Improving the Historical Knowledge and Writing of Students With or At Risk for LD

Susan De La Paz, PhD1, and Daniel R. Wissinger, PhD2

Abstract

In this study, we explored the potential of two forms of discussion (disciplinary vs. traditional) for 39 sixth- and seventh-grade students with or at risk for learning disabilities (LD), before writing historical arguments. Nine teachers who led small group discussions in six heterogeneous social studies classrooms implemented the intervention. Students who were involved in disciplinary discussions (n = 19) scored statistically higher than their peers who engaged in traditional discussions (n = 20) on a measure of historical knowledge (partial η² = .23); they also wrote essays with better persuasive quality (partial η² = .43) and greater evidence of historical thinking (partial η² = .40). A delayed posttest delivered 8 weeks after instruction ended revealed that students in the experimental condition continued to write in more historically sophisticated ways than did students in the comparison condition (partial η² = .19). Challenges, however, remain for struggling learners who must now meet basic and advanced disciplinary literacy goals.

Keywords

students with and at risk for LD, discussion, historical writing

Adopted by 43 states across the nation, the Common Core State Standards (CCSS) for English Language Arts and Literacy in History/Social Studies (Council of Chief State School Officers, 2010) are intended to create shared responsibility for teachers to incorporate literacy instruction in all academic subjects. Moreover, the College, Career, and Civic Life (C3) Framework for Social Studies State Standards emphasizes the importance of students working with historical evidence by reading primary source documents and communicating their conclusions in writing (National Council for the Social Studies, 2013). Together, the CCSS and C3 framework outline ambitious expectations for adolescents to acquire disciplinary literacy skills along with conceptual knowledge about history and the social studies. These are fundamental shifts from long-held views that reading and writing are the sole responsibility of English teachers (De La Paz et al., 2016; Shanahan & Shanahan, 2012).

While the C3 framework provides guidance on a broad range of practices, the development of discipline-specific reading and writing skills is outlined in the CCSS. For example, according to the CCSS, college- and career-ready students are able to “construct effective written arguments and convey intricate or multifaceted information,” as well as to support their ideas with “relevant evidence” (Council of Chief State School Officers, 2010, p. 3), which in history means choosing and establishing its relevance, in part, by explaining how that evidence supports a claim and by evaluating the source from which it arises—all in the context of an argument. To do this, students must “identify aspects of text that reveal an author’s point of view” (CCSS.ELA-Literacy.RH.6-8.8) as they read and “compare the point of view of two or more authors for how they treat the same topics, including which details they emphasize in their respective accounts” (CCSS.ELA-Literacy.RH.9-10.6). These expectations illustrate fundamental shifts from traditional expectations that students read for understanding and write for the purpose of summarizing (Bradford, 2011).

Given the consistency with which struggling learners (see Note 1) labor to meet grade-level expectations (National Center for Educational Statistics, 2012), many authors warn that increased demands on these youth will further frustrate learners as well as their teachers (Bulgren, Graner, & Deshler, 2013). Struggling learners often fail to master
basic, or domain general, reading and writing skills (Graham & Perin, 2007; Jitendra & Gajria, 2011) and have difficulty learning academic content in comparison with their normally achieving peers. In social studies, this has been traced to deficits in verbal learning and memory (Mastropieri & Scruggs, 2010) and limitations in being able to consider multiple perspectives (Bouck, Okolo, Englert, & Heutsche, 2008). Further cognitive and socioemotional delays constrain some students’ ability to differentiate time and develop historical empathy (Okolo & Ferretti, 1996). Many students with learning disabilities (LD) present difficulties differentiating relevant from irrelevant details (Bulgren, Deshler, & Lenz, 2007), which is critical in history and the social studies. These problems are likely compounded in middle and high school, where content area learning incorporates complex disciplinary standards (Moje, 2008). Because of these difficulties, there is a continuing need to identify instructional practice that closes the gap between struggling and typical learners, facilitates the development of both general and disciplinary literacy, and fosters greater acquisition of subject matter knowledge.

One way to simultaneously improve students’ knowledge about the past and apprentice them in the norms and discourse of an academic community is to infuse instruction with opportunities to engage in collaborative debate and discussion (Ferretti, MacArthur, & Okolo, 2001; MacArthur, Ferretti, & Okolo, 2002). However, research reveals that the skills needed to engage in disciplinary forms of dialogue do not come naturally, because students are uncertain about the goals and norms of argumentation in the disciplines and how these practices differ from other types of argumentation (Felton and Kuhn, 2001; Kuhn, Shaw, & Felton, 1997; Sampson, Grooms, & Walker, 2010). While there exists numerous methods and frameworks for facilitating the development of discipline-specific argumentation skills (Driver, Newton, & Osborne, 2000; Reznitskaya et al., 2001; Toulman, 1958), much of the work to date has concentrated on the processes involved in scientific argumentation, which differ in several important ways from the unique processes of historical argumentation. Arguments in the sciences tend to be rhetorical, where the objective is to persuade or convince others of a claim, providing them with reasons for accepting new or different views (Driver et al., 2000). In history, however, definitive claims may not exist (Carr, 1954; Shemilt, 1983), so argumentation becomes dialectical and requires that interlocutors consider, debate, and resolve differences in opinion about multiple explanations for events, trends, and controversies about the past (Wineburg, 1991). This collection of propositions forms an argument that contains a distinct structure and organizational scheme and serves as a critical standard for practicing historians.

Walton, Reed, and Macagno (2008) created a framework originally designed for philosophers that centers on dialectical argumentation and seems to provide a comprehensive tool for teaching students to construct and evaluate historical arguments. Within their dialectical framework, Walton and colleagues list argumentation schemes, or common forms of argument, with an accompanying set of critical questions that can be applied to specific circumstances and used to test the strength of an argument or the premises on which it is based (Macagno & Konstantinidou, 2013). Two particular schemes and their critical questions are consistent with processes used by expert historians to investigate the past, are present in the historical writing of middle and high school students (De La Paz, Ferretti, Wissing, Yee, & MacArthur, 2012), and align with goals outlined in the CCSS. The argument from expert opinion is one, providing a model for questioning the author of a document, examining personal reliability, and determining whether his or her assertions are consistent with evidence and what other experts have stated. Another scheme, called argument from consequences, prompts individuals to consider (a) the causal nature (cause–effect) of certain actions or inaction; (b) the evidence in the form of facts, data, or support that these consequences will or will not occur; and (c) whether these consequences are actually good or bad. We believed that these two schemes and their accompanying critical questions could be integrated into traditional instructional activities, such as discussions, and would function as a cognitive tool for scaffolding the development of arguments that were historical in nature, based on earlier findings by De La Paz et al. (2012) in which the more capable writers used these and similar strategies to help contextualize and corroborate evidence in their arguments.

Prior research suggests that introducing students to argument schemes like those described in Walton and colleagues’ (2008) framework not only facilitate the organization and retrieval of argument-relevant information during discussions (Felton & Herko, 2004; Nussbaum & Edwards, 2011) but also allow them to apply newly attained argumentation structures to their writing (Reznitskaya et al., 2001; Song & Ferretti, 2013). Together, these findings led us to investigate whether explicit instruction on historically related argument schemes and their accompanying critical questions could (a) function as a cognitive scaffold for younger middle school students to engage in sophisticated but structured argumentative discussions about controversial topics in history and (b) help them develop greater historical knowledge and craft more sophisticated written arguments.

**Efficacy of Teaching Schemes and Critical Questions**

In our original study, 151 sixth and seventh graders from a suburban district in the northeastern United States were either randomly assigned to an experimental condition—based on Walton and colleagues’ (2008) argument from expert opinion and argument from consequences schemes
Students then worked in small groups in the experimental or comparison discussions (see Wissinger & De La Paz, 2016). Finally, students independently wrote a historical argument, using their document sets, notes, and a version of De La Paz and Graham’s (1997) DARE heuristic for including important elements in their essays (i.e., develop a stance about the controversy, add evidence from the documents to support your stance, rebut arguments from the other side, and end by restating your stance). Our initial findings indicated that students who were randomly assigned to the experimental condition significantly outperformed their peers who were randomly assigned to the comparison condition on measures of historical knowledge (partial $\eta^2 = .44$), and historical writing (partial $\eta^2 = .45$). Follow-up analysis indicated that students who learned schemes and critical questions wrote historical essays that contained greater levels of substantiation (partial $\eta^2 = .24$) and more sophisticated rebuttals (partial $\eta^2 = .41$). Analysis of transcribed discussions aligned with these findings, demonstrating that experimental students provided explicit references to schematic language, substantiated claims, and consistently discussed limitations to counterevidence (i.e., rebuttals) more effectively than did students in the comparison condition. A delayed posttest administered 8 weeks later showed that experimental students continued to write with significantly greater levels of historical sophistication than did their peers in the comparison condition (partial $\eta^2 = .34$). The findings also revealed that students in both conditions were equally proficient on measures of persuasive writing quality and earned above-average to high scores for their responses on open-ended reading comprehension questions, thereby echoing the value of discussion for promoting a range of basic literacy learning outcomes.

Consistent with our original study, we tentatively hypothesized here that learning argumentative schemes and critical questions and engaging in disciplinary dialogue would also enhance struggling learners’ ability to learn more historical content and write historical arguments that were more sophisticated as compared with struggling learners who engaged in traditional forms of discussion; however, we were uncertain if this would be true, as the field continues to find differences in learning outcomes for students with or at risk for LD and for their peers who do not struggle with learning. In addition, and consistent with our original study, we predicted that struggling learners’ performance on generic outcome measures (i.e., reading comprehension, persuasive quality, and essay length) would be comparable across conditions because struggling learners in both conditions were exposed to the same historical controversies, a heuristic for including text structure in their arguments, and opportunities to discuss primary source documents in the same three historical investigations.

### Purpose and Research Questions

The purpose of this study was to examine the outcomes for students with and at risk for LD from the larger randomized control trial. This is important because an estimated 25% of students who participate in subject matter instruction in middle and high school struggle with reading and writing (Fuchs, Fuchs, & Compton, 2004) and because data from the National Longitudinal Transition Study–2 indicate that secondary students with disabilities, on average, earn 72% of their high school credits in general education classrooms, with students with LD earning the majority of these credits in English and social studies (Newman et al., 2011). Unfortunately, we lack adequate information about how these learners respond to the types of disciplinary forms of instruction called for in the CCSS for Literacy in History/Social Studies. If Walton and colleagues’ (2008) dialectical framework provides a promising means for facilitating discussions that teach students how to analyze primary and secondary sources (CCSS.ELA-Literacy.RH.6-8.1), identify aspects of text that reveal an author’s point of view (CCSS.ELA-Literacy.RH.6-8.8), and write arguments focused on discipline-specific content (CCSS.ELA-Literacy.WHST.6-8.1), it is important to understand whether it is equally effective for struggling learners who are a natural part of inclusive social studies classrooms. We therefore propose the following research questions:

**Research Question 1**: What are the effects of two forms of discussion (traditional vs. disciplinary) on struggling learners’ performance on general learning measures after reading and writing about controversial topics in history?

**Research Question 2**: Is there a relationship between the types of discussions in which struggling learners engage and the persuasive quality, historical thinking, and length of their writing?

### Methods

#### Participants

About half of the 39 students in this retrospective analysis ($n = 19$) were identified with specific LD (SLD) in reading and/or writing and received special education services through an individualized education program. The other half ($n = 20$) received Tier 2 interventions in the areas of
reading and writing, as they were considered at risk for academic failure. Students receiving Tier 2 interventions were consistently below (i.e., based on 8–10 weeks of progress monitoring with curriculum-based measures) and nonresponsive to Tier 1 interventions. Because of this heterogeneity, we hereafter use the term struggling learners to define our participants.

In addition, there were 19 struggling learners in the experimental group (12 with SLD and 7 receiving Tier 2 interventions) and 20 in the comparison group (12 with SLD and 8 receiving Tier 2 interventions). As in our larger study (Wissinger & De La Paz, 2016), all struggling learners were randomly assigned to either condition. To ensure that students in the two conditions were comparable with respect to academic ability, we compared their scores on the written expression portion of the Wechsler Individual Achievement Test–Third Edition (WIAT-III), the Gates-MacGinitie Reading Test–Fourth Edition (GMRT-4; MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2002), and state-administered reading assessment and their social studies grades for the first two semesters of the school year. A series of one-way analyses of variance revealed no significant academic differences between struggling learners in the two conditions (see Note 2). Chi-square goodness-of-fit tests also showed that students were comparable in terms of disability status, grade level, gender, ethnicity, and eligibility for free/reduced-price lunch prices. Table 1 lists demographic and academic characteristics by condition.

Before performing our analysis, we completed an a priori power analysis in G Power (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the probability of detecting a meaningful effect with a sample size of 39 students. Effects sizes from the literature on historical literacy research with struggling learners ranged between 0.11 and 0.80. The results of the analysis indicated that our sample size was sufficient for a multivariate analysis of covariance (MANCOVA) with two levels and three dependent variables based on an alpha of 0.05, a power of .90, and large ($f = 0.40$) to medium ($f = 0.25$) effect sizes.

**Table 1.** Demographic and Academic Characteristics by Condition Prior to Instruction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment</th>
<th>Comparison</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic characteristics, n</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification</td>
<td>19</td>
<td>20</td>
<td>.810</td>
</tr>
<tr>
<td>Students with LD</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Tier 2</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td>.657</td>
</tr>
<tr>
<td>Sixth</td>
<td>10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Seventh</td>
<td>9</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>.751</td>
</tr>
<tr>
<td>Female</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>12</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td>.443</td>
</tr>
<tr>
<td>Caucasian</td>
<td>14</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Eligible for free/reduced-price lunch</td>
<td>15</td>
<td>16</td>
<td>.854</td>
</tr>
<tr>
<td><strong>Academic characteristics, M (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing: WIAT-III</td>
<td>89.75 (12.1)</td>
<td>90.40 (15.1)</td>
<td>.712</td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMRT-4</td>
<td>5.17 (0.29)</td>
<td>5.26 (0.31)</td>
<td>.580</td>
</tr>
<tr>
<td>PSSA</td>
<td>1,175.79 (196.11)</td>
<td>1,211.55 (229.77)</td>
<td>.316</td>
</tr>
<tr>
<td>Performance level</td>
<td>Basic</td>
<td>Basic</td>
<td></td>
</tr>
<tr>
<td>Social studies: Grades (first quarter)$^a$</td>
<td>2.47 (0.61)</td>
<td>2.60 (0.82)</td>
<td>.591</td>
</tr>
</tbody>
</table>

Note. The $p$ values for demographic characteristics are based on chi-square goodness-of-fit models. The $p$ values for academic characteristics are based on analysis-of-variance models. WIAT-III = Wechsler Individual Achievement Test–Third Edition (essay composition subtest, scaled score); GMRT-4 = Gates-MacGinitie Reading Test–Fourth Edition (grade-equivalent score); PSSA = Pennsylvania System of School Assessment (reading scaled score).

$^a$Social studies grades were based on a 4-point scale, where 4.0 = A, 3.0 = B, 2.0 = C, and 1.0 = D.
disabilities were provided instruction in subject matter classes (e.g., science, social studies, language arts), which were co-taught by classroom and special education teachers. Classroom teachers reported that special education teachers worked with them during the regular school year to make modifications to tests (oral administration of tests, extended time), adapted reading assignments, and often supported students with disabilities as well as other struggling learners to complete assignments in subject matter areas outside regular class time. For the investigation, the six classroom teachers, two retired elementary school teachers, and the second author (then a special education teacher in the district) delivered instruction. Teaching experience among the group ranged from 3 to 35 years (\( M = 19.06, SD = 12.21 \)). All teachers held bachelor degrees, with five possessing a master’s degree at the time of the study.

**Instructional Procedures**

Three historical topics were chosen for instruction: Indian removal, the Mexican-American War, and the Gulf of Tonkin incident. Materials for each historical investigation included document sets that contained two primary sources, a secondary source with background information, and a question that related to the historical controversy. All three historical topics were explored across 15 days, or five 40-min sessions per week. Instruction across each 5-day sequence centered on three components: building background knowledge, discussion, and argumentative writing. Classroom teachers delivered instruction to their respective students in a whole-class setting on Days 1, 2, and 5. On Days 3 and 4, students transitioned to small groups (five to seven students) to participate in experimental or comparison discussions.

**Building background knowledge.** The first 2 days of each investigation sequence was used to develop students’ background knowledge. Classroom teachers followed scripted lesson plans and presented information through PowerPoint presentations to whole classrooms. Students were provided with document sets and response packets to record contextual information about the historical event. On the first day, teachers introduced the historical question and provided background information about major characters, events, and trends that occurred in the United States during the period. Teachers used the second day to examine primary source documents, highlight important features, and prompt students to record notes in response packets.

**Discussion.** Days 3 and 4 of each instructional sequence focused on group discussion. Students in each classroom participated in one of four discussion groups (two experimental and two comparison); in total, over the course of the study, nine teachers facilitated the four small group discussions. To control for teacher effects, teachers rotated conditions and groups at the end of each historical investigation.

**Experimental condition.** Students in this condition learned the argument from expert of opinion and argument from consequences schemes and a series of critical questions that accompanied each scheme in their small groups. Critical questions for the argument from expert of opinion were as follows: (a) Is the author a reliable source? (b) Is the information that the author is stating similar or different from what other authors are stating? and (c) Is what the author is stating based on sound evidence? Questions that accompanied the argument from consequences were as follows: (a) What are the good/positive consequences that are likely to happen if we follow through with the decision? and (b) What are the negative consequences that are likely to happen if we follow through with the decision? The language in the five critical questions was modified to align with historical authors and language in the source documents for each investigation.

The experimental instruction leveraged argumentation schemes and critical questions as cognitive scaffolds for students to make more informed decisions about historical controversies and gain a better understanding of each primary source. Teachers also emphasized that schemes and critical questions were tools to enable students to identify limitations in the “other side” argument and promote their ability to write rebuttals. To do this, teachers modeled how to use schemes and critical questions to examine each source in the first investigation, while gradually turning over more responsibility to students in subsequent investigations. As students progressed into the second and third (and final) historical investigations, teachers were coached (by the second author) to serve as facilitators of discussions, prompting students when necessary, redirecting, and correcting misconceptions when they occurred. Discussions were concluded with a review, a debriefing about notable student findings, and questions.

**Comparison condition.** Students in this condition also discussed the historical controversies on the third and fourth days of each week in small groups. As in the experimental condition, these students engaged in six discussions, for about 240 min across topics. However, students in this condition responded to a preexisting set of comprehension questions that were developed by the district to promote understanding of nonfiction and informational text. The questions prompted students to identify major historical actors, determine the author’s purpose and position, identify main or big ideas, and record details that support main or big ideas. These questions provided a clear comparison to those used in the experimental condition and had an additional advantage that students were familiar with them. As in the experimental condition, comparison students discussed one document per day for 2 days per week.
Argumentative writing. Students in both conditions wrote historical arguments on the final day of each sequence in their normal, large class setting. Similar to Days 1 and 2, the classroom teachers independently facilitated instruction on Day 5, first providing a brief overview of the investigation and a review of the historical controversy. Students were encouraged to use document sets and notes in response packets, and they were given 40 min to complete their written responses. Students with disabilities and those with 504 plans were provided with the specified accommodations outlined by the school district. To clarify the argumentative writing process, students in both conditions were introduced to a modified version of De La Paz’s (1997; De La Paz & Graham, 2005) DARE heuristic, and teachers modeled how to create a DARE graphic organizer and transfer information from response packets to their essays for Investigation 1: Indian removal. After students’ constructed argumentative essays for Investigation 2: Mexican-American War, support materials (i.e., DARE, graphic organizers, and guided response questions) were faded from instruction to promote greater independence.

Training and Fidelity of Implementation

Participating teachers engaged in three 1-hr training workshops before the study began, with a 1-hr review before each historical investigation. Training focused on the use of instructional materials and procedures for delivering instruction in whole group settings on Days 1, 2, and 5 of the sequence and to discussion groups on Days 3 and 4. Review sessions were used to introduce each new historical investigation and for discussion and clarification.

To ensure that instruction was implemented in a manner consistent with the scripted lessons, independent observers randomly observed and record notes 18 times across Days 1, 2, and 5 and randomly selected audio recordings of discussion sessions (one-third of total) that occurred on Days 3 and 4. The observers used a protocol with which they rated teachers’ implementation of the intervention for fidelity (i.e., the presence of the lesson components, procedures for building background knowledge, discussion and analysis, and writing) and quality (i.e., ratings of disagree strongly, disagree, mutual, agree, or strongly agree on a range of 10 instructional quality variables specific to the intervention). Ratings for fidelity and quality were averaged across items and across observations to yield a composite score for each teacher. Cronbach’s alpha between raters ranged from .92 to .97 ($M = .92, SD = .04$), and the overall quality of the lessons recorded was .92 (on a 100-point Likert-type scale, with 100 indicating that observers strongly agreed that instruction was being implemented as prescribed).

Measures

Five researcher-developed measures were used to compare outcomes among struggling learners in the two conditions. The measures represented forms of student knowledge in two areas: general learning and writing.

General learning. Two measures were used to examine general learning. These included tests of historical knowledge and reading comprehension.

Historical knowledge. We created a 12-item multiple-choice test on historical content that was delivered before and after instruction. Questions in the measure were constructed by classroom teachers and the second author, varied in difficulty, and derived from the primary sources that students examined in historical investigations and the pretest/posttest. Items were reviewed and further validated by a history teacher who read the primary sources and examined the questions generated for content and construct validity. Bivariate correlation with social studies grades, $r(151) = .524, p \leq .001$, suggested that the test had adequate criterion validity and may be considered representative of students’ knowledge about history. Historical knowledge was the total number of correct answers (range = 0–12 points). Internal consistency reliability was calculated through Cronbach’s alpha and came to .87 at pretest and .93 at posttest.

Reading comprehension. We developed a 10-item test to assess students’ factual comprehension of two primary and one secondary source document about life in America in the 1950s, which they were asked to read independently at pre- and posttest. Criterion validity was established through bivariate correlations—GMRT-4, $r(151) = .672$, $p \leq .001$, and Pennsylvania System of School Assessment (PSSA) reading scores, $r(151) = .633, p \leq .001$—which indicated that the test was representative of students’ comprehension. Students’ responses were scored on a scale of 0 to 3, where 0 = an insufficient/inaccurate response, 1 = an incomplete answer (e.g., indicating either a misunderstanding or no details), 2 = a partial answer to the task (e.g., indicates some awareness or at least one text-based detail), or 3 = a complete answer (e.g., correct with text-based support). Scores by independent raters were averaged, and agreement was tested with Cronbach’s alpha and found to be .91.

Writing. Three measures were used to evaluate students’ writing: historical writing, persuasive quality (based on standardized writing assessment criteria), and essay length. The first two measures were scored with an analytic scoring method that involved rating students’ historical writing on several traits.
**Analytic Rubric of Historical Writing.**  

<table>
<thead>
<tr>
<th>Substantiation</th>
<th>Perspective Recognition</th>
<th>Contextualization</th>
<th>Rebuttal</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 (a) Facts/quotes are explained AND linked explicitly to a conclusion. AND Explanations are consistently accurate.</td>
<td>(a) Evaluates the authors’ perspectives (e.g., discusses reliability/trustworthiness). OR (b) Reconciles multiple authors’ perspectives (e.g., compares the ideas in the documents).</td>
<td>(a) Integrates background information and evidence from the documents in an explanation or conclusion. OR (b) Uses background information and evidence from the documents together to draw a conclusion or make an inference.</td>
<td>Opposing side claims are clearly presented and drawn from the documents. Writer does not simply consider an opposing side but offers an explicit rebuttal, evaluation of evidence, or reconciliation of opposing views.</td>
</tr>
<tr>
<td>3 (a) Facts/quotes are presented without explanation but are explicitly linked to a conclusion. OR (b) Facts/quotes are explained but not explicitly linked to a conclusion. AND At least one explanation is accurate.</td>
<td>(a) Describes author’s perspective in a way that recognizes text as the author’s point of view. Note. The author’s name does not have to be mentioned.</td>
<td>(a) Describes background information alongside the evidence from the documents without explicitly connecting them.</td>
<td>Opposing side claims are presented and drawn from the documents, but there is no explicit justification for choosing one side over the other. The author clearly chooses a position, but stops short of explaining why his/her position follows from what is presented (i.e., opposing sides elaborated but not explicitly rebutted or reconciled).</td>
</tr>
<tr>
<td>2 (a) Facts/quotes presented without explanation (may implicitly support a conclusion that is stated). OR (b) Facts/quotes are presented with inaccurate explanations throughout.</td>
<td>(a) Mentions the author (e.g., &quot;According to Lynch, &quot; &quot;The author says&quot;).</td>
<td>(a) Mentions background information in the documents (example).</td>
<td>(a) Opposing side claims are not drawn from the documents. OR (b) Opposing side claims are distinguished or acknowledged but not elaborated on.</td>
</tr>
<tr>
<td>1 (a) Minimal evidence. OR (b) Irrelevant evidence. OR (c) Transcription of document (and nothing else) Note. There may be a claim.</td>
<td>(a) Presents evidence from documents as student’s own perspective. OR (b) Treats documents as authoritative (e.g., &quot;Document 1 says,&quot; &quot;It says in the document&quot;).</td>
<td>(a) Minimal background information mentioned. OR (b) Student uses anachronisms (e.g., makes a chronological mistake or uses information from another period without noting the different era).</td>
<td>(a) No mention of opposing side claims. OR (b) No clear overarching position, so no clear treatment of opposing sides on the issue of the prompt.</td>
</tr>
</tbody>
</table>

**Historical writing.** An analytic rubric was created to capture the extent to which students exhibited historical thinking in their writing. Essays were scored on the basis of four components: substantiation, perspective recognition, contextualization, and rebuttal. The analytic quality of each of the four components was based on a 3-point scale (De La Paz et al., 2016; Monte-Sano & De La Paz, 2012; see Table 2). We established criterion validity through bivariate correlations with the WIAT-III, $r(151) = .480, p \leq .001$, and students’ social studies grades, $r(151) = .327, p \leq .001$, which showed a significant relationship between students’ writing ability and their knowledge of social studies. The second author trained two raters in two 2-hr sessions. During this time, benchmark essays for each level on the rubric were shared; scoring distinctions were discussed; and raters practiced scoring students’ writing produced during the three historical investigations. The two raters then scored all pre- and posttest essays. The average of their scores was used in analyses of each element. Interrater agreement as calculated by Cronbach’s alpha was 0.86 (range = .81–.89) across all elements.

**Persuasive quality.** Each essay was separately scored with the PSSA persuasive writing rubric (Pennsylvania Department of Education, 2010b). This standardized persuasive quality index captures five dimensions of effective writing: focus, content, organization, style, and conventions.
The quality of the five dimensions is based on a 4-point scale (i.e., PSSA; Pennsylvania Department of Education, 2010a). We use bivariate correlations with WIAT-III scores to establish criterion validity, \( r(151) = .548, p \leq .001; \) the relationship seemed to suggest that persuasive quality scores were representative of students’ writing ability. In two 1-hr sessions, the second author trained two separate teachers in the district who had extensive experience scoring sixth- and seventh-grade students writing. The raters then scored all of the essays in the data set independently. Cronbach’s alpha was .93 (range = .89–.96 across the four dimensions).

**Essay length.** All essays were also scored on the total number of words written. Length was determined by counting the total number of words in students’ essays. The second author counted the total number of words in students’ essays at pre- and posttest. An elementary school teacher in the same school district completed reliability checks. Cronbach’s alpha was computed on a randomly selected pool (15% of the papers) for total number of words written. Reliability for this measure was .98.

**Data analysis.** We first conducted one-way MANCOVA on the two general learning measures (historical knowledge and reading comprehension) and the three writing measures (historical writing, persuasive quality, and essay length). The independent variable was treatment condition. If the main effects for treatment condition were significant, univariate tests were used to identify which knowledge measures were statistically different by condition. To verify the conservative nature of the \( F \) test for condition and to improve the statistical detection of treatment effects, we used pretest scores on dependent measures as covariates (Huck & McLean, 1975). Covariates met assumptions of randomization, linearity, and homogeneity of regression slopes for each outcome measure as described here. Means and standard deviations for historical knowledge, reading comprehension, historical writing, persuasive quality, and essay length by condition are presented in Table 3.

### Results

**Research Question 1:** What are the effects of two forms of discussion (traditional vs. disciplinary) on struggling learners’ performance on general learning measures after reading and writing about controversial topics in history?

To answer the first research question, we ran a one-way MANCOVA on historical knowledge and reading comprehension to determine whether differences in students’ general learning were contingent on discussion condition. The covariate was students’ pretest scores. Levene’s test showed that the assumption of homogeneity of variances was not violated for either measure, with \( p = .65 \) for historical knowledge and \( p = .61 \) for reading comprehension. Box’s M test indicated that variance–covariance matrices were not different across cells, \( F(3, 271286.9) = 1.22, p = .765. \) Finally, the Shapiro-Wilk’s test indicated that historical knowledge and reading comprehension scores were normally distributed for each condition (\( p > .05 \)).

The Wilks’s lambda revealed significant multivariate main effects for the experimental condition, after controlling for pretest scores, Wilks’s \( \lambda = .758, F(3, 37) = 5.60, p = .008, \) partial \( \eta^2 = .24. \) Follow-up univariate tests

<table>
<thead>
<tr>
<th>Measure</th>
<th>Treatment Pretest</th>
<th>Treatment Posttest</th>
<th>Comparison Pretest</th>
<th>Comparison Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>General learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading comprehension</td>
<td>— —</td>
<td>7.85 1.09</td>
<td>— —</td>
<td>7.68 1.16</td>
</tr>
<tr>
<td>Historical knowledge</td>
<td>5.10 1.37</td>
<td>9.20 1.24</td>
<td>5.58 1.92</td>
<td>7.95 1.39</td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persuasive quality</td>
<td>10.89 3.77</td>
<td>14.65 3.25</td>
<td>11.1 3.14</td>
<td>13.37 2.78</td>
</tr>
<tr>
<td>Historical writing</td>
<td>3.11 3.13</td>
<td>8.50 2.26</td>
<td>3.5 2.89</td>
<td>6.1 2.66</td>
</tr>
<tr>
<td>Substantiation</td>
<td>1.21 0.85</td>
<td>2.35 0.59</td>
<td>1.20 0.69</td>
<td>1.89 0.66</td>
</tr>
<tr>
<td>Perspective recognition</td>
<td>0.57 1.12</td>
<td>2.20 0.89</td>
<td>0.75 1.07</td>
<td>1.42 1.12</td>
</tr>
<tr>
<td>Contextualization</td>
<td>0.63 0.76</td>
<td>1.40 0.88</td>
<td>0.65 0.75</td>
<td>1.00 0.58</td>
</tr>
<tr>
<td>Rebuttal</td>
<td>0.68 1.16</td>
<td>2.35 0.87</td>
<td>0.90 1.21</td>
<td>1.74 0.99</td>
</tr>
<tr>
<td>No. of words</td>
<td>100.05 78.04</td>
<td>153.53 46.70</td>
<td>81.16 50.68</td>
<td>165.3 55.18</td>
</tr>
<tr>
<td>Delayed posttest</td>
<td>7.29* 1.65</td>
<td></td>
<td>5.9 2.87</td>
<td></td>
</tr>
</tbody>
</table>

Note. The \( p \) values are based on follow-up series of univariate tests where pretest scores served as the covariate. A dash (—) indicates that no pretest was administered.

\*\( p < .05 \) (treatment vs. control).
showed significant differences favoring struggling learners who learned schemes and critical questions on the historical knowledge measure, $F(1, 36) = 10.49, p = .003$, partial $\eta^2 = .23$, but not for reading comprehension. The results indicated that while students with and at risk for LD in the two conditions began the investigation with comparable reading levels, students who learned to use historically related schemes and critical questions during discussions gained more historical knowledge than did their struggling peers who used traditional comprehension questions during discussions. The findings also showed that struggling learners in both conditions achieved comparable outcomes on the open-ended reading comprehension test administered after instruction.

Research Question 2: Is there a relationship between the types of discussions in which struggling learners engage and the persuasive quality, historical writing, and length of their writing?

To address the second research question, a one-way MANCOVA was performed on measures of historical writing, persuasive quality, and essay length. The covariates were corresponding pretest scores. Levene’s test revealed that the assumption of homogeneity of variances was not violated for the dependent measures ($p’s > .05$), and Box’s M test indicated that variance–covariance matrices were similar across cells, $F(6, 9841.7) = 3.65, p = .766$. The Shapiro-Wilk’s test showed that measures were normally distributed for each condition ($p > .05$).

Main effects were significant for the combined Writing Measures × Discussion condition, after controlling for students’ writing ability at pretest, Wilks’s $\lambda = .67, F(4, 34) = 5.32, p = .004$, partial $\eta^2 = .33$. Univariate tests revealed main effects by condition for historical writing, $F(4, 34) = 5.76, p = .001$, partial $\eta^2 = .40$, and persuasive quality, $F(4, 34) = 6.49, p = .001$, partial $\eta^2 = .43$, but not for length of essays. Together, the findings for writing showed that struggling learners who used schemes and critical questions to engage in historical forms of dialogue with their peers crafted essays that demonstrated greater evidence of historical thinking and were judged to be more persuasive but were comparable in length to those produced by comparison students.

**Historical Writing**

To better understand what components of struggling learners’ historical writing was most affected by the types of discussions they engaged in, we ran a separate set of analysis of covariance tests on each of the four elements of the analytic rubric of historical writing (i.e., contextualization, substantiation, perspective recognition, and rebuttal). The independent variable was condition, with pretest scores on each element serving as the covariates. The analysis revealed that as compared with struggling learners in the comparison condition, students who learned schemes and critical questions included significantly more sophisticated levels of substantiation, $F(5, 33) = 3.27, p = .017$, partial $\eta^2 = .33$, recognized authors’ perspectives more effectively, $F(5, 33) = 4.64, p = .003$, partial $\eta^2 = .41$, and crafted more mature rebuttals, $F(5, 33) = 3.25, p = .017$, partial $\eta^2 = .33$, but were comparable in terms of degree of contextualization, after controlling for initial scores on this measure.

**Delayed Posttest**

Eight weeks after instruction ended and 1 to 2 weeks after completing the state’s annual high-stakes assessments, all sixth- and seventh-grade learners in the district wrote an end-of-the-year essay after completing a final historical investigation that focused on conflicting perspectives about the Battle at Gettysburg in July 1863. Because teachers were concerned about the need to prepare students for these exams, all instruction related to this project ended 8 weeks before our delayed posttest. Students received a traditional form of social studies instruction rather than the experimental or comparison condition that had taken place earlier in the year. Students read document sets that included a historical question, background information, and two primary source documents. Classroom teachers reviewed documents and background information in one 50-min class period, but students wrote arguments without discussing the documents or working in small groups. Grade-level teams including reading specialists and special education and classroom teachers scored each essay using the analytic rubric of historical writing.

Delayed posttests were available for 33 of the 39 struggling learners (16 experimental and 17 comparison students). Comparisons between the partial and full samples suggested that there were no significant differences in terms of gender, $F(1, 32) = 0.56, p = .816$, social studies grades, $F(1, 32) = 1.01, p = .364$, PSSA reading, $F(1, 32) = 0.56, p = .810$, GMRT-4, $F(1, 32) = 0.59, p = .784$, or WIAT-III essay composition, $F(1, 32) = 1.40, p = .254$. Table 2 presents the means and standard deviations for historical writing on the delayed posttest measure. We ran a one-way analysis of covariance to examine the effects of the discussion condition on the delayed posttest after adjusting for historical writing scores at pretest. The results revealed a significant effect for condition, $F(2, 31) = 3.46, p = .044$, partial $\eta^2 = .18$. Struggling learners in the experimental condition continued to write essays with more sophisticated historical reasoning than did struggling learners in the comparison condition approximately 2 months after instruction ended.
Discussion

The purpose of this study was to learn more about the use of a structured yet disciplinary form of discussion to support struggling students’ disciplinary learning. The current results extend prior findings with heterogeneous learners (Wissinger & De La Paz, 2016) by demonstrating that this form of instruction also benefits sixth- and seventh-grade students with and at risk for LD by leading them to learn more historical knowledge (partial $\eta^2 = .23$) than that of students who participated in traditional discussions. These results replicate those found in our larger randomized control trial and add to other findings that disciplinary discussions can enhance the learning of historical content of students who struggle with reading and writing in general education classrooms (MacArthur et al., 2002; Okolo et al., 2007). An added strength of the current study is that we were able to document, through an a priori power analysis, that 39 students with or at risk for LD from a larger randomized control trial and add to other findings that disciplinary discussions can enhance the learning of historical content of students who struggle with reading and writing in general education classrooms (MacArthur et al., 2002; Okolo et al., 2007). An added strength of the current study is that we were able to document, through an a priori power analysis, that 39 students with or at risk for LD from a larger randomized design formed an adequate sample to obtain statistical judgments about the impact of our intervention.

Our current results show that while struggling learners began the investigation with comparable writing skills (as demonstrated by scores on the WIAT-III), those who participated in disciplinary discussions wrote arguments that contained greater evidence of historical thinking than that of their peers who engaged in traditional forms of discussion. Partial $\eta^2$ was .40 for historical writing. In particular, essays written by these students evidenced greater levels of substantiation, authors’ perspectives, and more mature rebuttals. Partial $\eta^2$ was .33 for substantiation, .41 for authors’ perspectives, and .33 for rebuttal.

Furthermore, outcomes for struggling learners in the experimental condition as related to gains in historical writing remained significant on a delayed posttest measure given 8 weeks after instruction ended. Partial $\eta^2$ was .19 on the delayed posttest. The outcomes related to writing seemed to illustrate the power of our intervention for fostering disciplinary literacy among a diverse community of learners. While encouraging, however, these findings are not altogether surprising, given prior successful intervention work with students with and without LD in history (De La Paz, 2005; Ferretti et al., 2001; Gersten et al., 2006; MacArthur et al., 2002).

In addition, in contrast to the differential results on our disciplinary literacy measures, struggling learners in both conditions earned comparable scores on a general reading comprehension measure and on their length of essay, which were also consistent with our 2016 findings. In other words, on these measures, students’ domain-general reading and writing learning outcomes were comparable. In fact, we had anticipated this outcome, as instruction in the comparison condition included a defensible yet traditional form of discussion as well as efforts to teach students to plan before composing.

However, in contrast to our original study, which did not provide outcomes for different subgroups of students, in this study we found that the experimental instruction also resulted in significant improvement on our persuasive quality measure, again favoring students who were exposed to disciplinary discussions. The largest partial $\eta^2$ for the writing measures was recorded for persuasive quality (.43). This finding indicates that between the two groups of students, some elements in this study—likely, the text-structure heuristic DARE—were especially helpful for struggling learners versus comparison students in terms of developing a clear stance, supporting claims with evidence, and including other-side arguments and rebuttals in their writing. Whereas in our original study, all students demonstrated the same persuasive quality of essays post instruction, in this study, struggling learners received a specific disciplinary boost from the instruction.

Furthermore, to our knowledge, this study is the first to indicate that explicit instruction on historically related argument schemes and critical questions from Walton and colleagues’ (2008) dialectical framework can function as a cognitive scaffold for struggling learners to develop greater historical knowledge and write essays that evidence significantly higher generic (i.e., persuasive quality) and disciplinary (i.e., historical writing) qualities. Given general writing requirements in the CCSS for Grade 6–8 students to “write arguments focused on discipline-specific content” (CCSS.ELA-LITERACY.WHST.6-8.1) and disciplinary goals for students to “cite specific textual evidence to support analysis of source documents” (CCSS.ELA-LITERACY.RH.6-8.1) and “identify aspects of a text that reveal an author’s point of view” (CCSS.ELA-LITERACY.RST.6-8.4), our work seems to highlight a viable approach for teaching these skills in general education classrooms.

Thus, the major contributions that this study makes are that it demonstrates that students who are at risk for or who have LD are able to use specialized forms of discussion to learn important and ambitious content that they are expected to acquire via instruction in general education classrooms. Although it is fair to say these students do not always make gains as large as their normally achieving peers (more on this in the Limitations section), this population of students often do not measurable gains when participating in inclusive secondary settings. The fact that students in our study made gains in general and disciplinary outcomes is welcome because struggling writers need to make improvements in both types of writing. We now provide examples of discussion excerpts that facilitated this type of learning.

Illustrating the Effectiveness of Schematic Instruction

To illustrate the underlying mechanisms responsible for improvements to struggling learners’ historical knowledge
and writing, we included two randomly selected excerpts of the experimental group’s discussions. The group contained five students without disabilities (Kyra, Johnathan, Kaitlyn, Joe, and Nasir) and one student with LD who was also an English-language learner (Raul; note that all names are pseudonyms). The dialogue provides evidence for how students’ understanding of schemes and critical questions formed generalized patterns of historical thinking and reasoning that were evidenced in their writing. In what follows, we briefly highlight instances where students gathered evidence from sources and evaluated the credibility of information on the basis of who crafted the document (i.e., argument from expert opinion) and how students understood the interplay of human interaction and the constraints on human actions in terms of causing change (i.e., argument from consequences). It is equally important to note how the group’s utterances were interspersed with explicit references to words such as “reliability,” “their point of view,” “trust,” “evidence,” and “consequences.”

**Excerpt: Mexican-American War**

_T—So what was the battle about, what were they fighting about? Kyra—Land. Nasir—Basically Texas._

_T—the land right, about the border._

_T—The first question is can we trust what President Polk is saying about Mexico invading our territory? Is he a reliable source? Raul—not really, he’s a politician. . . . Kaitlyn—People are one-sided. And they’re going to tell a story completely from their side . . . oh they attacked us so let’s go attack them [reference to President Polk’s speech]. But Polk is telling us from our point of view [the United States], and they’re not going to automatically tell the truth. Johnathan—Unless they’re a reliable source. Kyra—Polk isn’t reliable, he’s lying. Just like Indian Removal with Jackson [Investigation 1], there was no evidence. Jonathan—Yeah, people just wanted that land. . . . T—Good, remember our discussion about Manifest Destiny. . . . This might help to explain our growing desire for the land. Raul—And Gibbons [Joshua Giddings] said that Mexico did not invade our territory. . . . Polk said that they didn’t._

_T—How many of you agree we should’ve went to war, and what about the consequences? Kyra—Nobody wants to go to war. Joe—I don’t, only because I think this could have been solved without fighting, they could have just talked with them, and talked through it. Kaitlyn—Yeah. War kills people. Joe—they might have an argument, but after the war they’re not going to be like “Hi best friend!” and we [the United States] are right next to them [Mexico]. It would be like getting into a fight with your next door neighbor. There will be bad consequences like we will see them every day and have to deal with them. Raul—Yeah, if you just killed their family, they’re going to be hating [hostile] on anyone who moves [relocates] into that territory._

_T—You’re on the right track. Think about it, after war occurs, then you have a lot of angry feelings. This might take years to resolve. Kyra—So it really doesn’t matter who wins and who loses, there’s going to be hostility._

_T—[however] if you are the United States, what might be the good consequences of going to war? Nasir—They would get land. I mean, Texas is a big state. Kyra—There might be resources there. Mining resources, gold, you might even find oil. Raul—[The] main thing is that they’ve got land. Kyra—Yes, we [the United States] could spread out, instead of being closed in, because we only had 10 . . . or 20 states at that time, and there was a lot of people._

The first series of interactions illustrates how students used the argument from expert opinion to draw conclusions about the primary source authored by President Polk. For example, Kaitlyn responds to the teacher’s question about whether they can trust what President Polk stated: “Polk is telling us from our point of view [the United States], and they’re not going to automatically tell the truth” (see Note 3). Referencing the investigation on Indian removal, Kyra extends the group’s sentiments of distrust for President Polk’s statement by arguing that the government was pursuing war because “people wanted that land.” This was a direct reference to the secondary source (Document 1) that discussed Americans’ belief in Manifest Destiny and the nation’s desire to expand. More important, perhaps, Raul recognizes that Representative Gidding’s point of view (Document 3) did not match that of President Polk, thereby illustrating his awareness of the need to compare multiple points of view. In response to these limitations, the group formed a consensus that it could not trust Polk’s message on the war with Mexico (Document 2).

In a similar manner, the second series of interactions evidenced how students used the argument from consequences to explore the interplay of intentional human action (e.g., President Polk) and constraints (e.g., geography, historical legacies, and political relationships) that contributed to the decision to go to war with Mexico. Students also developed counterfactuals (e.g., if the Americans do not go to war in Mexico, then . . . ) before drawing conclusions about whether the United States had a reasonable argument for going to war. Kaitlyn, for example, comments, “War kills people.” In response, Kyra extends this reasoning by discussing that, beyond the death of soldiers and families, “it really doesn’t matter who wins and loses . . . there’s going to be hostility” for all who relocate and for current residents of the disputed border area. Yet, Nasir, Kyra, and Raul list counterevidence for why President Polk and Congress would rationalize the decision to go to war (e.g., the possibility of oil, gold, and, more important, land that would allow the United States to expand). Together, this example shows how the group analyzed cause and consequence before concluding whether it was reasonable to enter into war.

**Summary.** The two excerpts provide a glimpse of the types of contributions that struggling learners provided to the historical discussion, as well as a comprehensive view of the types of interactions shared during heterogeneous group dialogue. Through explicit instruction, the critical questions became a cognitive tool that students used to examine and record inconsistencies between sources.
(Critical Question 2); question the reliability of statements made by historical authors (Critical Question 3); and evaluate cause, consequence, and craft counterfactuals about the reasons for historical events (Critical Questions 4 and 5). From these examples, we conclude that students’ knowledge of two common forms of argument, along with their critical questions, formed generalized patterns of thinking and reasoning that were later integrated in their writing.

Limitations and educational implications. Despite these insights, some limitations are also evident. One is that while students with and at risk for LD demonstrated improvements in outcomes related to historical learning and provided evidence that they were capable of engaging in disciplinary forms of discussion, these outcomes do not suggest full mastery in learning. To explain, some gaps in performance between normally achieving and struggling learners remained after instruction ended, as the effect sizes in the current analysis were smaller than those reported in Wissinger and De La Paz (2016) for gains in historical knowledge (differences in partial η² between normally achieving students and students with or at risk for LD on these measures: .21 for historical knowledge, .05 for historical writing, and .16 for historical thinking on the delayed posttest). Persistent gaps between students with and without LD are consistently reported in literature on teaching strategies for historical reasoning (De La Paz, 2005; De La Paz et al., 2016; Ferretti et al., 2001; Gersten et al., 2006; Hernandez-Ramos & De La Paz, 2009; Reisman, 2012). However, in this study, struggling learners also made gains in their general writing quality, coming up to the level of their normally achieving peers on this measure.

Finally, future work will need to address several limitations. First, while students demonstrated strong scores in the area of substantiation, researchers should focus on teaching younger adolescents how to interpret facts and quotes in their writing. Similarly, instruction should continue to emphasize that history is interpretive (Mink, 1987; Monte-Sano, 2008) and often explained differently by those who witnessed the same event. As a result, students must further learn that information taken from a source document (e.g., newspaper articles, diary excerpts, public addresses) contains the perspectives of the person who authored it, which must be recognized as so in their writing. The greatest challenge, perhaps, is to find ways to teach struggling and novice learners how to understand the contexts of certain historical periods and how the perspectives and actions of these individuals were shaped by that period (van Drie & van Boxtel, 2008). In short, the standards of the past often do not align with those of today; therefore, it is important to teach students how to wrestle with issues of causality, connections, and context when explaining the past.

Conclusions

There is a widely recognized need to identify teaching strategies in history and the social studies that will engage struggling learners and those without LD in their learning and equip them with an understanding of how experts in the discipline think about the past imaginatively and with integrity (Bulgren, Graner, & Deshler, 2013). Researchers in special education must continue their efforts to develop instruction that enables students who struggle most with learning to meet the same academic challenges as students without learning difficulties in contemporary history and social studies classrooms; as of yet, this goal has been difficult to realize (Buckley, 2005; van Hover & Yeager, 2003). However, the provision of disciplinary approaches to instruction is similar to that described by Wissinger and De La Paz (2016) and in this analysis appears to be one avenue for struggling learners to have access to meaningful learning opportunities in history and social studies classrooms.

Declaration of Conflicting Interests

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

Funding

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

Notes

1. Struggling learners included students with learning disabilities (LD) and students who were receiving tiered support in the school for reading and writing difficulties.
2. Struggling learners in the sample scored comparably to a national sample of students with disorders of written expression (M = 88.9, SD = 12.9; Wechsler Individual Achievement Test–Third Edition; Psychological Corporation, 2009).
3. Example of what Seixas (2015) refers to as historical perspective, or the understanding that any particular historical event or situation involves people who may have diverse perspectives.

References

Sampson, V., Grooms, J., & Walker, J. P. (2010). Argument-driven inquiry as a way to help students learn how to participate in


