gates MacGinitie Reading Comprehension subtest did not differ across condition ($\gamma_{01} = .36, SE = 1.84, p = .84$). The Hedges’ $g$ was .03.

In Study 2, the mean implementation of the TBL treatment across teachers was 3.51 ($SD = 1.48$) with a mean quality of implementation of 2.84 ($SD = .79$). The mean implementation and quality ratings for targeted instruction were 3.70 ($SD = 1.17$) and 3.05 ($SD = .97$), respectively. The mean implementation rating for the comprehension checks was 4.92 ($SD = .28$) with all or nearly all of the required elements and procedures implemented in more than 92% of the observations. Quality of the comprehension check implementation averaged 2.69 ($SD = .85$). The
knowledge application component was implemented with a mean rating of 3.04 (SD = 1.54) and a mean quality of implementation of 2.77 (SD = .65). Teachers’ mean implementation of the treatment as well as their quality of implementation were significantly correlated with the treatment students’ posttest content and supporting details scores in the written essays (rs = .230 to .298), but were not correlated with students’ posttest ASK scores (rs = -.044 to -.091).

In the comparison condition, two teachers did implement one element of targeted instruction, reviewing or reteaching prior content, resulting in a mean targeted instruction implementation of 1.67 (SD = .94) across the six teachers. No other TBL elements were observed during any teacher’s comparison classes. There was a significant difference in global instruction (p = .013) and in global classroom management (p < .001) between treatment and comparison classes. Teachers’ general instructional quality and classroom management were stronger when they were implementing their typical practices.

There was a significant main treatment effect on the number of key content ideas provided by students in their written essay (γ010 = .52, SE = .19, p = .01). The Hedges’ g effect size estimate was .31. The conditional model accounted for 78% of the between-class variance and 50% of between-teacher variance. Conversely, the effect of treatment on students’ use of supporting details of key content in writing was not significant (γ01 = .53, SE = .34, p = .14). The model-adjusted mean number of supporting details was 5.46 for students in the TBL condition and 4.93 for comparison condition students. Hedges’ g was .16.

**Conclusions:**

The findings suggest some potential for TBL practices to improve student content knowledge in middle and high school social studies classes, but also indicate additional teacher and/or student supports may be needed, particularly in middle school, for effective implementation and the possibility of consistent improvement in student outcomes. The findings of this study provide direction for future research regarding elements of TBL that were challenging for teachers to implement, including extending student thinking during discussions and providing feedback on discussions. In addition, the 8th grade teachers faced challenges in maintaining their instructional quality and classroom management while learning to implement TBL.

**References**


http://dx.doi.org/10.1080/07380569.2012.658347


http://dx.doi.org/10.1177/00222194070400020601


http://dx.doi.org/10.3102/00346543069003315


