Learning and Instruction 71 (2021) 101388

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Exploring the relationship between dialogic teacher talk and students' persuasive writing

Shireen Al-Adeimi [a,*], Catherine O'Connor [b]

[a] Michigan State University, College of Education, 620 Farm Lane Erickson Hall 327, East Lansing, MI, 48824, USA
[b] Boston University Wheelock College of Education & Human Development, 2 Silber Way SED 330, Boston, MA, 02215, USA

ARTICLE INFO

Keywords:
Dialogic discourse
Persuasive writing
Teacher talk
Writing
Discussion

ABSTRACT

Studies have documented the importance of dialogic classroom discourse for supporting academic outcomes such as reading comprehension and vocabulary development. This study examines the relationship between teacher talk during whole-classroom discussions in 42 classrooms and post-discussion persuasive essays from students in grades four through seven (n = 471). Teacher talk was coded by type of question (contestable, semi-open, or quiz-like) and teachers' follow-ups (prompting, pressing for reasoning, active listening), and further categorized as indicating either high or low levels of dialogic talk. Multilevel modeling that accounts for student participation rate, discussion topic, and students' demographic information shows that high dialogic teacher talk positively predicts students' persuasive essay scores, while low dialogic teacher talk negatively predicts their scores. The study supports existing hypotheses about the role of teacher dialogic talk in whole-class settings.

1. Introduction and literature review

Classrooms are dynamic environments in which a multitude of interactions and activities co-occur. The complexity at the heart of classrooms calls for the investigation of multiple factors that relate to various student outcomes. At the same time, this complexity poses a challenge to those who wish to accurately isolate and capture aspects of classroom interactions. This study focuses on one aspect of classrooms — teacher talk during whole-classroom discussion — and investigates the ways in which various features of teacher talk relate to students' post-discussion persuasive writing.

Writing is an essential, yet difficult component of literacy. It is a cognitive, social, and linguistic task that requires writers to manage several demands at once, as writers seek to communicate their thoughts to a non-present audience (Gee, 2007). Students' academic success is in part dependent upon their mastery of academic writing, and in upper-elementary and middle school, students encounter a new and more challenging form: persuasive writing.

To persuade an audience, writers take on a position, then present evidence and reasoning to support their position, while also considering counterarguments to their claims. This process carries more cognitive and linguistic demands than narrative writing (Nippold, Ward-Lonergan, & Fanning, 2005; Salahu-Din et al., 2008). Persuasive writing is emphasized across subject areas in standards adopted by most U.S. states (the Common Core State Standards). Students are expected to gain proficiency in this genre by the end of high school (National Governors Association, 2010) in preparation for university entrance exams which often test students' written skills through persuasive writing tasks (College Board, 2012). Beyond classroom settings, persuasive writing also prepares students for the work environment and for tasks encountered in society at large (Nippold, Ward-Lonergan, & Fanning, 2005).

Writing can pose a challenge for students as it is a productive, rather than a receptive task (Kress, 1994). Additionally, many students struggle with proficiency in the persuasive writing genre (National Commission on Writing, 2006), perhaps because the demands of writing may not be taught explicitly. This complex task is also oftentimes difficult to instruct, partly due to the lack of comprehensive research on writing instruction as well as the potential disconnect between research and classroom practice (Myhill & Fisher, 2010). For these reasons, much is yet to be understood about classroom factors that can promote persuasive writing.

1.1. Dialogic classroom talk

One factor that logically should support persuasive writing is classroom talk — a ubiquitous feature of any classroom. Over the years,

* Corresponding author.
E-mail addresses: aladeimi@msu.edu (S. Al-Adeimi), mco@bu.edu (C. O'Connor).

https://doi.org/10.1016/j.learninstruc.2020.101388
Received 31 December 2019; Received in revised form 22 July 2020; Accepted 23 July 2020
Available online 12 August 2020

pedagogical approaches that prioritize dialogue-based talk have emerged; these include accountable talk (Michaels et al., 2008), collaborative reasoning (Waggoner et al., 1995), Quality Talk (Murphy et al., 2017), dialogic teaching (Alexander, 2005), dialogic discussions (Tharp & Gallimore, 1988), exploratory talk (Mercer & Littleton, 2007), instructional conversations (Cazden, 1988), and others. While specific definitions vary (see Alexander, 2020 for an extended discussion), to various degrees, these approaches are illustrative of the Bakhtinian notion of dialogic discourse (Bakhtin, 1984), Vygotsky's social constructivist theory (Vygotsky, 1978), and Dewey's call for inquiry-oriented learning (Dewey, 1938, p. 18). They also stand in contrast to the Initiation-Response-Evaluation approaches that remain a common feature in classrooms despite criticism for their inability to involve students in learning that promotes their thinking (Alexander, 2008; Cazden, 1988; Mehan, 1998; Nystrand, 1997).

We use the term *dialogic talk* to encompass a range of dialogicity in teachers' interactions with students during whole-classroom discussions. In contrast to monologic talk, dialogic talk provides students with opportunities to internalize and appropriate linguistically conveyed reasoning and argumentation skills used during discussion (Vygotsky, 1978).

1.2. Operationalizing classroom talk

Several theory-driven instruments have been created in order to operationalize the varieties of classroom talk at the heart of K-12 classrooms. These include subject-area specific measures such as the Protocol for Language Arts Teaching Observations (PLATO; Grossman et al., 2009) and the Science Discourse Instrument (SDI; Osborne et al., 2019), as well as cross-subject tools such as the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008), the Dialogic Inquiry Tool (DIT; Reznitskaya et al., 2012), and Quality Talk (QT; Murphy et al., 2017). Most of these require "high inference" judgements about the degree to which talk in a classroom is "dialogic" while some (e.g. SDI) also judge the quality of the dialogic interaction.

In examining the types of discourse moves coded in existing instruments, a dichotomy between *open-ended* and *closed* questions emerges. Open-ended questions (or authentic, contestable, questions) are questions that do not have a right answer. The point of discussion of these questions is to get students to externalize their reasoning and work with the reasoning of other students. In contrast, closed or test questions have a known answer and are therefore not conducive to dialogic interactions. Additionally, teacher moves such as those that *press for reasoning* (e.g. Michaels & O'Connor, 2015) or *link* students' responses to one another (e.g. Osborne et al., 2019) emerged.

1.3. Supporting persuasive writing through talk

Empirical studies have shown that whole classroom dialogic talk is related to content comprehension (Ketch, 2005; Murphy et al., 2009; O'Connor et al., 2015) a cornerstone for success in school and beyond. In this study, we explore whether teacher-initiated dialogic talk supports students' persuasive writing.[1] As dialogic classroom talk involves synchronous exchanges about thinking and reasoning among various interlocutors, we hypothesize that it may support persuasive writing, which itself is a form of (asynchronous) communication about reasoning and argumentation between writers and their audiences. Thus, we hypothesize that learning that is negotiated and constructed during a communicative activity such as classroom discussion may have bearing on how such learning is demonstrated in another communicative activity — writing to persuade.

[1] While we only focus on teacher talk here, this paper is part of a larger project that also explores the role of student dialogic talk in supporting persuasive writing.

Few studies have empirically explored whether dialogic approaches contribute to students' argumentation skills in writing; however, some evidence is accumulating. For example, Kuhn et al. (2016) have shown that *dialogic communication* (small group discussions, followed by electronic dialogues between student pairs, and culminating in a whole class discussion) in middle-school classrooms improved students' writing quality. Persuasive essays collected from 38 students over two years were found to contain significantly more arguments from both sides of the issue, an evaluation of the opposing position, and the use of *however* to connect opposing views, when compared to pre-intervention essays. The authors contend that more evidence is needed before they can "prove that the dialogic aspect of [the] intervention contributed to students' achievement" (p.115) but they draw their conclusions based on the transfer of students' oral statements (e.g. "Others might say …") to their writing.

The same dialogic approach utilized in the previous study was also found to contribute positively to adolescent students' spoken dialogic argumentation (Crowell & Kuhn, 2014) and students' metalevel comprehension of argumentative discourse (Kuhn et al., 2013). Additionally, Kuhn and Crowell (2011) found that over time, argumentative reasoning skills were enhanced for students who took part in dialogic communication prior to writing on the same topic.

It is worth noting here a recent large-scale study examining the relationship between classroom dialogue and various student outcomes in 72 classrooms (Howe et al., 2019). Specifically, Howe et al. (2019) analyzed the predictive association between 11 dialogue variables and six student outcomes including the Year 6 Standardized Assessment Tests that assess mathematics, grammar, spelling, and punctuation, as well as reading. They found that given high student participation, elaborated talk (whereby students build on or elaborate upon their own or another person's contribution) and querying (whereby a speaker doubts or challenges a statement) were positively associated with students' scores (Howe et al., 2019). Though their study does not focus on persuasive writing as an outcome, it highlights an important shift toward linking specific discussion features to student outcomes, an approach also taken by our study.

1.4. Present study

Building upon prior works that explore specific features of classroom talk (e.g. Murphy et al., 2009; O'Connor et al., 2015), and those linking talk content to specific student outcomes (e.g. Howe et al., 2019), this study aims to explore this relationship more explicitly, as specific features of dialogic talk will be examined. Furthermore, whereas the aforementioned studies employ a unique discussion structure that may be difficult to implement in regular classrooms, this study will concentrate on whole class, teacher-facilitated discussions only, a structure that is familiar to students and teachers in K-12 classrooms. Also, unlike studies such as Howe et al. (2019), where discussion features are examined against end-of-year assessments, this study examines a student outcome directly related to their topic of discussion. That is, this study allows for testing the hypothesis that if students engage in discussions about a given topic prior to producing persuasive essays where they take a position on the topic of discussion, perhaps they will be able to create argument schemas that will then transfer from verbal argumentation to persuasive writing (Reznitskaya et al., 2001).

2. Method

We empirically investigate the predictive association between various features of whole-classroom teacher dialogic talk and students' persuasive writing skills. Examining how classroom-level data predict students' individual academic outcomes in a writing assignment, this paper analyzes the contents of teacher utterances during whole class discussion. Specifically, the following research question will be examined: *When accounting for students' participation levels, prior academic*

Table 1
Topics by grade level (471 essays produced in 42 classrooms discussing 25 topics).

Grade	Topic	Essays	Classes	Prompt
4	Purchasing	45	3	Why do we buy what we buy?
	Language	36	2	Should everyone learn a second language?
	Origins	35	2	Who cares where we're from?
	Clothing	31	4	Why do we wear what we wear?
	Nutrition	24	2	Who should decide what we eat?
	Uniforms	22	1	Should students be required to wear uniforms?
	Fairness	20	1	What is fair?
	Behavior	9	1	Should students share responsibility for each other's behavior in school?
5	Fighting	36	3	Why do we fight?
	Privacy	35	2	Do we need to give up our privacy to protect our communities?
	Inclusion	33	5	Should everyone be included?
	Impact	31	2	How can the words and ideas of one person impact a community?
	Caring	12	2	Why should I care?
	Divisions	10	1	What divides us and how can we resolve our differences?
	Happiness	10	1	Why do communities have different ideas about what brings happiness?
	Power	6	1	The power of Power
	Laws	5	1	Do we need laws to regulate our behavior?
6	Plastic	9	1	Should our use of paper or plastic be an individual choice or be regulated by the government?
	Steroids	6	1	Steroids: Substance abuse or an innocent boost for athletes?
	Ideas	4	1	How do we test that idea?
7	School	17	1	Is an extended school day the right choice for U.S. students?
	Suicide	13	1	Should doctors be allowed to assist seriously ill patients with suicide?
	Units	12	1	So what exactly is a unit?
	Leaders	6	1	The Pharaohs of ancient Egypt: Protectors of order or sources of oppression?
	Spenders	4	1	Pyramids and other monumental structures: Great achievements or a waste of Egypt's surplus?

language skills, ethnicity, sex, socio-economic status, special education status, and English learner status, are students' persuasive essay scores predicted by teacher questions and follow-ups that we judge to be indicative of highly dialogic talk?

2.1. Research setting

All data analyzed in this study were gathered as part of the Catalyzing Comprehension through Discussion and Debate (CCDD) study, a large-scale intervention involving urban elementary and middle schools (grades 4–7 in the U.S. system, or ages 9–12) in the Northeastern United States. Data analyzed in this paper are from whole-classroom discussions and persuasive essays from classrooms that implemented the Word Generation program (WG), a supplementary, cross-disciplinary curriculum developed in partnership with the Strategic Education Research Partnership.

Word Generation aims to improve students' academic language and vocabulary through discussion (Snow et al., 2009). The program consists of units that introduce elementary and middle school students to social or civic dilemmas through short activities, along with academic vocabulary integrated into these activities. Examples of these dilemmas include, "Should students be required to wear uniforms?" and "Do we need to give up our privacy to protect our communities?" Over one or two weeks, students read materials related to a given topic, developed positions on the topic, identified reasons and evidence for their positions, and engaged in debates about the particular issue. These whole-classroom debates lasted approximately 45 min and were audio-recorded.

Table 2
Student demographic data.($N_{total} = 369$)

	N	Percentage (%)
Eligible for free or reduced lunch	319	86.50
Race		
Black	154	41.73
White	106	28.73
Latinx	89	24.12
Other	20	5.42
Gender		
Female	200	54.20
Male	169	45.80
English language learners (ELL)	18	4.88
Students with Individualized Education Plans	37	10.03

At the end of a unit, students were asked to produce a persuasive essay in response to a prompt asking them to take a stance on the given issue. Like argumentative writing, these "Take a Stand" essays encourage students to ground their writing in text-based evidence and reasoning, and are also purposefully aligned to standards that require students to "write arguments to support claims with clear reasons and relevant evidence" (National Governors Association, 2010; Snow et al., 2009).

All topics examined in this study are presented in Table 1 below. They include whether everyone should learn a second language (grade four), why people fight (grade five), whether laws should regulate our use of paper or plastic (grade six), and whether physician-assisted suicide should be allowed (grade seven). It is important to note that fidelity of implementation varied across schools. That is, while all students in the sample took part in WG activities and discussions, not all students completed their essays. Thus, the total number of essays produced per classroom ranged from minimal (4 essays) to full participation (22 essays). Out of the 25 topics in the sample, the topic "Why do we buy what we buy?" produced 45 essays (from three classrooms), while two others, "Pyramids and other monumental structures: Great achievements or a waste of Egypt's surplus?" and "How do we test that idea?," produced the fewest essays (with four essays each) (Table 1).

2.2. Participants and data

A total of 42 audiotapes of classroom discussions from two districts were analyzed, along with their corresponding essays (n = 471).[2] As the most dominant ethnic group was Black/African American (41.7%), all other groups are discussed in reference to Black students. A high percentage of students in this sample (86.5%) were eligible for free or reduced lunch. Other student demographic data are presented in Table 2.

2.3. Measures

Outcome variable. Though students ranged in ages from 9 to 12, they all studied the same curriculum that centered around contestable topics, debates, and "Take a Stand" essays. Students were not explicitly instructed on their writing, and all essays produced were first drafts. As such, we scored all students' persuasive essays using a holistic rubric that accounted for the presence of claims, reasoning, counterclaims, and conclusions. We did not, however, assess the quality of students' argumentative structure using this rubric. A score of 4 represents essays that included a claim(s), evidence or reasoning, acknowledgement of

[2] There were 369 unique students, 25 different topics, and 24 different teachers in this sample.

Table 3
Low-inference teacher utterance codes with examples.

Category	Code Type	Code Description	Examples
High Dialogic Talk (HighDT)	Follow-ups: prompts	Prompting student-to-student interactions about content	*What do you think about what she said?* *Who wants to agree, disagree, or add on to that idea?*
	Follow-ups: press for reasoning	Following up with a particular student by asking for clarification, examples, or evidence	*Why do you think that?* *Can you give us an example of that?*
	Follow-ups: active listening	Using active listening to keep a student talking	*Uh huh, ok.* *That's interesting.*
	Questions: contestable	Asking an open-ended, contestable question	*Should people be able to rent pets?* *Were the Pharaohs oppressive rulers or great leaders?*
Low Dialogic Talk (LowDT)	Questions: semi-open	Asking a semi-open, "how" question	*How do you find all the common factors?* *What is the process for becoming a citizen?*
	Question: quiz-like	Asking a quiz-like question with a known response	*Can someone define the word 'operable'?* *What's the opposite of negative six?*

counterclaims, and a stated conclusion. Incomplete essays or those that were written about an irrelevant topic were excluded from analysis. Essays with a score of 1 often included an unsupported claim or a series of claims, while essays with a score of 2 often included a claim and supporting reason(s)/evidence. A score of 3 indicated presence of three out of the four categories (e.g. claim with evidence and a conclusion.)

To achieve reliability, two raters independently scored a number of essays based on the scoring rubric above. Following initial scoring, raters discussed areas of disagreement before independently scoring 20% of essays in the sample (n = 95). The resulting Cohen's kappa score of 0.96 indicates a high inter-rater reliability between both raters.

2.4. Coding teacher talk

Given that data in this study were in the form of audio recordings of discussions across a number of grades and subject areas, we were unable to use instruments that require the use of video (e.g. CLASS), those that make high-inference judgements about the quality of talk (e.g. SDI), subject-specific instruments (e.g. PLATO), or those that are used to code small-group interactions (e.g. QT).

We therefore identified "low inference" features of teacher talk during whole-class discussions that capture a range of dialogic talk. Specifically, we focused on six types of teacher utterances categorized into **questions** and **follow-ups**, and hypothesized as indicating either high or low dialogic talk (Table 3). Additionally, as the focus of our study was dialogicity (rather than quality) of teacher-initiated talk, we coded the presence, rather than quality, of these features.

A team of research assistants was trained to code all six teacher moves. Following the training, coders were assigned practice files that they first coded independently and then compared and discussed their codes in pairs. After coders were judged to be ready to code independently, we used 25 double-coded files to calculate interrater reliability as percent agreement among coders in relation to the mean score for each talk category. This approach yielded the following agreement percentages: follow-up prompts (88%), press for reasoning (80%), active listening (68%), contestable questions (80%), semi-open questions (76%), and quiz-like questions (80%).

2.5. Predictors: teacher talk features

Coding at an utterance level, we focused on content-related teacher utterances during whole-class discussion that were either **questions** or **follow-ups** to student responses. We first coded teacher **questions** that constituted the focus of discussion in the surrounding utterances and coded three types of teacher-initiated questions. One was truly contestable, open-ended questions, which were typically a central part of the WG curriculum (See Table 3 for examples of each category.). This category of utterance was taken as an indicator of "high levels of dialogic talk."

In contrast, we also coded for questions that would be indicators of "low levels" of dialogic talk: (a) quiz-like questions with a single right answer, and (b) semi-open "how" questions—questions intended to get students to report on their approach to a problem, e.g.: "how do you find all the common factors?"

In addition to these three teacher question codes, we coded teachers' **follow-ups** to students' responses. First, we coded utterances in which teachers invited students to make connections to each other's ideas and contributions. Specifically, we looked for instances of teachers prompting student-to-student interactions about content, e.g.: "Giselle, what do you think about what Dani just said?" or "Who has a response to what Ali just proposed?" These, we reasoned, would support a key feature of dialogic discourse: participants engaging with one another's thinking, with a focus on understanding and sharing thoughts rather than competing for a "right answer."

Next, we focused on teacher follow-ups with a particular student. We tracked moves where teachers sought to dig into a particular student's reasoning by asking questions about evidence, reasoning, examples, etc. We also tracked teachers' follow-up moves intended to get a student to expand or keep talking without pressing for reasoning. These "active listening" utterances also have the potential to contribute to the dialogic nature of the discussion.

The first four teacher discussion features in Table 3 (all three follow-ups and contestable questions) are hypothesized to be more indicative of dialogic talk as they are intended to prompt students to think more critically, to bring their own reasoning to the discussion, and to engage with others on a given topic. On the other hand, semi-open and quiz-like questions are hypothesized to be less indicative of dialogic discourse, as they are less likely to engage students in extended discussion focused on understanding, sharing, and contesting one another's reasoning. While these six features do not represent the full range of teacher utterances in a given discussion, we chose them as the most indicative of the potential for dialogicity or lack thereof.

Presence of discussion features was coded minute-by-minute by trained coders. After coding, teacher **questions** and **follow-ups** were categorized by whether or not they were indicative of dialogic talk. To account for variability in discussion time, features in each category were summed and divided by discussion time. The resulting two prediction variables were low dialogic talk (LowDT) and high dialogic talk (HighDT). To illustrate, if there were four teacher questions indicative of high levels of dialogic questions (i.e. contestable questions), one instance of a teacher prompt that invited discussion among students, three instances of teacher follow-ups that pressed for reasoning and two instances of active listening follow-ups during a 45 min discussion, we would have 0.22 instances of high dialogic talk (HighDT) per minute of classroom discussion. On the other hand, if there were eight questions indicative of low levels of dialogic talk during 45 min of discussion (semi-open and/or test-like questions), the resulting LowDT score would be 0.17 instances of low dialogic talk per minute of discussion time. Importantly, the coding scheme described here accounts for the complexity of teacher-facilitated classroom discussions in that it captures a range of discourse moves, some of which are classified as

indicative of low dialogic discussion, while others of high dialogic discussion. Thus, the presence of HighDT in a given discussion does not imply the absence of LowDT and vice versa.

Participation rate (control). An additional component of this study, student participation rate, was included in anticipation of two scenarios. First, topics that students find more interesting will likely elicit more participation. This may in turn better prepare students for persuasive writing through supporting their claims with evidence or reasoning and including counterarguments they may have encountered during the discussion. Second, participation rate itself may influence the nature of teacher-initiated questions and follow-ups; for example, if students respond to a teacher's contestable question with a yes/no response, the teacher may ask semi- or quiz-like questions instead to encourage broader participation. Thus, to account for the effect of participation on the outcome variable, we included participation rate as a control variable in the models described below.

Participation rate was captured through a minimal measure: the percentage of students in the class who made at least one contribution during the discussion. Research assistants were present during data collection of classroom discussions and marked each time a student made a contribution. Noting the number of different students who participated out of the total number of students present yielded a percentage that we subsequently classified as the following: Low levels of participation (0–24%) were assigned a score of 1; participation rates of 25–49% were assigned a score of 2, participation rates of 50–74% were given a score of 3; and finally, participation rates of 75% or higher were assigned a score of 4.

Academic language (control). We included students' academic language scores to understand whether teacher talk predicts students' persuasive writing regardless of their prior academic language skills. Students were given a validated measure of overall academic language skills, the Core Academic Language Skills Instrument (CALS-I) (Uccelli et al., 2015), prior to discussions. The language skills captured by the CALS-I are defined as "knowledge and deployment of a repertoire of language forms and functions that co-occur with oral and written school learning tasks across disciplines" (Uccelli et al., 2015, p. 1079). Specifically, the instrument assesses students on six skills: *unpacking complex words* (e.g. morphological skills), *comprehending complex sentences* (understanding complex syntax), *connecting ideas* (knowing how to use connectives and discourse-markers), *tracking themes* (anaphoric resolution[3]), *organizing argumentative texts*, and *awareness of academic register,* which was measured in two ways[4] (Uccelli et al., 2015).

Demographic variables (controls). Other student-level variables are students' socioeconomic status (SES) as indicated by receipt of free or reduced-price lunch, students' sex (Female), ethnicity (Black, Latinx,[5] White, and Other, with Black as a reference category), English language learner status (ELL), and special education status (students with an individualized education plan).

Fitted model. Given that certain topics are likely to elicit more discussions than others, which could then impact initial essay score (intercept), all models fitted in this study have a random intercept by topic. Intercepts were also allowed to vary randomly by teacher so as to account for any possible teacher effects.[6] Additionally, we fitted a random slope for participation rate by topic in order to account for the possibility that different topics might elicit different levels of participation, which may then also influence the relationship between participation and essay score (slope). The following 2-level multilevel (final) model was fitted using the statistical software R (R Core Team, 2016):

Level-1 model (Individual-level):

$$\widehat{EScore}_{ij} = \beta_{0j} + \beta_1 AL_{ij} + \beta_2 SES_{ij} + \beta_3 ELL_{ij} + \beta_4 SPED_{ij} + \beta_5 White_{ij} + \beta_6 Latinx_{ij} + \beta_7 Oth_{Raceij} + \beta_8 Female_{ij} + \varepsilon_{ij}\varepsilon_{ij} \sim N(0, \sigma_y^2)$$

Level-2 model (Group-level):

$$\beta_{0j} = \gamma_{00} + \gamma_{01} HighDT_j + \gamma_{02} LowDT_j + \gamma_{03} PAR_j + u_{0j}$$

$$u_{0j} \sim N(0, \sigma_\alpha^2)$$

where:

- *EScore* is the predicted essay score of student *i* in classroom *jAL*$_{ij}$*HighTD*$_j$.
- β_{0j} denotes a random intercept for each classroom and is further described by the level-2 model.[7]
- $\beta_1 - \beta_8$ are individual-level controls (369 observations)
- γ_{00} is the overall intercept across all groups
- $\gamma_{01} - \gamma_{04}$ are the classroom discussion predictors (42 observations):
  - *HighDT* is High Dialogic Talk;
  - *LowDT* is Low Dialogic Talk
  - *PAR* is overall student participation rate
- Errors are assumed to be normally distributed at both the student and classroom levels and are denoted by ε_{ij} and u_{0j} respectively

Table 4
Summary statistics for student-level and classroom-level variables.

	Mean	St. Dev.	Min	Max
Student-level variables (N = 471)				
Essay Scores (Outcome)	2.40	0.84	1.00	4.00
Academic Language (AL)	−0.04	1.00	−2.40	4.30
Classroom-level variables (N = 42)				
Discussion Features (per minute)				
High Dialogic Talk (HighDT)	0.45	0.33	0.00	1.20
Low Dialogic Talk (LowDT)	1.00	0.66	0.08	2.70
Other classroom variables	3.10	0.91	1.00	4.00
Participation Rate (PAR)				
Discussion Time (in minutes)	34	11	14	76

3. Results

Summary statistics for student- and classroom-level variables are shown in Table 4. Over 75% of students received a score of 2 (n = 241) or 3 (n = 122) out of 4 on their essays, while the rest either scored a perfect score (n = 65) or the lowest score (n = 43). The mean essay score was 2.4 out of 4 (*SD* = 0.84), or 60%.

3.1. Correlations

Next, correlations among variables of interest were explored (Fig. 1). While these correlations do not reflect the multilevel aspect of the relationship among predictors and outcomes, they serve as early predictors of hypothesized relationships in addition to ruling out potential collinearity among variables of interest.

Correlations between outcome and main predictors. As shown in

[3] Anaphors refer to words or phrases that refer to a previously stated idea or person (e.g. 'she' refers to 'Sarra'). Anaphors in academic texts can be conceptual (e.g. 'this' refers to "The discovery of insulin.').

[4] Students were asked to identify the academic register (i.e. academic vs. colloquial definition) and were also asked to produce the academic register (i.e. they were asked to write definitions of four words for a dictionary for adults.).

[5] This is a gender-neutral term used to refer to people of Latin American origin.

[6] There were 42 classrooms in this sample and 24 unique teachers taught those lessons. Of the 471 student essays, 369 were produced by unique students.

[7] Coefficients describing the magnitude and polarity of the relationships between predictors and essay scores are represented by β coefficients for student-level predictors (e.g. $\beta_2 SES_{ij}$) and γ coefficients for classroom-level predictors (e.g. $\gamma_{01} HighTD_j$).

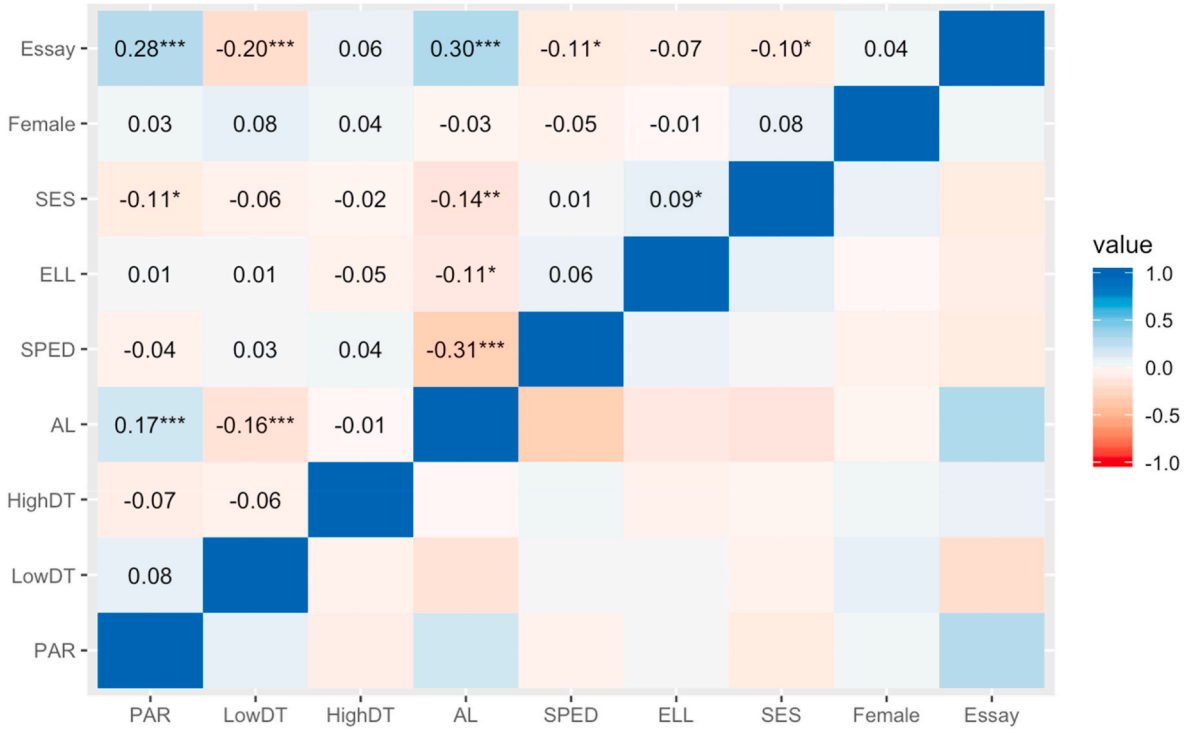

Fig. 1. Correlation matrix of essay scores, predictors (high and low dialogic talk), and controls (participation, academic language scores, special education status, English learner status, and sex)

Fig. 1, statistically significant negative correlations were found between low dialogic talk (LowDT) and essay scores ($r = -0.20$, $p < .001$). The direction and magnitude of these correlations are an early indication of the hypothesized relationship between low dialogic talk and essay scores. However, there was no simple correlation between essay scores and measures of high dialogic talk scores for the teachers' talk (HighDT).

Correlations between outcome and controls. The strongest correlation is observed with prior academic language skills (AL), where there is a positive, statistically significant correlation ($r = 0.30$, $p < .001$). Other control variables that significantly correlated with the outcome are participation rate (PAR) ($r = 0.28$, $p < .001$), special education status (SPED) ($r = -0.11$, $p < .05$) and socioeconomic status (SES) ($r = -0.11$, $p < .05$).

3.2. Multilevel model results

Predictors. A multilevel model with all predictors and controls was fitted to answer the research question outlined above (Table 5, Model 3). The inclusion of both high and low dialogic features of teacher talk in the same model reflects the co-existence of these forms of talk in any given classroom. That is, a teacher with many instances of HighDT (i.e. follow-ups and open-ended questions) may have also asked several semi-open and quiz-like questions during the same discussion (LowDT). Thus, to understand the contribution of high levels of dialogic teacher talk on persuasive essay scores, we controlled for the low levels of dialogic talk, and vice versa.

Measures of dialogic talk were regressed on essay scores using a random intercept (by topic, teacher, and student), random slope (by participation rate and topic) multilevel model to examine the predictive relationships among teacher discussion features and student writing. The hypothesized relationship between high dialogic talk and essay scores was confirmed. Specifically, high dialogic talk was found to positively predict essay scores when accounting for low levels of teacher dialogic talk, participation rate, prior academic language, and students' demographic features ($\beta_{MaxTD} = 0.36$, $p < .05$) (Fig. 2). Thus, on average, for every additional instance of high dialogic talk per minute of discussion time (e.g. asking contestable questions or encouraging students to respond to one another), we observe a 9% increase in essay scores.[8] Consistent with the hypothesis that low dialogic talk would be negatively related to persuasive essay scores, results show that semi-open and quiz-like teacher questions (low dialogic talk) were associated with a statistically significant decrement in essay scores, when accounting for measures of high teacher dialogic talk, student participation rate, prior academic language scores, and demographic variables ($\beta = -0.24$, $p < .05$) (Fig. 2). Thus, on average, every additional occurrence of low dialogic talk per minute of discussion was associated with a decrease of 6% in essay scores.

Controls. The only statistically significant relationship between control variables and the outcome was found between prior academic language skills (AL) and essay score ($\beta_{AL} = 0.17$, $p < .001$). Thus, participation rate, socioeconomic status, sex, ethnicity, special education status, and English learner status were not statistically significantly predictive of essay scores (Fig. 2).

4. Discussion

This study sought to explore a classroom feature that, although common among all classrooms in this sample, was implemented differently from one classroom to another. Though the classrooms investigated employed a discussion-and debate-focused curriculum, dialogic interactions varied. Even when talk was dialogic, it was not conducted uniformly across all classrooms, as teachers varied widely in how they facilitated classroom talk and which formats they used (e.g. fishbowl discussion vs. all-class). Nonetheless, the study aimed to control as many variables as possible and examine teacher questions and follow-ups in order to test the question, *What is the relationship between dialogic- and less-dialogic teacher talk and persuasive writing?* As both classroom discussion and debate and persuasive writing involve engaging in a

[8] Essays are scored out of four. Thus, $0.36 \div 4 = 0.09$ (9% increase per minute of discussion time.).

Table 5
Fixed and random effects multilevel models.

Variable	Dependent Variable: Essays		
	Estimate (SE)		
	Model 1	Model 2	Model 3
Fixed Effects			
Intercept	1.81***	1.96**	**1.99*** **
	(0.48)	(0.47)	**(0.26)**
High Dialogic Talk (HighDT)	0.02		**0.36***
	(0.17)		**(0.16)**
Low Dialogic Talk (LowDT)		−0.14	**−0.24***
		(0.08)	**(0.11)**
Participation Rate (PAR)	0.16	0.16	**0.14**
	(0.15)	(0.15)	**(0.09)**
Academic Language (AL)	0.17***	0.17***	**0.17*** **
	(0.04)	(0.04)	**(0.04)**
Socioeconomic Status (SES)	−0.00	−0.01	**−0.00**
	(0.09)	(0.09)	**(0.09)**
English Learner Status (ELL)	−0.18	−0.19	**−0.17**
	(0.14)	(0.14)	**(0.14)**
Special Education Status (SPED)	−0.08	−0.07	**−0.09**
	(0.10)	(0.10)	**(0.10)**
Ethnicity:	−0.09	−0.08	**−0.07**
White	(0.08)	(0.08)	**(0.08)**
Latinx	0.03	0.03	**0.03**
	(0.08)	(0.08)	**(0.08)**
Other	−0.08	−0.07	**−0.06**
	(0.14)	(0.14)	**(0.15)**
Female	0.09	0.09	**0.09**
	(0.06)	(0.06)	**(0.06)**
Random Effects			
Observations	471	471	471
Groups:			
Students	369	369	369
Topics	25	25	25
Teachers	24	24	24
Students (Intercept) Std.Dev.	0.23	0.23	**0.22**
Topic (Intercept) Std.Dev.	1.63	1.60	**0.33**
Participation (Slope) Std.Dev.	0.55	0.54	**0.25**
Teachers (Intercept) Std.Dev.	0.13	0.09	**0.26**
Residual Std.Dev.	0.54	0.54	**0.55**
AIC	972.3	970.9	**977.8**
DIC	874	870	**870**
Deviance	906.4	903.3	**905.9**

*p < .05 **p < .01 ***p < .001.

communicative activity where students externalize their thinking, develop a stance, and persuade an audience using reasoning or evidence, we tested the hypothesis that dialogic teacher talk would support students' persuasive writing in classrooms where both the topic of discussion and persuasive writing were the same.

In addition to classroom differences, there were individual differences among students. While most students in this sample came from low-income households, students differed in their race/ethnicity, sex, special education status, English learner status, and prior academic language skills. Though other differences certainly exist among students, these variables were accounted for so as to understand the role of discussion in writing, regardless of the ethnic, socioeconomic status, sex, and academic backgrounds of students.

By categorizing teacher talk moves into those indicative of higher and lower dialogic talk, and by controlling for demographic variables as well as students' prior academic language skills, we were able to identify specific talk features during whole classroom discussion that relate to post-discussion persuasive essays. The outcome in this study was also scored using a simple rubric that teachers can easily utilize to assess the argumentative structure shown in their students' persuasive essays.

Findings support the theoretical assumptions that place semi-open and quiz-like teacher questions at the lower end of the dialogic spectrum, and suggest that these discussion features are not as conducive to supporting the communicative activity that underlies persuasive writing. To successfully persuade an audience, writers present claims to a non-present interlocutor and support their claims with evidence and reasoning, while also acknowledging (and dismantling) counterarguments. While contestable questions allow for the exploration of rich and diverse ideas and perspectives upon which students can later draw in their writing, discussions that feature low levels of dialogic talk (through semi-open and quiz-like questions, in this case), stifle opportunities for extended engagement and may therefore limit students' capacity to persuade an audience in writing.

Similarly, findings partially support the hypothesis that teacher discussion features deemed to be dialogic are more conducive to promoting student writing (Table 5). Though the study does not allow for causal inferences, these results highlight possible pedagogical approaches (e.g. prioritizing open-ended, contestable questions) that may be beneficial for promoting students' writing outcomes on the topic of discussion.

If found to be causal, this finding could encourage teachers to add to their repertoire of questions or discussion moves during a whole class discussion, when their goal is to support students' persuasive writing. For example, one component of teacher talk explored in this study is type of teacher questions: a) contestable questions; b) semi-open questions that provide more latitude but are not truly contestable; and c) closed, quiz-like questions with known answers. Contestable questions are highly dialogic in that they prompt a wide range of responses. These types of questions have been shown to be positively associated with essay scores (as part of high dialogic talk). One mechanism through which this may occur is that dialogic questions, when asked in a whole class setting, invite students to explore various ideas and perspectives on a given topic. To use an example from the CCDD project from which these data were drawn, when a teacher asks, "Why do you think the Pharaohs were good leaders?" students will explore multiple views and opinions in their responses.

In many cases, contestable questions also require students to use evidence and/or reasoning to support their claims, as the very nature of these questions does not accept a response that is not supported by evidence or reasoning. That is, it is not sufficient to answer a question like "Do we need laws to regulate our behavior" by saying, "Yes, we need laws." In order to fully answer the question, students must support their responses with textual or other forms of evidence and reasoning. This is also illustrated through other contestable questions such as, "Should everyone learn a second language?" While this question may initially produce a "Yes/No" response by a student, it more easily allows a teacher to follow-up with students in order to press for reasoning, another highly dialogic teacher move explored in this study.

Additionally, when asked in whole class settings, contestable questions such as the ones examined in this paper expose students to various viewpoints that they may have otherwise not encountered or considered. When writing persuasively, exposure to different perspectives may in turn allow students to defend their positions more strongly, as they can draw upon multiple viewpoints which they can then either support or refute using evidence and reasoning they may have practiced during whole class discussion. It can be argued that even students who are not active participants may benefit from being part of classrooms that prioritize contestable questions, as their simple presence during a dialogic talk can help them gather evidence and viewpoints needed for persuasive writing.

These findings lend empirical support to earlier criticisms of quiz-like questions for their inability to promote critical thinking and reasoning skills (Alexander, 2008; Cazden, 1988). In comparison to quiz-like questions, semi-open questions offer more latitude in that they generally deal with processes and procedures. Because they elicit an expected set of responses, they are not truly contestable as they aim to arrive at a "correct" or agreed-upon response. For example, a teacher who asks multiple students to explain the process of becoming a citizen will expect to hear the "correct" response, even if students' responses will differ in form. Like quiz-like questions, semi-open questions may promote students' recall skills; however, in this study, both types of

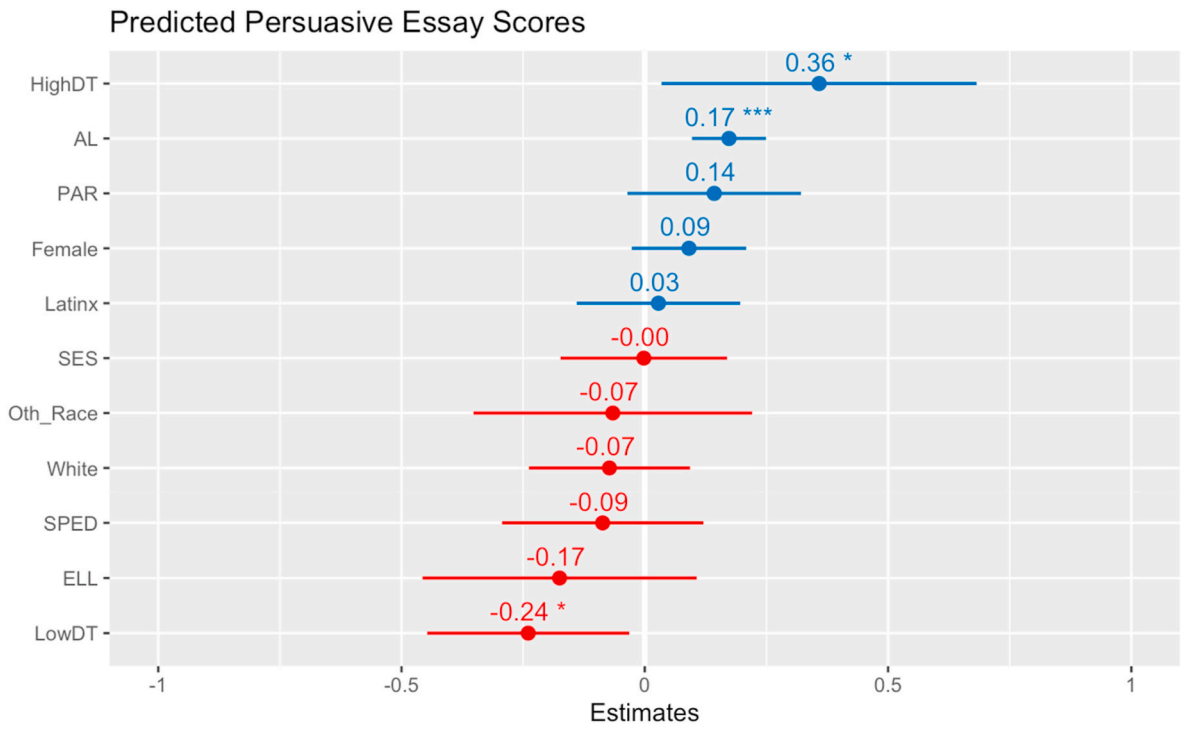

Fig. 2. Final model showing predicted persuasive essay scores with significance levels and confidence intervals for predictors and controls.

questions were negatively predictive of students' persuasive essay scores. Thus, it could be hypothesized that one mechanism for improving persuasive writing would be for teachers to center their discussions around contestable questions, thereby providing students with opportunities to refine their own thinking and engage with multiple perspectives, which may help students acquire the skills necessary to develop a stance and persuade an audience.

4.1. Limitations and future directions

Though this work echoes a shift in the field toward linking specific teacher talk moves to student outcomes (e.g. Howe et al., 2019), several limitations exist in this study. Given that this study employs data from the intervention group alone (due to the unavailability of post-discussion essays in control classrooms), teachers are aware of the Word Generation program's overall goals of fostering comprehension through discussion and debate. Thus, it would be worth comparing these results to classrooms that are not implementing a discussion-based curriculum in order to analyze whether results found in this study are replicated in classrooms that are more representative of average (non--discussion based) classrooms. It is important to note, however, that even in this sample, there was wide variation in what teachers considered 'discussion,' as some classrooms were characterized by high levels of dialogic talk (e.g. open-ended questions and many follow-ups with students) while others were characterized by minimally dialogic teacher talk (e.g. semi- or quiz-like questions that invite IRE rather than dialogic exchanges.)

Another limitation pertains to the fidelity of implementation of persuasive essays. It would be worthwhile for future analyses to consider only classrooms in which full implementation occurred, or at least compare those with high and low implementations. Furthermore, the use of one data-point as an indication for a teacher's dialogic talk score is not ideal, as an audio-recorded discussion on any given day may not reflect a typical classroom discussion. Longitudinal studies that explore various discussion factors, including classroom talk and participation levels over time, may help solidify our understanding of the relationship between talk and writing. Yet one strength of this study is the strong linkage between the focus of the discussion and the contents of the persuasive essay. Such studies can complement larger studies where talk is studied over time and the outcome variable is performance on more general measures of academic performance.

5. Conclusions

This study sought to empirically analyze the relationship between teacher classroom talk and students' persuasive writing. The work presented in this paper aligns with the increased emphasis on discussion and writing in all subject areas (National Governors Association, 2010), as well as the need to build upon existing scholarly work that points toward prioritizing certain forms of talk over others (e.g. Michaels et al., 2008; Waggoner et al., 1995) and examine predictive relationships among teacher talk features and any number of student outcomes (e.g. Howe et al., 2019; Kuhn et al., 2016; Murphy et al., 2009). These findings open up the possibility for further investigations that incorporate both student and teacher talk features in whole class settings to help uncover the best ways teachers can support students' literacy outcomes.

CRediT authorship contribution statement

Shireen Al-Adeimi: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Funding acquisition. **Catherine O'Connor:** Conceptualization, Resources, Writing - review & editing, Supervision, Validation.

Acknowledgements

This work was supported by the Institute of Education Sciences under Grant R305F100026 and by the American Educational Research Association Minority Dissertation Fellowship Program in Education Research.

References

Alexander, R. J. (2005). *Teaching through dialogue: The first year*. London: *London Borough of Barking and Dagenham*.

Alexander, R. J. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). Dialogos.

Alexander, R. (2020). *A dialogic teaching companion*. Routledge.

Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics (C. Emerson, Trans.* (Vol. 8). Minneapolis: University of Minnesota Press.

Cazden, C. (1988). Classroom discourse: The language of teaching and learning. In *(2012). Writing section*. Portsmouth, NH: Heinemann. College Board. Retrieved January 17, 2015 from http://professionals.collegeboard.com/testing/sat-reasoning/about/sections/writing.

Crowell, A., & Kuhn, D. (2014). Developing dialogic argumentation skills: A 3-year intervention study. *Journal of Cognition and Development, 15*(2), 363–381.

Dewey, J. (1938). *Experience and education*. New York: Touchstone.

Gee, J. P. (2007). *Social linguistics and literacies*. Routledge.

Grossman, P., Greenberg, S., Hammerness, K., Cohen, J., Alston, C., & Brown, M. (2009). April). Development of the protocol for language arts teaching observation (PLATO). In *Annual meeting of the American educational research association, san diego, CA*.

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher–student dialogue during classroom teaching: Does it really impact on student outcomes? *The Journal of the Learning Sciences*, 1–51.

Ketch, A. (2005, September). Conversation: The comprehension connection. In , *Vol. 59. The reading teacher*. Newark, DE: International Reading Association, 1.

Kress, G. (1994). *Learning to write* (2nd ed.). New York: Routledge.

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science, 22*(4), 545–552.

Kuhn, D., Hemberger, L., & Khait, V. (2016). Dialogic argumentation as a bridge to argumentative thinking and writing. *Journal for the Study of Education and Development, 39*(1), 25–48.

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction, 31*(4), 456–496.

Mehan, H. (1998). The study of social interaction in educational settings: Accomplishments and unresolved issues. *Human Development, 41*(4), 245–269.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion. *Socializing intelligence through talk and dialogue*, 347–362.

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies In Philosophy and Education, 27*(4), 283–297.

Murphy, P. K., Firetto, C. M., Greene, J. A., & Butler, A. M. (2017). *Analyzing the talk in quality talk discussions: A coding manual*. The Pennsylvania State University, University Park, PA. https://doi.org/10.18113/S1XW64. Unpublished manuscript.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101*, 740–764.

Myhill, D., & Fisher, R. (2010). Editorial: Writing development: Cognitive, sociocultural, linguistic perspectives. *Journal of Research in Reading, 33*, 1–3.

America's Schools and Colleges National Commission on Writing. (2006). *Writing and school reform*. Washington, DC: College Entrance Examination Board.

National Governors Association Center for Best Practices. (2010). *Council of chief state school officers. Common Core state standards*. Washington D.C: National Governors Association Center for Best Practices, Council of Chief State School Officers.

Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults. *Language, Speech, and Hearing Services in Schools, 36*(2).

Nystrand, M. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Teachers College Press.

Osborne, J. F., Borko, H., Fishman, E., Gomez Zaccarelli, F., Berson, E., Busch, K. C., … Tseng, A. (2019). Impacts of a practice-based professional development program on elementary teachers' facilitation of and student engagement with scientific argumentation. *American Educational Research Journal, 56*(4), 1067–1112. https://doi.org/10.3102/0002831218812059.

O'Connor, C., Michaels, S., & Chapin, S. (2015). "Scaling down" to explore the role of talk in learning: From district intervention to controlled classroom study. In L. Resnick, C. Asterhan, & S. Clarke (Eds.), *Socializing intelligence through academic Talk and dialogue. Washington D.C.* (pp. 111–126). American Educational Research Association.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment scoring System™: Manual K-3*. Paul H Brookes Publishing.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from https://www.R-project.org/ .

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2- 3), 155–175.

Reznitskaya, A., Glina, M., & Oyler, J. (2012). *Dialogic inquiry tool*. Montclair, NJ: *The Institute for the Advancement of Philosophy for Children*.

Salahu-Din, D., Persky, H., & Miller, J. (2008). *The nation's report card: Writing. 2007*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Snow, C. E., Lawrence, J. J., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research of Educational Effectiveness, 2*(4), 325–344.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York: Cambridge University Press.

Uccelli, P., Barr, C. D., Dobbs, C. L., Phillips Galloway, E., Meneses, A., & Sanchez, E. (2015). Core Academic Language Skills (CALS): An expanded operational construct and a novel instrument to chart school-relevant language proficiency in pre- adolescent and adolescent learners. *Applied PsychoLinguistics*, (36), 1075–1109.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Waggoner, M., Chinn, C. A., Yi, H., & Anderson, R. C. (1995). Collaborative reasoning about stories. *Language Arts, 72*, 582–589.

http://professionals.collegeboard.com/testing/sat-reasoning/about/sections/writing, (2012–. (Accessed 17 January 2015).