



Language-rich discussions for English language learners

Jie Zhang^{a,*}, Richard C. Anderson^b, Kim Nguyen-Jahiel^b

^a College of Education and Behavioral Sciences, Western Kentucky University, United States

^b Center for the Study of Reading, University of Illinois at Urbana-Champaign, United States

ARTICLE INFO

Article history:

Received 17 February 2012

Received in revised form 5 December 2012

Accepted 10 December 2012

Available online 5 January 2013

Keywords:

Collaborative Reasoning
English language learners
Classroom discussion
Language development

ABSTRACT

A study involving 75 Spanish-speaking fifth graders from a school in the Chicago area investigated whether a peer-led, open-format discussion approach, known as Collaborative Reasoning, would accelerate the students' English language development. Results showed that, after participating in eight discussions over a four-week period, the CR group performed significantly better than the control group on measures of listening and reading comprehension. The CR group produced more coherent narratives in a storytelling task. The reflective essays written by the CR group were longer; contained more diverse vocabulary; and contained a significantly greater number of satisfactory reasons, counterarguments, and uses of text evidence. CR discussions also enhanced students' interest and engagement in discussions, perceived benefits from discussions, and attitudes toward learning English.

Published by Elsevier Ltd.

The purpose of the study reported in this paper was to investigate the effects of an approach to classroom discussion, *Collaborative Reasoning*, on the oral and written English development of Spanish-speaking English language learners. English Language Learners (ELLs), particularly those who are Spanish-speaking and from low-income backgrounds, rarely attain the same levels of achievement in reading and writing as native English speakers. The gaps become profound and persistent by fourth grade and beyond when the instructional focus shifts from word-level decoding skills to comprehension and writing (National Center for the Educational Statistics, 2009). According to the report of the National Literacy Panel on Language-Minority Children and Youth, the major impediment for ELLs' reading comprehension is their limited oral English proficiency (August & Shanahan, 2006).

A broad range of oral language skills, including vocabulary, syntax, morphology, listening comprehension, and oral narrative skills, has been found to predict reading skills of Spanish-speaking ELLs (e.g., Carlisle, Beeman, Davis, & Spharim, 1999; Miller, Heilmann, & Nockerts, 2006; Proctor, Carlo, August, & Snow, 2005). Despite the importance of oral English proficiency for ELLs, literacy instruction practices for ELLs too often feature individual seatwork and teacher-directed whole class instruction. In the teacher Initiate–student Response–teacher Evaluation (IRE) structure, typical of classroom discourse, ELLs are less actively engaged in academic tasks (e.g., Arreaga-Mayer & Perdomo-Rivera, 1996; Simmons, Fuchs, Fuchs, Mathes, & Hodge, 1995). Arreaga-Mayer and Perdomo-Rivera reported astonishing low levels of active oral engagement among Latino students with limited English proficiency, only 4% of the school day. Whole class instruction (54%) and independent seatwork (32%) were the predominant instructional formats, while small group activities were infrequent, only 2% of the day. When students are being instructed directly by the teacher in whole-class groups, they spend about 70% of

* Corresponding author at: Western Kentucky University, Gary Ransdell Hall 3088, 1906 College Heights Blvd. # 41031, Bowling Green, KY 42101-1031, United States. Tel.: +1 270 745 2933.

E-mail addresses: [jiez35@gmail.com](mailto:jiezh35@gmail.com), jie.zhang@wku.edu (J. Zhang).

their time passively watching and listening (Simmons et al., 1995). It is not surprising that, having few chances for extended use of English, ELLs show low academic engagement and lag behind in reading achievement.

Furthermore, in schools with large enrollments of ELLs, language tends to be treated as a ‘formal’ subject (e.g., phonics, vocabulary, grammar) with little opportunity for interactive language that is comprehensible, interesting, and relevant (Ellis, 1986; Krashen, 1982). Teachers of ELLs typically rely on simple tasks that require literal recall or, at most, limited inference. Fast-paced, low-level, question–answer routines limit students’ opportunities to talk, generate questions, and express extended ideas. The emphasis on phonics, spelling, accurate oral reading, proper English pronunciation, vocabulary lists, grammar, and literal comprehension has probably been exacerbated by the perception of schools that these emphases are necessary to prepare students for high-stakes examinations (Assaf, 2006). There is reason to fear that the regimen in today’s schools inhibits the language development of ELLs and may retard their cognitive development and undermine their motivation for school by taking the meaning and enjoyment out of learning (Gersten, 1996).

There is a small but growing body of research on effective literacy instruction for English language learners (see August & Shanahan, 2006). Existing research has several limitations. First, extensive oral English development for ELLs has been largely overlooked. In a review of studies of English language learners in the U.S., Genesee, Lindholm-Leary, Saunders, and Christian (2006) concluded that fewer than 50 studies focused on English oral language outcomes and used sound methodology. This dearth of research contrasts with the very large numbers of studies of reading instruction involving native English speakers. Nonetheless, there is some evidence that English language learners benefit from comprehensive programs featuring enhanced literature discussion and substantial experience with spoken English (e.g., Kucer & Silva, 1999). For example, engaging in a classroom discussion approach, called Instructional Conversations, where students and teachers interact with one another in a joint meaning-making process, improved fourth- and fifth-grade Spanish-speaking ELLs’ reading comprehension (Saunders & Goldenberg, 1999). The oral English support that is available for ELLs has neglected academic language and higher-order cognitive tasks – that is, language for explaining, knowledge building, reasoning, problem solving, and decision-making (Snow & Uccelli, 2009).

To address the inadequacy of conventional literacy instruction for ELLs and some of the limitations of literacy research for this ever-growing group of children, Collaborative Reasoning (CR), an alternative peer-led small-group discussion approach, was employed in the current study aiming to promote ELLs’ oral and written English development. Unlike traditional classroom discussion which emphasizes mastery of the information in texts, CR is intended to stimulate critical reading and thinking and to be personally engaging (Clark et al., 2003).

In CR, students learn to use the discourse of reasoned argumentation to discuss stories or texts that they have read. The texts contain multi-layered issues such as friendship, fairness, justice and equality, duty and obligation, honesty and integrity, winning or losing, ethnic/racial identity, and child-friendly policy issues. A Big Question central to the issues in each text is discussed in small heterogeneous groups. For example, the text entitled *A Trip to the Zoo* (Reznitskaya & Clark, 2001) tells about two girls, Lilly and Anna, discussing their upcoming field trip to the zoo. Lilly is excited at the prospect of seeing and maybe petting the animals. However, Anna does not share Lilly’s feelings and explains to her friend why she thinks that zoos are terrible places for the animals. The Big Question is: Are zoos good places for animals?

Students are expected to take public positions on the Big Question, support the positions with reasons and evidence, carefully listen, evaluate and respond to one another’s arguments, and challenge when they disagree. CR features open participation in which students speak freely without raising hands and waiting to be nominated by the teacher. Students hold interpretative authority and are responsible for their own judgments about which positions and arguments are stronger (Chinn, Anderson, & Waggoner, 2001; Reznitskaya et al., 2001). Students are responsible for managing their own discussions, negotiating turn-taking in an orderly fashion and attending to relevant issues. The teacher’s role is to facilitate and scaffold student argumentation and social skills, usually sitting outside the group.

It stands to reason that participation in Collaborative Reasoning discussions will have positive effects on ELLs’ oral and written English development. First of all, we share the widespread belief among second language educators that language is best learned in the context of extended meaningful communication (e.g., Pica, 1987). The regular use of cooperative learning groups provides students with many meaningful and structured opportunities to master the use of academic language (Calderón, Hertz-Lazavowitz, & Slavin, 1998). Cooperative learning promotes the use of a wider range of communicative functions, such as paraphrasing the ideas of others, asking for clarification, summarizing, indicating agreement or disagreement (McGroarty, 1993). CR is a highly interactive approach to discussion that may promote academic language development because children must learn to take and yield the floor, speak clearly and listen carefully, express reasons and cite evidence to justify positions, issue challenges and respond to the challenges of others.

Second, Collaborative Reasoning discussions offer extended opportunities for open discussions of complex issues, which provide young ELLs more opportunities to practice English in extended discourse. In comparison to typical forms of classroom discussion, students’ rate of talk almost doubles during CR and the talk more frequently involves cognitive processes known to be productive for learning, such as elaborating text propositions, making predictions, using evidence, and expressing and considering alternative perspectives (Chinn et al., 2001; Reznitskaya et al., 2001). Increased use of English tends to be associated with the subsequent gains in English oral language proficiency and language learning strategies (Chesterfield & Chesterfield, 1985; Saville-Troike, 1984).

Third, learning a new language is hard work and may be a source of frustration or embarrassment. Collaborative Reasoning may enhance ELLs’ motivation and engagement, which in turn may promote language development because of interest from engaging stories, intrinsic motivation from an authentic goal, sense of personal agency from working

independently, satisfaction from collaboration with peers, and excitement from socio-cognitive conflict (Wu, Anderson, Nguyen-Jahiel, & Miller, *in press*). Collaborative Reasoning discussions may also promote L2 learning attitudes because group dynamics in the cooperative learning of second language have the potential to increase L2 learning motivation, L2 self-confidence, perceived L2 competence, and to decrease language learning anxiety and stress (Dörnyei, 1997; Long & Porter, 1985).

The overall goal of this quasi-experimental study is to investigate whether Collaborative Reasoning discussions impact the development of oral and written English proficiency of young English language learners. Three specific questions are addressed: first, do Collaborative Reasoning discussions improve ELLs' English listening, speaking, reading, and writing, their motivation and engagement in discussions, and L2 learning attitudes? We anticipate that engaging in CR discussions will accelerate ELLs' oral and written English, and enhance their motivation, engagement, and English learning attitudes.

The second question is, does variation in initial English proficiency influence the benefit ELLs receive from CR discussions? In the current study, we targeted ELLs in both mainstream classes and sheltered bilingual classes. In Illinois, placement of ELLs into the two types of class is based on their performance on the annual statewide English proficiency test. Only students who pass the test can be transferred to a mainstream class where they receive instruction entirely in English. In sheltered or transitional bilingual classes, students are taught in English with Spanish support. We expect that students in mainstream classes, or ELLs with a higher level of English proficiency, will benefit from CR discussions for the reasons already presented at length. However, it is less certain that students in sheltered bilingual classes will benefit. It could be that there is a threshold of English proficiency (Cummins, 1991; Hakuta & Diaz, 1985) required for an approach such as Collaborative Reasoning to be successful. The idea of a language threshold was supported in a study showing more benefit from Instructional Conversations for high and middle achieving than low achieving Hispanic fourth graders (Saunders & Goldenberg, 2007).

The third question is, does motivation and engagement mediate the effects of CR on language outcomes? A plausible hypothesis is that CR improves children's motivation and engagement, and the heightened motivation and engagement in turn promote listening, speaking, reading, and writing.

1. Method

1.1. Participants

Four fifth-grade classrooms ($N=90$), two mainstream classrooms and two sheltered bilingual classrooms, in an elementary school in Illinois participated in the study. The school serves urban low- to middle-SES families. One mainstream and one sheltered bilingual classroom were randomly assigned to implement CR discussions while the other two classrooms served as controls. Of the total of 90 students in these classrooms, 75 (83.3%) were Latina/o and 12 (13.3%) were African Americans, and 3 (3.3%) were from other ethnic groups (1 European American, 1 Brazilian, 1 Pacific Islander). Students in the mainstream classes were predominantly Latina/o and African American. Students in the sheltered bilingual classes were all Latina/o. Because this study targeted Spanish-speaking English language learners, only the 75 Latina/o students were included in the data analysis. The mean age of the ELL sample was 11.2 years old. There were 32 Spanish-speaking ELLs in the CR condition (19 mainstream, 13 sheltered bilingual) and 43 (24 mainstream, 19 sheltered bilingual) in the control condition. This demographic information is summarized in Table 1.

According to the home language and literacy practice background survey (return rate: 70%), ELLs in the sample tended to use mostly Spanish or an equal combination of Spanish and English when speaking with their parents and other adults in the home. However, the children usually spoke to their siblings in English. Most ELLs in the sample had limited home literacy resources. For instance, more than half the families (52%) reported having fewer than 40 children's books and only 6.7% families reported having more than 60 children's books at home. Literacy-based interactions between children and parents were limited in most families; 25% of parents reported that they never read books with their children, nor did they help their children with homework in either Spanish or English. Some parents (13%) never went to the library with their children. Storytelling in the home, in either Spanish or English, was infrequently reported (24% said "never") for these fifth-grade students. Home language use and literacy practice information are provided in detail in Table 1.

Most parents had a high school education (mother: 30.7%, father: 40%). Only a few parents had received a two-year college education (mother: 13.3%, father: 5.3%). Most children in the sample were born in the U.S. and only 14 (18.6%) children were born outside of the U.S. Most families (81.4%) had lived in the U.S. for over ten years. Regarding the language of instruction in the first school year, 16 children (21.3%) were reported to have received initial reading instruction in Spanish, 19 (25.3%) to have received the initial reading instruction in both Spanish and English, and 18 (24%) to have received initial instruction in English.

1.2. Procedure and materials

Prior to the intervention, the four participating teachers attended a full-day workshop on CR. The workshop introduced the theoretical framework of CR. They learned instructional moves designed to facilitate student independent thinking and self-management of discussions and to promote argumentation and reasoning skills. These moves include prompting, thinking out loud/modeling, asking for clarification, challenging, encouraging, fostering independence, summing up and

Table 1
Demographic information and pretest scores of the Latina/o sample (N = 75).

	Mainstream class		Bilingual class	
	CR	Control	CR	Control
Number of participants	19	24	13	19
Gender				
Boy	52.6%	47.1%	38.5%	83.3%
Language child speaks to mother				
Only Spanish	42.1%	12.5%	38.5%	47.4%
Mostly Spanish	5.3%	12.5%	30.8%	5.3%
Equal amount	21.1%	12.5%	30.8%	15.8%
Mostly English	0	8.3%	0	5.3%
Only English	15.8%	20.8%	0	0
Frequencies of child reading at home				
Never	0	4.2%	0	5.3%
Once a month	5.3%	8.3%	7.7%	10.5%
Once a week	0%	12.5%	7.7%	5.3%
2–3 times a week	5.3%	16.7%	46.2%	10.5%
Everyday	73.7%	25%	23.1%	21.1%
No. of child books at home				
1–20	31.6%	16.7%	61.5%	15.8%
21–40	31.6%	25%	15.4%	21.1%
41–60	5.3%	16.7%	7.7%	5.3%
More	15.8%	4.2%	0	5.3%
Mother education level				
Elementary	10.5%	4.2%	15.4%	21.1%
Middle school	15.8%	12.5%	15.4%	5.3%
High school	36.8%	33.3%	38.5%	15.8%
College	21.1%	12.5%	15.4%	10.5%
Language used in 1st school year				
English only	42.1%	33.3%	7.7%	5.3%
Spanish only	10.5%	4.2%	61.5%	26.3%
Both	31.6%	25%	23.1%	21.1%
Child born outside the US	10.5%	8.3%	38.5%	15.8%
Pretests				
Vocabulary	.62 (.12)	.53 (.31)	.45 (.12)	.40 (.19)
Sentence grammar	.87 (.04)	.86 (.09)	.82 (.06)	.79 (.08)
Gates-MacGinitie Reading	.73 (.13)	.62 (.16)	.54 (.16)	.46 (.15)

refocusing, and debriefing (see Clark et al., 2003). The teachers watched video clips of exemplary CR discussions highlighting key elements of the approach and studied problematic episodes identifying weaknesses and generating solutions to improve the discussion. They read a CR story and prepared an argument outline to facilitate a discussion. The timeline for the study, an explanation of procedures, and a description of assessments were also presented during the workshop.

At the end of the workshop, one teacher from each pair of mainstream class and bilingual class was randomly assigned to implement CR, and the other two served as wait-list control teachers. The two teachers assigned to implement CR were provided with the CR manual and a set of eight stories with enough copies for all the students in their classes. The control teachers agreed not to use CR until after the study and they were not provided with the set of stories until study had been completed. There are several reasons for including all four participating teachers in the workshop and withholding condition assignment until the end of the workshop. Waiting until the end of the workshop allowed all teachers to begin on an equal footing. Having the same in-depth information about the project enabled the control teachers to address parental and student questions equally as well as the experimental teachers. This helped mitigate differentials in student recruitment in the two conditions. By treating control teachers as full and equal participants, we hoped to sustain their enthusiasm for the project, reduce attrition, and maintain the motivation of students to do their best on the assessments.

Before the intervention, all students completed three pretests assessing English vocabulary, syntactic knowledge, and reading comprehension. After the intervention, all students completed posttest assessments of English reading, writing, speaking, and listening, as well as surveys of motivation, engagement in discussions, and English learning attitudes.

Following the pretests, students in the two CR classrooms participated in 8 CR discussion sessions in small groups, heterogeneous in race, gender, talkativeness, and English reading level. Discussions took place over a period of four weeks with 2 sessions per week, each approximately 20 min long. The participant observer, the first author, videotaped all the discussions, observed classrooms, and took field notes on discussion days. In addition, the participant observer provided on-going support by offering teachers suggestions when needed. Students in the two control classrooms continued their regular language arts lessons.

Eight stories used in the study were chosen to be personally engaging to the children, relevant to their life experience, and appropriate for their reading level. These stories describe dilemmas faced by the characters and create contexts for serious consideration of important topics including friendship, fairness, justice, honesty and integrity, duty and obligation, winning and losing, ethnic/racial identity, and public policy. The stories ranged in reading level from second to fifth grade. The gender

of the main protagonists was balanced, four with female protagonists, four with male protagonists. Six stories centered on personal or moral dilemmas. Examples of these include *Ronald Morgan Goes to Bat* (Giff, 1990) [grade 2], *My Name is Maria Isabel* (Ada, 1993) [grade 4], and *On My Honor* (Bauer, 2004) [grade 5]. For the discussion of *Ronald Morgan Goes to Bat*, children deliberate the big question, “Should the coach let Ronald [a horrendous baseball player who has great team spirit] play on the team?” The big question for *My Name is Maria Isabel* is, “Should Maria [who has the same name as a classmate] change her name to Mary as the teacher suggested?” *On My Honor* invites children to consider the question, “Should Joel go to Starved Rock with Tony [a friend who is a reckless adventure seeker]?” Two texts, *Crystal’s Vote* (Nguyen-Jahiel, 1996) and *A Trip to the Zoo* (Reznitskaya & Clark, 2001), were written in narrative form at the third- to fourth-grade reading levels by the CR research team to introduce children to child-friendly policy issues. In the discussions of these texts, children consider whether a school should buy computers or new textbooks to replace worn out math textbooks, and whether zoos are good places for animals.

The following is a description of the assessments employed as pretests and posttests:

Vocabulary checklist. This pretest is a wide range, general vocabulary test in a checklist format (Anderson & Freebody, 1983). Students were asked to read through a long list of words and nonwords and indicate whether they know the meaning of each item. The list consisted of 180 real words with increasing difficulty, 50 nonwords (e.g., *pensile, jerbal*), and 30 pseudo-derivatives (e.g., *acceptment, inhappy*). Scores were corrected for guessing using the ‘high threshold’ formula described by Anderson and Freebody (1983).

Sentence grammaticality judgment test. This pretest is a modified version of the sentence grammaticality judgment test developed by Johnson and Newport (1989). Forty sentences, including twenty correct sentences and twenty ungrammatical sentences, were orally presented to students. Students were asked to decide if each sentence was expressed the right way in English. Examples of ungrammatical sentences are: *Bill and Joe is good friends; The little boy is speak to a policeman.*

Gates-MacGinitie reading comprehension test. This pretest is the reading comprehension subtest from the Gates-MacGinitie Reading Tests (MacGinitie, MacGinitie, Maria, & Dreyer, 2000). Students read short passages and then answered multiple-choice questions.

SVT listening and reading comprehension. These posttests employed the Sentence Verification Technique (SVT) developed by Royer and colleagues (Marchant, Royer, & Greene, 1988). Students listened to or read a passage and then judged a series of sentences. The task was to decide whether a given sentence had the same meaning as a sentence in the story they had just heard or read. In each modality (listening or reading), the test contained one below-grade level passage, one passage at grade level, and one above-grade level passage. These passages had *not* been previously read or discussed by any of the students. Each passage was accompanied by 16 test sentences, mixing original sentences, paraphrased sentences, meaning changed sentences, and distracters. This is the same SVT test shown to have satisfactory reliability and validity for a sample of fifth grade Spanish-speaking ELLs by Royer and Carlo (1991). As compared to a multiple-choice test, an SVT test has advantages in assessing ELLs’ language comprehension because it minimizes irrelevant task demands, such as test-taking strategies, and reduces the load on verbal working memory, which in principle ought to allow a more valid assessment of comprehension.

Cloze reading comprehension. Post intervention reading comprehension was also assessed with three fill-in-the-blank cloze passages, none of which had been previously read or discussed, with ten content words deleted at random from each passage. Students provided the most appropriate word to fill in the blank given the meaning of the sentence or passage. Responses were counted as correct if they made sense in context, whether or not they were exact replacements for deleted words.

Storytelling. Post-intervention English speaking ability was assessed with a storytelling task using a wordless picture book, *Frog, where are you?* (Mayer, 1969). The frog story was selected because it has been successfully used with school-age children in a wide range of countries and has been shown to be valid for assessing multifaceted language abilities of ELLs at the extended discourse level (see Berman & Slobin, 1994). The book consists of 24 pictures representing a hierarchically organized story with a main episode (a boy losing, searching for, finding his frog) and 13 sub-episodes. The interview procedures followed Berman and Slobin (1994). Students were individually pulled out from the classroom into a quiet room. During the task, students were asked by their teacher to look through all pictures in the book and then tell a story based on the pictures. Students were prompted by their teacher saying: “Can you tell me what is happening in this story?” If students stop telling the story in the middle, the teacher may ask: “Can you tell me more?” or “What happened next?” Narratives were collected in English. Children were allowed to switch between English and Spanish, although this seldom happened. Children’s narratives were audio taped using a digital voice recorder.

Reflective essay writing. The prompt for this posttest was a three-page story, *The Pinewood Derby* (McNurlen, 1998). The story had *not* previously been read or discussed by any of the students. It tells about a boy named Thomas who wins a contest building and racing model cars, but breaks the rules by not making his car by himself. He confides to his classmate, Jack, that he has received help from his older brother in making his car. Students were asked to write an essay reflecting on whether or not Jack should tell on Thomas. Students were given 40 min to write.

Attitude questionnaires. Finally, all students completed two five-point Likert scale questionnaires about their attitudes toward discussion and attitudes toward learning English. The discussion questionnaire, developed by the Collaborative Reasoning research team, evaluated students’ interest in discussions, engagement in discussions, and the benefit they

perceived from discussions. Items for CR students were phrased “Collaborative Reasoning discussions.” Items for control students were phrased “Classroom discussions,” and control students were asked to think about discussions they had had over the previous month. The questionnaire about English learning attitudes, modified from Gardner’s *Attitude/Motivation Test Battery* (1985), was used to investigate students’ attitudes, perceived competence, language-use anxiety, and risk taking in L2 learning.

1.3. Transcription and coding

Oral and written language samples generated from storytelling and reflective writing tasks were transcribed by the first author and a Spanish-English bilingual speaker, both of whom were blind to whether children had received the CR intervention, following the Systematic Analysis of Language Transcripts (SALT) conventions (Miller & Chapman, 2003). Each language sample was segmented into T-units. A T-unit is defined as one main clause and all of its subordinate clauses (Hunt, 1977). For example, “*The boy put the frog in a jar when he went to sleep*” is one T-unit containing one main clause (“*The boy put the frog in a jar*”) and one subordinate clause (“*when he went to sleep*”). The transcripts were coded for bound morphemes, omissions, pauses, errors, and mazes including repetitions, false starts, and filled pauses.

The transcripts were reviewed by a third transcriber who searched for and corrected transcription errors. The final transcripts were analyzed using SALT software and all of the standard measures were computed including transcript length (total number of T-units, total number of words), vocabulary diversity (number of different words), syntactic complexity (MLU, mean length of utterance; SI, subordination index: average number of clauses per T-unit), verbal fluency (words per minute, number of within-utterance pauses), and mazes (number of utterances with mazes). Another hand-coded score, Narrative Scoring Scheme (NSS), an index of child’s ability to produce a coherent narrative based on story grammar analysis (Stein & Glenn, 1979), was also computed as a measure of quality of storytelling. Narrative quality was scored manually by the first author. To establish the reliability of NSS scoring, 20% of the story transcripts were randomly selected and independently scored by the third author who received training to follow the scoring rubric and was blind to the intervention condition. The percent agreement for the NSS total scores between the two raters was 88%.

Following Chinn et al.’s (2001) coding scheme, written language samples were coded for argument quality using NUD*IST 6 (QSR, 2002) by the first author, as well as the same language measures employed with the storytelling task. The coding involved three stages. First, the essays were chunked into idea units. An idea unit, as defined by Mayer (1985), “expresses one action or event or state, and generally corresponds to a single verb clause” (p. 71). Next, each idea unit was assigned a position code (e.g., *Jack should tell on Thomas*) or a reason code (e.g., *cheating is not right*). In the final step, counterarguments and rebuttals were identified. A counterargument was defined as the idea unit in support of the opposite of the chosen position. A rebuttal is an idea unit that responds to the counterargument and further supports the chosen position after recognizing a possible opposing argument. Another argument stratagem, the use of text evidence (e.g., “*It is said in the story [EVIDENCE]”*, or “*The author stated [EVIDENCE]”*), was also coded. Four measures, representing the number of idea units coded as reasons, counterarguments, rebuttals, and uses of text evidence, were generated. Twenty percent of the essays were randomly selected and coded independently by the third author who is experienced in the coding scheme and was blind to intervention condition. The percent agreement between the two coders was 94%.

2. Results

Table 1 summarizes performance on the pretest measures of English proficiency. A 2 (CR vs. Control) \times 2 (mainstream vs. bilingual) MANOVA showed no significant difference between the CR and the Control groups on the pretests, $F(3, 62) = 2.03$, $p = .11$, $\eta_p^2 = .09$. Thus, the two groups can be regarded as comparable in initial English proficiency.

Table 2 displays children’s performance on SVT listening, SVT reading, and the cloze test, as well as reliability coefficients (Cronbach’s alpha) for each measure. Using the pretest scores as covariates, a 2 (CR vs. Control) \times 2 (mainstream vs. bilingual) MANCOVA showed a significant overall intervention effect on the three outcome measures, $F(3, 56) = 2.94$, $p = .04$, $\eta_p^2 = .14$, and a significant interaction between intervention condition and whether students were enrolled in the mainstream or bilingual program, $F(3, 56) = 2.98$, $p = .03$, $\eta_p^2 = .14$.

Follow up ANCOVAs found interactions between intervention condition and program on the SVT listening and cloze measures. For the students in mainstream classrooms, the CR condition was significantly better than the control condition on the SVT listening test, $F(1, 69) = 11.08$, $p = .00$, and nearly so on the cloze test, $F(1, 69) = 3.16$, $p = .08$. For the students in bilingual classrooms, however, there was no CR vs. Control difference in SVT listening, $F(1, 69) = .03$, $p = .85$, or the cloze test, $F(1, 69) = .90$, $p = .34$. The results altogether indicate that CR discussions have a greater impact on the students in mainstream classrooms than on the students in bilingual classrooms on these two measures of language comprehension. No interaction between intervention condition and program was found on the SVT reading test. The significant main effect of intervention condition on SVT reading, $F(1, 58) = 7.36$, $p = .009$, $\eta_p^2 = .11$, and lack of interaction, indicates that both mainstream and bilingual CR students performed better than control students on SVT reading comprehension.

In the following analyses, instead of classifying students according to program (mainstream vs. bilingual), initial English proficiency was treated as a continuous variable using the first principal component of the three pretests, which explained 72.5% variance in the pretest scores. To investigate if CR influences listening and reading differently for children with varying levels of initial English proficiency, three hierarchical regression analyses were conducted predicting SVT listening, SVT

Table 2

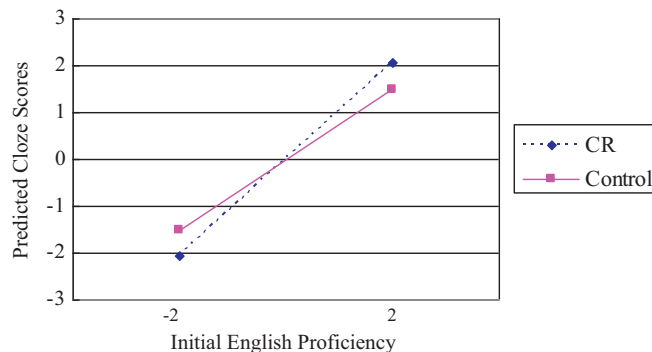
Means (SDs) of performance on SVT listening, SVT reading, and cloze tests.

Measures	α	Mainstream		Bilingual	
		CR (n = 19)	Control (n = 24)	CR (n = 13)	Control (n = 19)
SVT listening	.47	.71 (.05)	.62 (.09)	.64 (.09)	.65 (.08)
SVT reading	.72	.73 (.08)	.63 (.14)	.66 (.09)	.57 (.11)
Cloze reading	.87	.61 (.11)	.47 (.20)	.37 (.18)	.35 (.16)

Table 3

Hierarchical regression analyses of performance on SVT listening, SVT reading, and cloze reading tests.

Step		R^2 change	β	Sig.
<i>SVT listening</i>				
1	Initial English proficiency	.15	.39	.00
2	Contrast (CR vs. Control)	.03	.16	.17
3	Initial English proficiency \times contrast (CR vs. Control)	.01	.10	.08
<i>SVT reading</i>				
1	Initial English proficiency	.25	.42	.00
2	Contrast (CR vs. Control)	.08	.29	.01
3	Initial English proficiency \times contrast (CR vs. Control)	.00	.005	.96
<i>Cloze reading</i>				
1	Initial English proficiency	.70	.89	.00
2	Contrast (CR vs. Control)	.00	.01	.83
3	Initial English proficiency \times contrast (CR vs. Control)	.02	.14	.04

**Fig. 1.** Cloze reading performance as a function of initial English proficiency and intervention condition.

reading, and cloze test, respectively. In each analysis, the principal component scores of pretests were entered first, followed by the contrast, CR vs. Control, and the interaction of initial English proficiency and intervention condition. Table 3 displays the results of hierarchical regression analyses.

For the SVT listening test, a non-significant trend toward an interaction between initial English proficiency and the CR vs. Control contrast was found, $p = .08$. For the SVT reading test, consistent with previous ANCOVA, the CR vs. Control contrast was significant but no significant interaction between initial English proficiency and the CR vs. Control contrast was found. For the cloze reading test, consistent with previous ANCOVA, a significant interaction between initial English proficiency and the CR vs. Control contrast was found. Fig. 1 shows that on cloze reading comprehension children with higher initial English proficiency benefited from CR discussions, whereas children with lower initial English proficiency did not benefit from CR or even did slightly worse than controls.

Table 4 displays the statistics for the language measures from children's storytelling in six categories: story length, syntax, vocabulary diversity, verbal fluency, mazes, and narrative quality. MANCOVA documented a significant overall intervention effect, $F(9, 51) = 2.47$, $p = .02$, $\eta_p^2 = .30$ and a significant program effect, $F(9, 51) = 2.44$, $p = .02$, $\eta_p^2 = .30$. No interaction between intervention condition [CR vs. Control] and program [mainstream vs. bilingual] was found. Follow-up ANCOVA showed that the CR group produced significantly more coherent narratives in the storytelling task than did the control group, $F(1, 59) = 5.52$, $p = .02$, $\eta_p^2 = .09$. There were non-significant trends for CR students to produce more total T-units and more total words, indicating that the CR group tended to tell longer and more complex stories. No CR vs. Control difference was observed on the measures of vocabulary diversity (number of different words) and syntax (mean length of utterance, subordination index). Surprisingly, the stories produced by students from bilingual classes were significantly

Table 4
Means (SDs) of storytelling language measures.

Measures	Mainstream		Bilingual	
	CR (n = 19)	Control (n = 24)	CR (n = 13)	Control (n = 19)
<i>Story length</i>				
Number of T-units	31.52 (9.25)	31.50 (11.35)	43.00 (10.66)	33.85 (7.89)
Number of words	309.8 (91.7)	279.5 (106.5)	355.2 (104.6)	328.4 (108.5)
<i>Syntax</i>				
Mean length of utterance	8.75 (.99)	8.74 (2.06)	7.70 (.67)	8.59 (1.59)
Subordination index	1.18 (.12)	1.21 (.23)	1.14 (.07)	1.16 (.14)
<i>Vocabulary diversity</i>				
Number of different words	91.7 (19.63)	87.2 (24.3)	94.6 (14.3)	94.6 (18.0)
<i>Verbal fluency</i>				
Words per minute	121.8 (23.2)	128.3 (23.2)	94.6 (20.6)	109.2 (12.6)
Within-utterance pauses	5.52 (5.65)	2.61 (3.05)	13.23 (16.01)	6.81 (3.85)
<i>Mazes</i>				
Number of utterances with mazes	11.47 (5.87)	7.55 (4.11)	13.15 (10.18)	14.87 (6.04)
<i>Story quality rating</i>				
Narrative Scoring Scheme	22.26 (5.10)	17.94 (5.54)	21.76 (3.74)	19.55 (4.78)

Table 5
Means (SDs) of language and argument measures based on reflective essays.

Measures	Mainstream		Bilingual	
	CR (n = 19)	Control (n = 24)	CR (n = 13)	Control (n = 19)
<i>Essay length</i>				
Number of T-units	8.84 (4.20)	5.43 (3.48)	11.61 (3.73)	7.76 (5.32)
Number of words	118.8 (79.2)	69.4 (33.2)	135.5 (44.7)	79.6 (55.2)
<i>Syntax</i>				
Mean length of utterance	13.31 (3.68)	14.72 (6.52)	11.95 (2.99)	10.39 (2.60)
Subordination index	2.13 (.42)	2.16 (.67)	1.90 (.41)	1.67 (.39)
<i>Vocabulary diversity</i>				
Number of different words	63.8 (23.3)	42.2 (14.7)	59.5 (17.2)	44.1 (15.6)
<i>Argument quality</i>				
Reasons	10.58 (4.24)	5.69 (3.11)	7.92 (3.35)	5.89 (3.12)
Counterarguments	1.15 (2.56)	.31 (.60)	.76 (1.16)	.27 (.75)
Rebuttals	.84 (2.75)	.06 (.25)	.15 (.37)	.38 (1.03)
Evidence	.42 (.60)	.37 (.88)	1.46 (1.19)	.00 (.00)

longer than the stories produced by the mainstream-class students. However, bilingual-class students spoke at a slower rate (fewer words per minute), they paused more often, and their utterances contained more mazes, $ps < .05$.

Table 5 lists the descriptive statistics for the language and argument measures scored from the reflective essays. MANCOVA showed a significant overall intervention effect, $F(9, 50) = 3.42, p = .00, \eta_p^2 = .38$. A significant interaction between intervention condition and program was also found, $F(9, 50) = 2.87, p = .00, \eta_p^2 = .34$. Further ANCOVAs indicated that the reflective essays written by the CR group contained a significantly greater number of T-units, total words, and different words than the essays written by the control group, $ps < .05$. The CR group also produced significantly greater numbers of satisfactory reasons, counterarguments, and citations of text evidence than did the control group, $ps < .05$.

A significant interaction between intervention condition [CR vs. Control] and program [mainstream or bilingual] was found in the use of text evidence, $F(1, 58) = 14.22, p = .00, \eta_p^2 = .19$. For the students from bilingual classrooms, those who received CR used significantly more evidence from the story than those who did not, $F(1, 67) = 25.99, p = .00$, but no CR vs. Control difference was found for the students from mainstream classrooms. It should be noted that none of the students in the bilingual control classroom used the evidence stratagem in their reflective essays.

The essay below, written by a student from the bilingual CR classroom, illustrates the argument elements utilized by CR students. The essay is typical in the sense that the total number of idea units coded and the total number of argument elements approximated the means for the CR group. Codes for argument categories are displayed.

Yes, I think Jack should tell on Thomas because he didn't do the work by himself, and the other classmates worked very hard and didn't win. *It says in the story* [use of evidence] that Thomas didn't do the work by himself. If he didn't do the work by himself then that would be cheating because the teacher said do the work yourselves. *It also says in the story* [use of evidence] that the other classmates worked very hard and didn't win. *If* I didn't win it will be okay [counterargument], *but* it would not be okay if somebody cheated and won [rebuttal]. At last I think Jack should tell on Thomas because he didn't do the work by himself, and the other classmates worked very hard and didn't win.

Table 6
Summary of regression analyses predicting writing.

Model	Dependent variable	Predictors	R ² change	β	SE (β)	Sig.
1	Writing	CR vs. Control	.27	.52	.10	.00
2	Motivation	CR vs. Control	.29	.54	.10	.00
3	Writing	Motivation/CR vs. Control	.00	.04	.12	.74
4	Writing	CR vs. Control/motivation	.17	.50	.12	.00

The essay contains well-developed supporting reasons, counterarguments, and rebuttals. The position is clearly stated and several supporting reasons are presented. The essay cites text evidence through the use of such phrases as “It says in the story [EVIDENCE]”. The essay also uses the argument stratagem ‘Place oneself in [SCENARIO]’ (i.e., “If I didn’t win it will be okay [COUNTERARGUMENT], but it would not be okay if somebody cheated and won [REBUTTAL]”).

MANCOVA of the 10-item discussion attitude questionnaire showed that the students participating in CR discussions were significantly more motivated and engaged in discussions than the students who did not experience CR, $F(10, 51) = 3.23$, $p = .00$, $\eta_p^2 = .38$. The CR group also perceived more benefits from classroom discussions on their thinking, English speaking, listening, reading, and writing skills, $ps < .05$. Compared to the mainstream-class students, the bilingual-class students were more motivated and engaged in discussions, $F(10, 51) = 1.36$, $p = .06$, $\eta_p^2 = .24$. They also perceived more benefits from discussions, $ps < .05$.

Participation in CR discussions also significantly enhanced ELLs’ attitudes toward learning English (e.g., *Learning English is really great; I love learning English*, $ps < .05$), perceived competence in learning English (e.g., *I would rate my ability in understanding English as (low, below average, average, above average, high)*, $p < .05$), and decreased their language-use anxiety (e.g., *I am afraid the other students will laugh at me if I don’t say things right*, $p < .05$). Compared to mainstreamed students, students from the bilingual classrooms had more positive attitudes toward English learning and showed more language-use anxiety (e.g., *I get nervous and confused when I am speaking English in class*, $p < .05$).

Principal component analysis with a variance maximizing (varimax) rotation was conducted on the 10 items of the motivation questionnaire. Two principal components with eigenvalues greater than 1 were extracted. The first and second principal components, named as perceived benefits of discussion and motivation and engagement in discussion, uniquely explained 45% and 20% variance in the 10 items, respectively. Principal component scores were used in the analyses that will be described next.

We were interested whether there is a causal link between CR, motivation and engagement, and language outcomes. Motivation and engagement measures, as represented by second principal component scores, were weakly or moderately correlated with most of the language outcome measures in this study [$r = .02-.27$ depending on the measure]. Reflective essay writing performance was most sensitive to motivation and engagement ($r = .21-.27$), so if a causal link exists it is most likely for writing.

To test if CR has effect on writing mediated by its effect on motivation and engagement, a series of regression analyses were conducted. The first principal component scores of the nine writing measures represent writing. The second principal component scores of the ten motivation items represent motivation and engagement. Table 6 summarizes the regression results. In the first model, CR vs. Control contrast significantly predicted writing ($R^2 = .27$). In the second model, the CR vs. Control contrast significantly predicted motivation and engagement. In the third model, motivation and engagement did not significantly predict writing after controlling for the CR vs. Control contrast. In the fourth model, the CR vs. Control contrast still significantly predicted writing after controlling for motivation and engagement, but the predictive power dropped ($R^2 = .17$) as compared to Model 1.

To test for the mediation of motivation and engagement, its indirect effect was calculated and tested for its significance. According to Judd and Kenny’s (1981) difference-of-coefficients approach, the indirect effect between CR and writing through motivation and engagement is determined by the difference between two regression coefficients in Models 1 and 4 in Table 6 ($\beta_{\text{indirect}} = .02$). The Sobel test (1982), the most commonly used test for indirect effects, was performed to see if the difference in the two regression coefficients was significant. The results did not support the idea that enhanced motivation and engagement mediate the effect of CR on reflective essay writing, $p > .05$.

Similar analyses were conducted on the other language outcomes: SVT listening, cloze reading, and speaking (the first principal component scores of speaking measures), and the same results were obtained. Motivation and engagement was not on the path from CR to better listening, reading, and speaking. Taken together, CR discussions lead to both heightened motivation and engagement and enhanced language outcomes, but there does not appear to be a causal link between the two types of consequences.

2.1. A closer look at collaborative reasoning discussions

In this section, we present excerpts from two discussions of one group in the bilingual class and two discussions of one group in the mainstream class to complement the quantitative analyses already described. The excerpts, taken from the first discussion and the fourth discussion, were selected because we judge that they are representative of the dynamics of the entire series of discussions. The selection of the first discussion is an obvious one because it was the baseline. The fourth

discussion, rather than the last, was targeted because it was a major turning point for the bilingual students and their teacher.

The groups discussed the same stories in the same order. For the first discussion, the story was *Ronald Morgan Goes to Bat* (Giff, 1990). Ronald is completely inept at playing baseball: he cannot hit the ball, lets the ball roll between his legs, and even runs to the wrong base. He does, however, have great team spirit and is good at cheering on the team. The Big Question is: Should the coach let Ronald play on the team? For the fourth discussion, the story was *Crystal's Vote* (Nguyen-Jahiel, 1996). In this story, Crystal and Marcos are student members of a committee in their school. The committee has to decide whether they should replace the worn out fifth-grade math textbooks or buy a computer program that teaches mathematics. The Big Question is: Should Crystal vote for the textbooks or the computer program?

Teachers and students' names have been replaced with pseudonyms. We use // to denote interruptions; |n| to mark overlapping speech segments, with *n* referring to the order of the overlapping occurrence within a contiguous run; [] to note transcriber comments; () to note an uncertain segment of an utterance; and ... to indicate pauses less than 3 s within a speaking turn. Each speaking turn is numbered for ease in reference.

2.2. Discussions of bilingual group

The group we will highlight was made up of six Latino/a ELLs; two boys and four girls. The excerpt below contains the first four minutes from the beginning of the first discussion.

1	Mr. Herrera:	Now here is the question for you guys to uh... discuss. Okay? Now you don't have to agree, just go with what you believe. All right? Just because your friend says one thing, you don't have to go with his or her opinion. Okay? Just go with what you believe. All right?
2	Mr. Herrera:	Here is the question for you guys. You know that Ronald Morgan... Was he a very good baseball player?
3	Students:	No.
4	Mr. Herrera:	Should the coach let Ronald Morgan play with the team?
5	Julieta:	Yes.
6	Mr. Herrera:	Okay. Now go ahead and discuss. You said yes, why? [directs at Julieta]
7	Julieta:	I think this coach should let Ronald Morgan be in the team because he has tried a lot to... he has tried to... tried to hit the ball.
8	Mr. Herrera:	Okay? Anybody else?
9	Mariel:	And also because it is not fair, it is not fair that only like people that are old can get to the team, he can learn how to play the baseball in there too.
10	Salvino:	I think yes because I don't know how the coach could reject him, he has good spirit.
11	Orazio:	I think yes because he needs to try... he, he has... he has been trying. [silence 4 seconds]
11	Mr. Herrera:	What do you think? ...Nelli? [silence 6 seconds]
12	Mr. Herrera:	If you were the Ronald Morgan of the classroom, what would be your opinion? [silence 5 seconds]
13	Mr. Herrera:	Or if you were the one kid who really wants to win and there is really a bad kid in this classroom that you know uh, he, uh, he is probably gonna cost you the game. What do you think?
14	Nelli:	To practice more, not close his eyes when he is playing. [soft voice]
15	Mr. Herrera:	Liliana, what do you think?
16	Liliana:	Yes, so he could //
17	Mr. Herrera:	// Yes, what?
18	Liliana:	Yes, he could play in the team so he could learn more. [looks at teacher]
19	Mr. Herrera:	Okay? [silence 6 seconds]
20	Mr. Herrera:	Anybody else? [silence 3 seconds]
21	Mr. Herrera:	Let me ask you another question. Do you guys like to win?

At Turn #1, Mr. Herrera introduced Collaborative Reasoning but he did not give a complete picture of what was expected. He focused on encouraging the children to express their own opinions saying, *You don't have to agree with others. Just go with what you believe*. He then instructed the children to *go ahead and discuss*. But before the students could begin discussing, he asked a closed-end question at Turn #2, *Who was the main character of the story?*

During the 4-min run of discussion Mr. Herrera had 12 speaking turns, whereas altogether the 6 children had a total of just 9 speaking turns. Student responses were brief and the rate of student talk was quite low, in aggregate just 60 words per minute. Except for Turns #9, #10, and #11, when three children spoke consecutively without prompting, the group was reluctant to talk. Several times, Mr. Herrera felt he had to coax someone to speak. For example, he made several attempts to get Nelli to talk in Turns 11–13. Mr. Herrera held on to interpretive authority as well as control of the group's participation. The children looked toward him for confirmation after their brief speaking turns. The children did not present arguments with supporting reasons and evidence, nor did they consider others' arguments or respond to one another.

After the first discussion, the participant observer reiterated to Mr. Herrera that in CR the teacher is supposed to step back and serve as a facilitator. She reminded him that the guiding assumption of CR is that children are better able to develop reasoning and social participation skills when they manage discussions themselves. Starting with the second discussion, Mr. Herrera physically removed himself from the group and sat off to the side. Before the discussion, he gave a more comprehensive view of CR emphasizing (1) everybody participates, (2) do not raise hands, and (3) use evidence from the story. Over the next two discussions, he began to transfer discussion management and interpretative authority to the children.

The following excerpt contains 4 min from the middle of the same bilingual group's fourth discussion, *Crystal's Vote* (Nguyen-Jahiel, 1996). This discussion proved to be pivotal for Mr. Herrera and the children. By our standards, this was a full-fledged CR discussion.

- 1 Salvino: I think textbooks are better because you can take them home and study about them and you can raise money to go to field trips and not worry about the field trips and less about your study.
[silence 5 seconds]
- 2 Mariel: I think computer is better because if you make a mistake, they tell you what mistake you make and why you made and in the book they don't tell you.
- 3 Orazio: I think she should vote for computers because it says in the story as a matter of fact, I have here done sort of study which shows that students learn more from the computers than from textbooks.
- 4 Julieta: I think she should vote for the computers because it can help her search, search stuff and when she needs something she could write it down on a piece of paper.
- 5 Liliana: I think that she needs to vote for ...co...computers ...so... she could learn more.
- 6 Nelli: I think she should vote for computer because if she needs something she could search it in the computer and find about what she needs about it.
[silence 15 seconds]
- 7 Orazio: What do you think? [looked at Salvino]
- 8 Salvino: Because texts are better, you could study about them.
- 9 Orazio: What do you think? [looked at Julieta]
- 10 Julieta: Because she could... she could learn... she could learn more searching, searching on the computer.
[silence 5 seconds]
- 11 Julieta: You? Nelli?
- 12 Nelli: I think she could learn more because if she needs help on searching, she could search, uh, that the thing she needs.
- 13 Julieta: You? [looks at Liliana]
- 14 Liliana: I think yeah, because it might help her... to study more.
- 15 Orazio: Yeah, because the computers are better and they can use them again.
- 16 Julieta: You? [looks at Mariel]
- 17 Mariel: I think yeah, computers because... you can get help in there.
[silence 6 seconds]
- 18 Julieta: What made you change your decision? [looks at Salvino]
- 19 Salvino: Because you could study about them (textbooks) and computers you cannot take them home. That's why I changed my decision.
- 20 Orazio: Do you change your mind? [looks at Liliana]
- 21 Liliana: No, cause it is more better doing on computers than textbooks.
- 22 Orazio: Do you change your mind? [looks at Julieta]
- 23 Julieta: No because the, the computers... the computers help you better and learning and you can... some people can have them at homes.
- 24 Julieta: You? [looks at Nelli]
- 25 Nelli: I think the computers are better because it can still help you to search the things that you need. Because they, they... may help you the things you need for your homework and projects.
- 26 Julieta: You? [looks at Liliana]
- 27 Liliana: I think yes because it could help you... so you could study more and learn more.
- 28 Julieta: You?
- 29 Mariel: Yeah, I think computers are better.
- 30 Julieta: Why?
- 31 Mariel: Because uh... because if you get something wrong it tells you right there why you get it wrong.
- 32 Salvino: I think the textbooks are better because... when, when you... you learn about how you could look inside information about textbooks.
- 33 Orazio: Do you agree with him? [looks at Liliana]
- 34 Liliana: No, cauz ... computers help you more.
- 35 Julieta: You agree with him? [looks at Nelli]
- 36 Nelli: No. Computers are better because you can search for the things you need for the project.

Mr. Herrera talked very little in this discussion. In fact, in the 4-min excerpt above, he took no speaking turns while there were 36 student turns. As Mr. Herrera minimized his talk, the children produced longer utterances. The rate of student talk increased to 108 words per minute. The pace of the discussion, however, was still a bit slow with within-utterance pauses and occasional long silences between speaking turns.

By the fourth discussion, Mr. Herrera had stepped back and relinquished control. The children had assumed responsibility for turn management and they exercised interpretive authority. Given more social and intellectual space, two of the children, Orazio and Julieta, had emerged as leaders and assumed functions that might otherwise have been performed by the teacher (see Li et al., 2007). As shown in Turns 7, 9, 11, 13, 16, 24, 26, and 28, Orazio and Julieta nominated other children to speak and challenged them to develop better arguments. They pressed others to provide reasons for their positions with *Why?* and *What made you change your decision?* A few argument stratagems snowballed (see Chinn et al., 2001) from one of the leaders to the other. For example, when Orazio initiated the move *What do you think?* Julieta soon picked up and adopted a minimalist form of the move, directing *You?* at a particular groupmate. Likewise, Orazio picked up a stratagem initiated by Julieta. After Julieta asked, *What made you change your decision?* Orazio began to inquire of others, *Do you change your mind?* Orazio attempted to get his groupmates to consider and respond to one another's arguments. For instance, he asked Liliana, *Do you agree with him [Salvino]?*

From the fourth discussion onward, the children in the bilingual group no longer presented simple “yes” “no” positions on the Big Question. They provided extended arguments, supporting their positions with reasons, as all of them did in Turns 1–6. The children cited textual evidence, as Orazio did, stating *it says in the story as a matter of fact*, in Turn 3. When someone failed to provide supporting reasons or evidence, he/she was pressed to do so. Julieta was not satisfied with Mariel’s unsupported assertion that computers were better (Turn 29). She pressed Mariel to support her assertion. Julieta also noticed that Salvino had changed his mind at some point during the discussion without explaining why. In Turn 18, she prompted Salvino to explain his shift, *What made you change your decision?*

2.3. Discussions of mainstream group

The mainstream group was made up of five Latino/a ELLs and two African American children; three boys and four girls. An excerpt of the first four minutes of the group’s initial CR discussion, *Ronald Morgan Goes to Bat* (Giff, 1990), is presented below. Although this was the group’s first CR discussion, it already contained features typical of a CR discussion.

- 1 Mrs. Lehmann: This is our first CR discussion, and we are discussing the book Ronald Morgan Goes to Bat. Since this is your first CR discussion, I am trying to remind you a couple of things. The more discussions we have, the more natural these things will become. Make sure when you are talking to your group that you are looking at your group members. Okay? I am not part of your group. [Students: oh?!] I am sitting OUTSIDE of your group because I am not part of your group. I might throw some ideas in now and then but I am NOT part of your group. Okay? Video camera is not part of your group. [Students laughter] So don't look at the video camera. Okay? Make sure that when you are giving your thoughts or opinions, you are not talking over anybody else. You are waiting for somebody else to finish their ideas or opinions before. Uh, eye contact. Remember you don't have to raise your hand. Okay. All right?
- 2 Mrs. Lehmann: Okay. You've read Ronald Morgan Goes to Bat. And your big question is, Should the coach let Ronald Morgan play with the team?
- 3 Isabela: I think no. I think he should just let him practice and then when he is better he can get to play.
- 4 Shantell: I think he should let him play because it is not fair everybody get a chance and maybe he will get better one day.
- 5 Yasmin: I think he should get more opportunity.
- 6 Shantell: At least one or two opportunities.
- 7 Antonio: I think he should let him play too.
- 8 Charlayne: I think the coach should give him one or two chances and if he is not better by those chances, he should not get play.
- 9 Shantell: I think the coach should give him like one or two chances but if he does not get better at least a little practice a little more.
- 10 Students: True.
- 11 Jorge: I think Ronald Morgan needs to practice a lot. He needs more confidence and so and stops closing his eyes.
- 12 Shantell: I think he is going to get better. He just needs more confidence.
- 13 Yasmin: I think he should have a chance because, uh, it would not be fair, like because he will probably feel left out or something.
- 14 Shantell: The people on this team do like him because he has more spirit.
- 15 Yasmin: Yeah.
- 16 Antonio: I think he should practice more so he could get better.
- 17 Isabela: Yeah, because in the story it says he held his bats like in the wrong way and he ran in the wrong way. so I think he need to practice a little bit more.
- 18 Shantell: But he should get some chance //
- 19 Isabela: // Yeah he should get a chance but first he should practice a lot more.
- 20 Shantell: But he has a spirit I know since he has a lot of confidence he should be better.
[Several lines omitted]
- 21 Shantell: Does anyone has anything else to say?

Before launching into the Big Question, Mrs. Lehmann told the children, Turn 1, *I am NOT part of your group. I am sitting OUTSIDE of your group* (emphasis hers). She focused on a few expectations for CR discussions: (1) make eye contact, (2) listen carefully, (3) do not talk over one another, and (4) do not raise hands. Recognizing that the first CR discussion could be challenging for the children, Mrs. Lehmann told them reassuringly, *the more discussions we have, the more natural these things will become*. With this encouragement and an explicit statement of the norms, the children evidently knew how they were to go about the discussion.

Although this was their first CR discussion, the children were immediately able to take the reins, manage their own turntaking, and address the Big Question. The children conducted themselves just as one would in an adult conversation. They were attentive and respectful listeners, frequently giving one another supportive feedback like *true* and *yeah*. When the group began to recycle ideas, Shantell (Turn 21) prompted for alternative ideas: *Does anyone has anything else to say?* The rate of student talk was high, 137 words per minute. Mrs. Lehmann entered the discussion only at a few critical moments to remind the children of the participation rules and pose challenges.

Productive argumentation was present in this first discussion. The children provided good reasons to support their positions, as Shantell and Yasmin did at Turns 4 and 13. Isabela cited textual evidence, *because in the story it said*, Turn 17. There were already signs of Collaborative Reasoning and co-construction of ideas among the children. For example, six children engaged in the deliberation about Ronald needing more opportunity to practice. The children countered one another’s reasoning with *but*, as Shantell and Isabela did at Turns 18–20. One shortcoming in the first discussion was the children’s tendency to be one-sided and to cycle through the same ideas repeatedly: The coach should let Ronald have more chances, he just needed practice.

At the end of the discussion, Mrs. Lehmann held a debriefing session for the children to reflect on their participation and reasoning in the discussion. She kept her instructions short and to the point: *What is debriefing? I'm going to ask you how well*

you did in your discussions and what you need to work on. She asked the children to be specific in their evaluations of one another and to give concrete suggestions.

With such clear instructions, the children were able to independently review the strengths and weaknesses of the discussion. They provided feedback to one another and made suggestions for others to improve. The children praised each other for making personal connections to the story, talking in a respectful manner, making eye contact with each other, and using textual evidence. They noted that not all group members had an equal chance to talk. Mrs. Lehmann worked together with the children to generate ways that would allow the more quiet members to share their ideas. We expect that when children have opportunities to evaluate their own performance and set their own goals for improvement, they will make greater gains over time.

The following excerpt is taken from the middle 5 min of the mainstream group's fourth discussion, *Crystal's Vote*. The excerpt illustrates how the same group of children had honed their social participation and reasoning skills after just three discussions.

- 1 Shantell: Jorge has already changed for the textbooks.
 2 Charlayne: There are five textbook people.
 3 Shantell: Why do you think we should take the tex-textbook, Jorge?
 4 Yasmin: Yeah.
 5 Jorge: Cuz there will be more, there will, maybe there will be extras for the all, for the all fifth grade classes.
 [Several lines omitted]
 6 Mrs. Lehmann: Antonio, did you say you switch to textbook?
 7 Antonio: Yeah.
 8 Students: [laughter]
 9 Shantell: Robert, do you switch or you still want computer? You switch or you still want computers?
 10 Robert: I still want computers.
 11 Students: Why?
 12 Isabela: You can use facts.
 13 Students: Why?
 14 Yasmin: Why do you think computers are better?
 15 Robert: I just said too. I just said like //
 16 Isabela: // Why?
 [Several lines omitted]
 17 Isabela: Okay. So, Robert?
 18 Robert: Ok. and what if they write in the textbook, how are they gonna read it?
 19 Shantell: Doesn't, but doesn't //
 20 Charlayne: // (Tell them) just use a pencil?
 21 Shantell: Doesn't it, doesn't it say that things are only about math, the downloads aren't only about math, there's also about biology, geography, English. So what if they can't learn that?
 22 Antonio: Y'know, I'm still gonna say computers, never mind, (it's just on my mind) [RH] I'm sti- still gonna say computers.
 23 Shantell: Why?
 24 Yasmin: Yeah, why?
 25 Antonio: Because they are better than textbooks. Textbooks, you take it home, and then what hap-, what happ, what happens if something happens to the textbook?
 26 Shantell: But it doesn't matter.
 27 Yasmin: You will get your trust taken away |1| and you just never get it back |1|
 28 Isabela: |1| but computer, you can get virus |1|.
 29 Mrs. Lehmann: One person talks at a time. Let everybody finish their talk before you jump in.
 30 Charlayne: Besides, if you, if you, uh, use computer too much, you eyes will uh, strain because you use computer too long.
 [Several lines omitted]
 31 Robert: But you get bored when you read out of textbook.
 32 Shantell: It doesn't matter as long as you learn.
 33 Jorge: That's true, |1| but I still, I would still |1|
 34 Charlayne: |1| But it's exciting that's way they have it in the story. |1|
 35 Isabela: |1| Okay. So you'd rather be bored |1| and get an F, or learn and get an A + ?
 36 Jorge: Yeah
 37 Robert: Get a F, get bored
 38 Students: [laughter]
 39 Isabela: Oh, wow. Oh, wow. |1| Are you listening to this? |1|
 40 Antonio: |1| Computers are better |1| because when you read a textbook, you get bored and then you might fall asleep.
 41 Charlayne: You think people fall as- //
 42 Mrs. Lehmann: // That's a good point.
 43 Jorge: Yeah [laughter].
 44 Charlayne: You think, you think it might be boring, but in the story here it says it's really exciting, with colorful pictures.

Just as she did in the first discussion, Mrs. Lehmann entered into this discussion only at critical moments, to remind the group of participation rules and pose challenges when necessary. Teacher turns during this 5-min episode were low compared to student turns, 3 teacher turns vs. 45 student turns. The rate of student talk in the first discussion was already high, but it was even higher in this discussion, 198 words per minute. As indicated by several interruptions and overlapping turns and the high rate of student talk, the children were exceptionally engaged and frequently competed for the discussion floor. When too much overlapping talk occurred, Mrs. Lehmann stepped in to remind the students of the ground rules of CR.

Despite the spirited dynamic, the children were not unaware that some group members, like Robert, had not had a chance to talk. Two girls, Isabela, a Latina, and Shantell, an African American, emerged as co-leaders of the group. Together they managed turntaking and control of the topic. Shantell tried to engage Robert when he remained quiet for long stretches during the discussion. And, when Robert was not forthcoming after several students had invited him to speak, Isabela offered him more explicit help, *You can use facts*, Turn 12. The prompting of the two co-leaders and coaxing by other group members brought Robert into the discussion. He managed to counterargue against the prevailing line of argument that textbooks were the better option, *But you get bored when you read out of textbook*, at Turn 31.

As in the first discussion, the children supported their opinions with reasons and cited textual evidence, *Doesn't it say?* (Shantell at Turn 21), *that's the way they had it in the story* (Charlayne at Turn 34), and *in the story here it said* (Charlayne at Turn 44). When needed, the children prompted one another for reasons and evidence (e.g., *why? you can use facts*). The children frequently used argument stratagems to challenge and justify positions: *What if* [COUNTERARGUMENT], *It does not matter* [COUNTERARGUMENT], and, *As long as* [REBUTTAL]. These stratagems are more complex than those used in the first discussion.

Unlike the first discussion, in this discussion the children considered both sides of the issue and some children changed their minds about the Big Question. At the beginning of the discussion, the children lined up by gender: the 3 boys favored computers, the 4 girls textbooks. As the discussion progressed, the children were able to persuade a few groupmates to switch sides, and the switches did not go unnoticed by other group members. This excerpt, taken from the middle of the discussion, began with Shantell pointing out, *Jorge has already changed for the textbooks*. Another boy, Antonio, changed his initial position from supporting the purchase of computers to favoring the purchase of textbooks. He then, reconsidered and returned to his initial position, at Turn 22.

To summarize, children in the mainstream class adapted to CR more quickly and had more fluid discussions than did the children in the bilingual class. Compared to the bilingual class, the interactional dynamics were smoother. In our judgment, however, the quality of the reasoning displayed in the bilingual class does not compare unfavorably with that in the mainstream class.

3. General discussion

The major finding of this study is that engaging in language-rich Collaborative Reasoning discussions accelerates fifth grade Spanish-speaking English language learners' oral and written English, as well as their motivation, engagement in discussions, and English learning attitudes. This study extends the previous research on the effects of language-rich classroom discussions for English language learners (e.g., Saunders & Goldenberg, 1999). Despite the short duration of the intervention, only eight discussions over four weeks, positive effects were obtained not only on language comprehension (listening and reading), but also on language production (speaking and writing).

Collaborative Reasoning discussions significantly improved the SVT reading comprehension of students of all levels of initial English proficiency. In contrast, SVT listening comprehension was improved only for students with higher initial proficiency (more often in mainstream classes), but not students of lower initial proficiency (concentrated in sheltered bilingual classes). This finding seems to support Saunders and Goldenberg's (2007) recommendation that less proficient students may need more time to make large language gains in Instructional Conversations and that additional support or modifications of instruction are required to maximize the language gains of these students.

One possible reason of the differential intervention effects on listening comprehension is the difference in the CR discussions dynamics in the mainstream class and the bilingual class. As shown in the discussion excerpts, students in the mainstream class had a faster adaptation to CR and more fluid CR discussions than did the students in the bilingual class. In the first few discussions, bilingual class students seemed more reluctant, or less able, to elaborate ideas in extended language. They generated short answers to respond to the teacher instead of interacting freely with one another. They were observed rehearsing to themselves before participating overtly. Thus the flow of the first few discussions in the bilingual classroom was halting and the pace of discussion was slow. That the bilingual children were slower to adapt is not necessarily attributable to their being bilingual. Their teacher was slow to adapt, too, and his uncertain implementation of CR in the early discussions may have left the children confused about what was expected of them.

An alternative explanation for the differential intervention effect is that listening and reading make differential demands on verbal working memory. The discrepant effect of Collaborative Reasoning on SVT listening and SVT reading may be attributed, in part, to the heavier reliance on the phonological loop, an integral component of verbal working memory (Baddeley, 1986), in SVT listening than in SVT reading. When reading, problematic words can be attacked again, phrases can be reread, and the odds of complete processing are higher. In contrast, when listening, any hiatus in grabbing words out of the speech stream may result in a sentence slipping out of working memory before processing is complete. Thus, perhaps due to constrained phonological loop capacity, bilingual-class students did not show as much benefit from CR as the mainstream-class students in English listening.

This explanation is not testable in the current study since no verbal working memory tasks were included. Nevertheless, it stands to reason that the poor verbal working memory of the students with low initial English proficiency (concentrated in the bilingual classes) explains for null effects on the SVT listening test because SVT performance varies as a function of verbal working memory (Lynch, 1987). Working memory tasks have been shown to distinguish proficient bilingual children from children with limited L2 proficiency (e.g., Gutiérrez-Clellen & Weismer, 2004; Swanson, Sáez, Gerber, & Leafstedt, 2004).

A differentiated intervention effect was also found on the cloze reading comprehension test. Students in the mainstream class seemed to benefit more from CR discussions than students in the bilingual class. Without a pre-test cloze task, however, this finding should be interpreted with caution.

As compared to control students, both bilingual and mainstream students who participated in Collaborative Reasoning discussions produced oral narratives with higher quality of story structure. The narratives produced by CR students were more coherent in hierarchical thematic structuring and global plot organization. Narratives produced by CR students contained more detailed descriptions of setting, main characters, conflicts, and resolutions critical to advancing the plot in a logical order. CR students also more often expressed mental states of characters and more clearly identified referents. Children's ability to produce coherent stories probably was improved because CR leads to deeper reading of the set of stories, enabling students to obtain a better understanding of narrative structure. Unlike the ordinary reading of stories in school, where mastery of details is emphasized, our interpretation is that students who participated in CR discussions read stories more deeply in order to be better prepared for discussions. To address the big question, students needed to be sensitive to the elements of 'story grammar' (Stein & Glenn, 1979): setting, initial event, plot, conflict, and resolution to the dilemma contained in a story.

Collaborative Reasoning discussions may also have promoted ELLs' narrative skills because, in addition to deeper reading of stories, students had more opportunities to express extended ideas in English during CR discussions, sharing their understandings of stories, making personal connections, and citing the evidence from stories, which should improve students' skill at expressing narratives (Miller et al., 2006). Collaborative Reasoning discussions improved students' performance at a global narrative structure level but not at the lexical and syntactic level. No difference between CR and no-CR group was found on the number of different words and mean length of utterance in the oral narratives. This result may be due to the fact that only a narrow range of vocabulary is pertinent to the frog story and that students in this study generally produced short and simple sentences.

Replicating previous research (see Reznitskaya et al., 2008, for a review), the present study shows that engaging in Collaborative Reasoning discussions improved students' reflective essay writing. The essays written by CR students were longer, contained more diverse vocabulary, and contained a significantly greater number of relevant reasons, counterarguments, and uses of text evidence than the essays written by non-CR students. These findings suggest that participation in CR discussions allowed students to transfer the reasoning skills acquired from small-group oral discussions to an individual written task. Our theory about the mechanism of the transfer from oral argument to written argument is that students develop an argument schema, an abstract structure that represents the knowledge about the components of a complete and sound argument and relationships among the components, through socialization into argumentative discourse in CR discussions (Reznitskaya et al., 2008).

The findings of this study adds to the previous research in that English language learners appear to get a greater benefit in reflective essay writing from CR than native-English speaking students. In a summary of four CR studies involving mostly native-English speaking students in fourth and fifth grade, Reznitskaya et al. (2008) reported the effect size of the total number of argument components (reasons, counterarguments, rebuttals) pooled over the four studies, Cohen's $d = .57$. The current study found a larger effect size for the total number of argument components, Cohen's $d = .89$ between the CR and no-CR condition. This suggests that Collaborative Reasoning may have stronger effects with ELLs, although the finding will have to be confirmed in a larger, more robust study.

The current findings did not support the hypothesized link that improved motivation and engagement in turn promote language development for ELLs. Our findings are consistent with Wu et al.'s (in press) study involving 20 fourth- and fifth-grade classrooms in Illinois, which found that motivation and engagement of the children who participated in CR discussions was not correlated with the total number of argument elements in their reflective essays.

A major limitation of the present study was the small sample size and the fact that the study was a quasi-experiment enrolling just four classrooms. Therefore, there is no way to rule out, for instance, variations in the skill and enthusiasm of the teachers or positive or negative social dynamics of particular classes of students as factors influencing results. Until a larger study can be run, the results must be regarded as suggestive rather than definitive. Another limitation is the lack of a pre-assessment of oral English and, with the benefit of hindsight, the lack of measures of verbal working memory.

The current study has implications for effective literacy instruction of English Language Learners. The results offer preliminary evidence suggesting that engaging ELLs in language-rich discussions accelerates receptive and expressive language development. Despite the short duration of the intervention, participating in Collaborative Reasoning discussions positively impacted ELLs' language development, thinking and reasoning, motivation and engagement, and English learning attitudes. It should be noted, however, that the benefits of CR were more consistent for ELLs with higher initial levels of English proficiency. Assuming that the findings from this small study can be replicated, it appears that Collaborative Reasoning could help to bridge a serious gap in the education of ELLs, providing them opportunities elsewhere limited in today's schools to use English for extended meaningful communication.

Acknowledgments

The research reported in this paper was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A080347 to the University of Illinois at Urbana-Champaign. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

- Ada, A. F. (1993). *My name is Maria Isabel*. New York: Atheneum Children's Books.
- Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In Hutson, B. (Ed.), *Advances in reading/ language research* (Vol. II, pp. 231–256). Greenwich, CT: JAI Press.
- Arreaga-Mayer, C., & Perdomo-Rivera, C. (1996). Ecobehavioral analysis of instruction for at-risk language minority students. *The Elementary School Journal*, 96(3), 245–258 <http://dx.doi.org/10.1086/461826>.
- Assaf, L. (2006). One reading specialist's response to high-stakes testing pressures. *The Reading Teacher*, 60(2), 158–167 <http://dx.doi.org/10.1598/RT.60.2.6>.
- August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Baddeley, A. D. (1986). *Working memory*. New York, NY: Oxford University Press.
- Bauer, M. (2004). *On my honor*. New York, NY: Yearling.
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carlisle, J. F., Beeman, M., Davis, L. H., & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics*, 20, 459–478 <http://dx.doi.org/10.1017/S0142716499004014>.
- Calderón, M., Hertz-Lazarowitz, R., & Slavin, R. (1998). Effects of bilingual cooperative integrated reading and composition on students transitioning from Spanish to English reading. *Elementary School Journal*, 99, 153–165 <http://dx.doi.org/10.1086/461920>.
- Chesterfield, R. A., & Chesterfield, K. B. (1985). Natural order in children's use of second language learning strategies. *Applied Linguistics*, 6(1), 45–59.
- Chinn, C. A., Anderson, R. C., & Waggoner, M. A. (2001). Patterns of discourse during two kinds of literature discussion. *Reading Research Quarterly*, 36(4), 378–411 <http://dx.doi.org/10.1598/RRQ.36.4.3>.
- Clark, A.-M., Anderson, R. C., Kuo, L.-J., Kim, I.-H., Archodidou, A., & Nguyen-Jahiel, K. (2003). Collaborative reasoning: Expanding ways for children to talk and think in school. *Educational Psychology Review*, 15(2), 181–198 <http://dx.doi.org/10.1023/A:1023429215151>.
- Cummins, J. (1991). Interdependence of first-language second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children*. New York, NY: Cambridge University Press.
- Dörnyei, Z. (1997). Psychological processes in cooperative language learning: Group dynamics and motivation. *Modern Language Journal*, 81(4), 482–493 <http://dx.doi.org/10.2307/328891>.
- Ellis, R. (1986). *Understanding second language acquisition*. Oxford, UK: Oxford University Press.
- Gardner, R. C. (1985). *The attitude/motivation test battery: Technical report*. Canada: University of West Ontario, Department of Psychology.
- Giff, P. R. (1990). *Ronald Morgan goes to bat*. New York: Puffin Books.
- Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (2006). *Educating English language learners: A synthesis of research evidence*. New York: Cambridge University Press.
- Gersten, R. (1996). Literacy instruction for language-minority students: The transition years. *The Elementary School Journal*, 96(3), 227–244 <http://dx.doi.org/10.1086/461825>.
- Gutiérrez-Clellen, V. F., & Weismer, S. E. (2004). Verbal working memory in bilingual children. *Journal of Speech, Language, and Hearing Research*, 47(4), 863–876 [http://dx.doi.org/10.1044/1092-4388\(2004\)064](http://dx.doi.org/10.1044/1092-4388(2004)064).
- Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: a critical discussion and some new longitudinal data. In K. E. Nelson (Ed.), *Children's language* (pp. 319–344). Hillsdale: Lawrence Erlbaum Associates.
- Hunt, K. W. (1977). Early blooming and late blooming syntactic structures. In C. R. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, and judging* (pp. 91–104). Buffalo, NY: National Council of Teachers of English.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational stage on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99 [http://dx.doi.org/10.1016/0010-0285\(89\)90003-0](http://dx.doi.org/10.1016/0010-0285(89)90003-0).
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluation. *Education Review*, 5(5), 602–619 <http://dx.doi.org/10.1177/0193841X8100500502>.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Oxford University Press.
- Kucer, S. B., & Silva, C. (1999). The English literacy development of bilingual students within a transitional whole language curriculum. *Bilingual Research Journal*, 23(4), 347–371 <http://dx.doi.org/10.1080/15235882.1999.10162740>.
- Li, Y., Anderson, R. C., Nguyen-Jahiel, K., Dong, T., Archodidou, A., Kim, I.-H., et al. (2007). Emergent leadership in children's discussion groups. *Cognition and Instruction*, 25(1), 75–111 <http://dx.doi.org/10.1080/07370000709336703>.
- Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207–225 <http://dx.doi.org/10.2307/3586827>.
- Lynch, D. J. (1987). Reading comprehension under listening, silent and round-robin reading conditions as a function of text difficulty and working memory. Paper presented at the *Annual meeting of the American educational research association*, Washington, DC.
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Level 4, Form S. Itasca, IL: Riverside Publishing.
- Marchant, H. G., Royer, J. M., & Greene, B. A. (1988). Superior reliability and validity for a new form of the sentence verification technique for measuring comprehension. *Educational and Psychological Measurement*, 48, 827–834 <http://dx.doi.org/10.1177/0013164488483032>.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Books for Young Readers.
- Mayer, R. E. (1985). Structural analysis of science prose: Can we increase problem solving performance? In B. K. Britton & J. B. Black (Eds.), *Understanding of expository text* (pp. 65–87). Hillsdale, NJ: Erlbaum.
- McGroarty, M. (1993). Cooperative learning and second language acquisition. In D. D. Holt (Ed.), *Cooperative Learning: A response to linguistic and cultural diversity* (pp. 19–46). McHenry, IL: Center for Applied Linguistics and Delta System, Inc.
- McNurlen, B. (1998). *Pine Wood Derby*. Champaign, IL: Center for the Study of Reading.
- Miller, J., & Chapman, R. (2003). *Systematic analysis of language transcripts [computer software]*. Madison, WI: Language Analysis Laboratory, Waisman Center, University of Wisconsin.
- Miller, J., Heilmann, J., & Nockerts, A. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research and Practice*, 21(1), 30–43.
- National Center for the Educational Statistics. (2009). *National Assessment of Educational Progress, 2009, reading assessments*. Washington, DC: US Department of Education Institute of Education Sciences.
- Nguyen-Jahiel, K. (1996). *Crystal's vote*. Champaign, IL: Center for the Study of Reading.
- Pica, T. (1987). Second-language acquisition, social interaction, and classroom. *Applied Linguistics*, 8(1), 3–21 <http://dx.doi.org/10.1093/applin/8.1.3>.
- Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, 97(2), 246–256 <http://dx.doi.org/10.1037/0022-0663.97.2.246>.
- QSR. (2002). *QSR NUD*IST 6 (computer software)*. Victoria, Australia: Qualitative Solutions and Research.
- Reznitskaya, A., Anderson, R. C., Dong, T., Li, Y., Kim, I., & Kim, S. (2008). Learning to think well: Applications of argument schema theory. In C. Block & S. Parris (Eds.), *Comprehension instruction: Research-based best practices* (2nd ed., pp. 196–213). New York: Guilford Publishing.
- Reznitskaya, A., Anderson, R., McNurlen, B., Nguyen-Jahiel, K., Archidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes*, 32(2–3), 155–175 <http://dx.doi.org/10.1080/0163853X.2001.9651596>.
- Reznitskaya, A., & Clark, A. (2001). *A trip to the zoo*. Champaign, IL: Center for the Study of Reading.
- Royer, J. M., & Carlo, M. S. (1991). Assessing the language acquisition progress of limited English proficient students: Problems and a new alternative. *Applied Measurement in Education*, 4(2), 85–113 http://dx.doi.org/10.1207/s15324818ame0402_1.
- Saunders, W. M., & Goldenberg, C. (1999). Effects of instructional conversations and literature logs on limited- and fluent-English-proficient students' story comprehension and thematic understanding. *The Elementary School Journal*, 99(4), 277–301 <http://dx.doi.org/10.1086/461927>.

- Saville-Troike, M. (1984). What really matters for second language learning for academic achievement? *TESOL Quarterly*, 18(2), 199–219 <http://dx.doi.org/10.2307/3586690>.
- Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P., & Hodge, J. P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *The Elementary School Journal*, 95(5), 387–408 <http://dx.doi.org/10.1086/461851>.
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *Cambridge Handbook of Literacy* (pp. 112–133). New York: Cambridge University Press.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). Washington, DC: American Sociological Association.
- Stein, N., & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing* (pp. 53–120). Norwood, NJ: Ablex.
- Saunders, W., & Goldenberg, C. (2007). The effects of an instructional conversation on English Language learners' concepts of friendship and story comprehension. In R. Horowitz (Ed.), *Talking texts: How speech and writing interact in school learning* (pp. 221–252). Mahwah, NJ: Erlbaum.
- Swanson, H. L., Sáez, L., Gerber, M., & Leafstedt, J. (2004). Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *Journal of Educational Psychology*, 96(1), 3–18 <http://dx.doi.org/10.1037/0022-0663.96.1.3>.
- Wu, X., Anderson, R. C., Nguyen-Jahiel, K., & Miller, B. (in press). Enhancing motivation and engagement through collaborative discussion. *Journal of Educational Psychology*.