# Classroom Observation Opportunity to Learn November 2013

**Project READI Technical Report # 4**

Cindy Litman, Stacy Marple, Cynthia Greenleaf,
With Michael Bolz, Irisa Charney-Sirott, Lisa
Richardson, and Allison Hall

PROJECT READi

inquirium

Northern Illinois University

NORTHWESTERN UNIVERSITY

UIC UNIVERSITY OF ILLINOIS AT CHICAGO

WestEd

Citation for this Report:. Litman, C., Marple, S., & Greenleaf, C., with Bolz, M., Charney-Sirott, I., Richardson, L., & Hall, A. (2013). *Classroom observation opportunity to learn analysis, November 2013.* Strategic Literacy Initiative, WestEd. READI Technical Report #4.
Retrieved from URL: projectreadi.org

The study reported here involved an extensive team of observers. We wish to acknowledge the support of Megan Hughes, Jackie Popp, Diane Puklin, Tanya Solomon, Teresa Sosa, and Mary Pat Sullivan in collecting the observational data analyzed here. In addition, we wish to acknowledge the contributions of MariAnne George, Susan Goldman, Carol Lee, Taffy Raphael, and Cyndie Shanahan to the observational and interview protocols developed to guide this data collection. Finally, we thank Angela Fortune for her contributions to the research questions guiding the analysis of this data.

Please send us comments, questions, etc.: info.projectreadi@gmail.com

Project READI was supported by the *Reading for Understanding (RFU)* initiative of the Institute for Education Sciences, U. S. Department of Education through Grant R305F100007 to the University of Illinois at Chicago from July 1, 2010 – June 30, 2016. The opinions expressed are those of the authors and do not represent views of the Institute or the U. S. Department of Education.

Project READI operated as a multi-institution collaboration among the Learning Sciences Research Institute, University of Illinois at Chicago; Northern Illinois University; Northwestern University; WestEd's Strategic Literacy Initiative; and Inquirium, LLC. Project READI developed and researched interventions in collaboration with classroom teachers that were designed to improve reading comprehension through argumentation from multiple sources in literature, history, and the sciences appropriate for adolescent learners. Curriculum materials in the READI modules were developed based on enacted instruction and are intended as case examples of the READI approach to deep and meaningful disciplinary literacy and learning.

As of October 2013, Project READI researchers have segmented 71 videos from year 1 classroom observations. Each video was segmented on six dimensions: (1) *Grouping*; (2) *Teacher-Student interaction*; (3) *Opportunity to Learn: Task*; (4) *Opportunity to Learn: Content Delivery*; (5) *Student Behavior: Learning;* and (6) *Student Behavior: Engagement*. This report provides a description of the dimensions and codes within each dimension, coder training and reliability, and findings from a descriptive analysis of opportunity to learn based on the segmented lessons.

## I. Background

The primary goal of the Year 1 observations was to facilitate rapid prototyping of Evidence-Based Argument Instruction Models (E-B AIMs) based on the kinds of texts, tasks, participation structures, and tools that are associated with evidence of substantial student engagement with disciplinary literacies and reasoning with multiple sources. Observations were conducted in 79 history, literature, and science lessons. We approached the observations with the understanding that many of the classrooms we would observe did not necessarily have established argumentation routines, or may only have emergent ones. However, we reasoned that the observed lessons could reveal other disciplinary literacy practices as well as pitfalls that could potentially inform the rapid-prototyping work. Observed lessons reflected a wide range of established and emergent disciplinary literacy and argumentation practices, and student engagement and learning.

A preliminary analysis based on field notes, teacher interviews, and classroom materials suggested that the classroom observation data holds considerable promise for advancing knowledge of specific features of instruction and classroom life that mediate student engagement and learning from higher level disciplinary literacy tasks. Consequently, the Year 1 classroom observation strand of the READI project expanded beyond its original goal of informing rapid prototyping of E-B AIMS, to potentially illuminate more generally how instructional features of tasks, texts and classroom climate may influence students to acquire argumentation literacy in multiple subject areas and varying grade levels in real instructional settings.

### Recruitment and selection of teachers

Observations were conducted in classrooms located in the San Francisco Bay area and in the greater Chicago area. Identification of teachers/classrooms for observations followed somewhat different procedures in the two locations so we describe them separately here. A more detailed description of the recruitment and selection process, teacher characteristics, observation protocol, and distribution of observed lessons by month, subject area and grade level is found in Part 1 of the Classroom Observation Report.

San Francisco Bay Area Sites: From the network of Reading Apprenticeship teachers, Project READI team members identified experienced teachers in middle and high school whose literacy implementation in subject areas was believed to hold some promise to inform the development of new interventions (E-B AIMS). These teachers were invited to participate in classroom-based research with the aim of identifying features of instruction that were marked by high engagement and appeared to develop advanced comprehension skills. Reading Apprenticeship (RA) is a model of academic literacy instruction designed to improve literacy skills and academic achievement for all students, including struggling readers. In this instructional framework, students are given extended

opportunities to engage in close reading of a wide range of texts with instructional support—both textbooks and ancillary materials, such as primary sources, lab manuals, journal articles and trade books. Through an "apprenticeship" process and ongoing metacognitive conversation focused on disciplinary reading and thinking processes, content-area teachers explicitly model and guide students in practicing the tacit reasoning processes, strategies, and discourse rules used by successful readers and writers. However, prior to Project READI, the Reading Apprenticeship framework and professional development had not focused explicitly on cross-textual analysis and argumentation.

Because the sample of Reading Apprenticeship teachers included only one science teacher, we identified additional science teachers for our observations who had not participated in Reading Apprenticeship professional development but were recommended by WestEd colleagues as strong teachers of science.

Chicago Area Sites: Teachers and schools for observations in the Chicago Area were nominated by Project READI team members who had worked with area schools and teachers. Team members nominated teachers based on a variety of attributes, including teachers (1) they knew to be engaging in instruction designed to foster disciplinary literacies in history, science, and /or literature; (2) who were participating in implementing Cultural Modeling practices; and/or (3) who were reported to have established classroom participation structures that supported high student engagement. We also solicited teacher nominations from the CPS district leadership in literacy, social sciences, and sciences.

These differences in recruitment and selection thus resulted in a quasi-experimental design with two groups of teachers nominated by Project READI team members—approximately half of whom had participated in long-term professional development around an instructional framework focused on support for close reading, and the other half, likewise identified as good teachers in their disciplines, who did not share a common professional development experience or instructional framework.

## II. Review of Preliminary Analysis

The theoretical framework and research questions articulated in the Project READI proposal provided a starting point for our data collection. This lens directed our attention broadly to lesson texts, tasks and classroom culture, but stipulated few specific indicators of the classroom context that may influence the development of high levels of literacy engagement and achievement in real classrooms. In order to explain how features of instruction and classroom life mediate student engagement and learning from higher level disciplinary literacy tasks, in a preliminary analysis overlapping data collection, observations were coded for dimensions of text use, tasks and classroom culture. The preliminary analysis was based on field notes, lesson artifacts and teacher interviews, and did not include systematic coding of audio- or videotapes of observed lessons.

Initial coding and analysis utilized a "start list" of broad descriptive categories reflecting the conceptual framework and research questions articulated in the Project READI proposal (Miles & Huberman, 1994): lesson architecture, texts and text characteristics, tasks and task support, classroom culture, and student behavior. Within these broad categories, we approached the analysis using open and axial coding from grounded theory research.

*Open Coding* is "the interpretive process by which data are broken down analytically" (Corbin and Strauss, 1990, p. 423). In open coding, incidents, actions, interactions and other features of the data are compared for similarities and differences and grouped together to form categories and subcategories. For example, our theoretical framework and research questions directed our attention to argumentation, which we defined *a priori* as making a claim or assertion that is supported by evidence that connects to the claim in a principled way (Toulmin, 1958). Observers identified numerous episodes in which students generated claims and presented evidence to support their claims in principled ways. Researchers analyzing the observation data coded these as "argumentation." However, argumentation episodes varied in a number of ways and were further broken down into subcategories and labeled. Some argumentation arose informally in the context of collaborative meaning-making around text. Researchers coded these as "interactive argumentation." Some argumentation arose from tasks designed to teach key disciplinary principles, frameworks and understandings. Researchers labeled these as "arguing to learn." Some argumentation emphasized the acquisition of canonical forms of argument. We labeled these "learning to argue" tasks. We also identified tasks that *potentially* offered students the opportunity to make a claim or assertion supported with evidence, but failed to do so as a result of elements of task design or enactment. We labeled these as "missed opportunities." Missed opportunities for argumentation were further broken down into subcategories such as "classification"—tasks that focused on selecting the right answer from a limited set of teacher-generated possibilities and did not require students to provide supporting evidence.

Our analysis also included axial coding to "scrutinize the data to determine, what are the conditions that gave rise to that kind of work, in what context was it carried out, by what action/interactions did it occur, and what were the consequences?" (Corbin and Strauss, 1990, p. 423). Because our research questions focused on how features of instruction may influence the development of high levels of literacy engagement and achievement in real classrooms, we were especially interested in making conceptual linkages between texts, tasks, classroom culture and student engagement and learning, as well as linkages among and between texts, tasks and classroom culture.

## VI. Preliminary Findings

Major themes and practices related to the acquisition of argumentation literacy that emerged from the preliminary analysis are summarized in Table 1.

A major finding from the preliminary analysis was the importance of attending to "building blocks" of evidence-based argumentation as well as to more explicit E-BA activities. While E-BA was easily recognized in explicit argumentation tasks, the roots of E-BA were likewise present in close reading and discussion activities that required students to read with attention to evidence and interpretation, such as generating questions about a text, continuously revising a KWL chart as students read multiple texts on a topic, evaluating a source, or generating an essay topic and defending its importance with quotes from the text. Multiple close readings of a text supported students in moving to more elaborated meanings required for mature E-BA. Students benefited from an initial reading for meaning, followed by subsequent readings focused on interpretive practices of the discipline. In addition, our analysis revealed a close relationship between E-BA and collaborative meaning-making routines. Much rich argumentation took place in the form of interactive argumentation in the context

of close reading and collaborative meaning-making (Berland & Reiser, 2008; Chinn & Anderson, 1998). These building blocks were often not framed by teachers in argumentation language.

Argumentation tasks themselves assumed many forms, including argumentative inquiry as a tool for the construction and understanding of disciplinary knowledge and practices, tasks focused on teaching the language and structure of argument, tasks where attention to the form of argument was an outgrowth of argumentation inquiry, and interactive argumentation that was a byproduct of collaborative meaning-making.

Finally, it was clear that supports for argumentation literacy emerging from these classroom observations were situated in all three components of the observation and analytic framework—texts, classroom activities, and classroom culture—and in the synergy between these elements of instruction.

Table 1. Themes and Practices that Support the Acquisition of Argumentation Literacy from the Preliminary Analysis

| |
|---|
| • Epistemological orientation that positions tasks and texts as inquiry vs. as fact acquisition, and promotes and facilitates students construction, representation and evaluation of knowledge |
| • Close reading characterized by approaching texts to understand vs. to find information |
| • In-class reading and comprehending that affords opportunities for teacher and peer support and collaborative meaning-making |
| • Reading routines, tools and strategies that support negotiation of meaning about texts and explicit connections between and across multiple sources |
| • Participation structures that support student ownership, agency, engagement and participation, and convey authority to students to shape the topic and conversation, evaluate ideas—i.e., to do the work of sense-making |
| • Recursive cycles of whole class, individual and group work that provide students with opportunities to practice discipline-specific ways of discussing the texts under study individually, and in teacher- and peer-directed settings |

This preliminary analysis also led to a set of broad codes and operational definitions related to evidence-based argumentation with multiple texts. The codes capture instructional practices that appeared to mediate student engagement and learning from high level literacy tasks: (1) texts; (2) close reading; (3) argumentation; (4) disciplinary knowledge building tasks; (5) teacher support for learning; (6) instrumental support for learning; (7) epistemological framing; and (8) participation structures. Subcodes capture both promising practices and missed opportunities, cases where lesson features have some potential to foster argumentation literacy but fail to do so. In addition, we generated two codes related to student behavior—(9) engagement and (10) learning.

Definitions of these codes, which laid the foundation for subsequent analysis, are found in Table 2.

**Table 2. Operational Definitions of Broad Codes Emerging from the Preliminary Analysis**

| Theme | Definition |
|---|---|
| Features of instruction and classroom climate | |
| Texts | Features of texts and text use including:<br>a. Texts and text properties<br>b. How texts are used<br>c. How texts are used in relationship with other texts |
| Close reading | Interactive negotiation of meaning at the local and global levels to unearth and evaluate possible meanings, individually or collaboratively. Characterized by approaching texts to understand vs. to find information. *Missed opportunities are tasks with the potential to foster close reading because there are possible supports for unearthing and evaluating possible meanings, but features of task and/or classroom life fail to elicit these.* |
| Argumentation | Making a claim or assertion that is supported by evidence that connects to the claim in a principled way. Involves consideration/deliberation of multiple possibilities and/or viewpoints. *Missed opportunities are tasks with the potential to foster argumentation because there are multiple possible understandings to negotiate, but features of texts, task and/or classroom life fail to elicit these.*<br>Identify emphasis of argumentation:<br>a. Arguing to learn: Argumentation as a tool for the construction and understanding of disciplinary knowledge and practices<br>b. Learning to argue: Explicitly teaching language, structure and principles for argument and asking students to apply the structure to learn disciplinary argument |
| Disciplinary knowledge | Discipline-specific epistemologies and inquiry practices in reference to the overarching frameworks, concepts and themes of the disciplines. *Missed opportunities are tasks with the potential to foster disciplinary knowledge, but features of task and/or classroom life fail do this.* |
| Teacher support for learning from texts and higher level literacy and disciplinary knowledge tasks | Teacher modeling, guidance and support for learning and practicing meaning-making about text, argumentation and disciplinary knowledge. *Missed opportunities are instructional moves with the potential to support learning, but that fail do this.* |
| Instrumental support for learning from texts and higher level literacy and disciplinary knowledge tasks | Routines, tools and strategies that support learning, such as metacognitive reading routines (e.g., Talking to the Text/annotating, think aloud), notetakers (evidence/interpretation, disciplinary notetakers), etc. *Missed opportunities are routines, tools and strategies with the potential to support learning, but that fail do this.* |
| Epistemological framing | Signals communicated by teacher and students through tone of voice, word choice, interactions, routines, and explicit instructions and comments that convey understandings and expectations of a task or activity (e.g., "doing science" vs. "doing the lesson" (Jiménez-Aleixandre, Rodríguez, and Duschl, 2000)).<br>Identify framing that instantiates a(n):<br>a. *Procedural display orientation* that positions tasks and texts as information vs. inquiry, and promotes and rewards "doing school" over reading and learning for understanding |

| | |
|---|---|
| | (Philips, 1974).<br>Identify participation structures that:<br>a. Communicate that the teacher vs. students has authority to set the topic, direct conversation, evaluate ideas—i.e., to do the work of sense-making<br>b. Support student ownership, agency, engagement and participation, and convey authority to students to shape the topic and conversation, evaluate ideas—i.e., to do the work of sense-making |
| Indices of student engagement and learning | |
| Student engagement | Evidence of engagement and effort in relation to reading, argumentation and disciplinary knowledge building, including persistence and grappling, student ownership, agency and extended instructionally focused student talk. *Missed opportunities are evidence of lack of agency, engagement and participation* |
| Student learning | Evidence of reading comprehension, argumentation and disciplinary knowledge building reflected in construction, representation and evaluation of knowledge, and appropriation and use of disciplinary language, literacies, thinking and reasoning dispositions, skills and knowledge. *Missed opportunities are evidence that the enactment of the lesson does not result in reading comprehension, argumentation and disciplinary knowledge building* |

## III. Computer Assisted Segmentation of Classroom Videos

Unlike findings from the preliminary analysis summarized above, findings that are the focus of the research reported here are based on video data. Observed lessons were audio- and videotaped to capture evidence of student engagement in processes we hypothesized to be central to content learning and argumentation discourse. In order to more systematically investigate opportunities for students to learn evidence-based argumentation with multiple texts, we segmented the video data[1] using 31 codes related to six lesson dimensions: 1. Grouping, 2. Teacher-Student interaction, 3. Opportunity to Learn: Task, 4. Opportunity to Learn: Content Delivery, 5. Student Behavior: Learning, and 6. Student Behavior: Engagement.

Project READI researchers coded all Year 1 classroom observations from both sites for which we had video data of sufficient aural and visual quality for accurate coding. This resulted in 71 coded lessons—37 from Chicago and 34 from San Francisco Bay Area schools. Table 3 shows characteristics of segmented lessons at the two sites.

Table 3. Bay Area (N=34) and Chicago (N=37) Segmented Lessons

| | Discipline | | | | | | |
|---|---|---|---|---|---|---|---|
| Grade Band | History (N=24) | | Literature (N=34) | | Science (N=13) | | Grade Total |
| | Chicago | Bay Area | Chicago | Bay Area | Chicago | Bay Area | |
| 11-12 | 2 | 6 | 6 | 6 | 1 | 0 | 22 |
| 9- 10 | 7 | 3 | 5 | 5 | 6 | 2 | 28 |
| 6-8 | 2 | 4 | 7 | 4 | 1 | 3 | 21 |
| Total | 11 | 13 | 18 | 16 | 8 | 5 | 71 |

[1] Coding and analysis of video data was done using NVivo9 qualitative analysis software (QSR International).

## IV. Instrumentation for Video Segmentation

### A. Development and Definitions of Video Segmentation Dimensions and Codes

The segmentation work reported here was intended as a first pass through the video data, and dimensions and codes reflect broad features of the classroom environment related to the acquisition of argumentative literacy. The coding scheme went through multiple iterations from July to November 2012. Modifications focused on finding the appropriate grain-size and generating broad dimensions and smaller-grained codes within each dimension with the potential to support both low inference descriptive analysis and explanatory analysis around a range of research questions.

Video segmentation dimensions and codes drew on the themes and coding scheme that emerged from the preliminary analysis. Correspondence between themes from the preliminary analysis and the video segmentation dimensions and codes is shown in Appendix B.

The following section elaborates video segmentation coding scheme, providing an overall explanation of the six lesson dimensions and descriptions of the codes within each dimension.

**Video Segmentation Coding Scheme**

*Grouping*: The grouping dimension, which serves as a proxy for opportunities and expectations for participation, focuses on how students are receiving input or engaging in an activity or task, not on seating arrangement. For example, if students are engaged in individual silent reading, grouping is coded as *Individual*, even if students are seated at a table group. Every segment of video is coded to a grouping code. The Grouping dimension consists of consists of four codes:

1. *Individual*: Students work independently.

2. *Pairs*: Students work in pairs. Partner interactions of two minutes or more embedded in another grouping structure (e.g., pair-shares in the context of whole class discussion) are coded to both groupings.

3. *Small group:* Students are divided in small groups that they generally run themselves. Small group interactions of two minutes or more embedded in another grouping structure are coded to both groupings.

4. *Whole class*: Teacher interacts with the whole class at once.

*Teacher-Student Interactions*: It was not possible to code nuanced features of teacher modeling, guidance and support for learning and practicing meaning-making about text, argumentation and disciplinary knowledge during our first pass through the video data. Rather, this dimension consists of three broad-stroke codes reflecting the nature of the teacher-student interaction during the lesson:

1. *Housekeeping/management:* Segments of at least 30 seconds with a procedural focus related to general classroom business or non-instructional activity. Examples include passing out and collecting materials and student work, changing groupings, reviewing class assignments and school activities, dispensing rewards, etc.

2. *Teacher gives instructions*: This code is used exclusively for setting up a task. If teacher provides additional direction or instructional support while students are engaged in a task, teacher-student interaction is coded to *Teacher modeling, guidance and support*.

3. *Teacher modeling, guidance and support*: This code is used for segments in which the teacher provides instructional support for student learning. If the focus of teacher-student interaction is exclusively on procedures or classroom management, the interaction is coded to *Housekeeping/classroom management*.

Student opportunity to learn in observed lessons were coded on two dimensions: *Task* and *Content Delivery*.

While video data did not afford a close look at text features and usage, we did look at opportunities students had to work with text, in contrast to other ways of learning content. *Opportunity to Learn: Content Delivery* refers to the mechanism by which students are exposed to content, both in the input phase that generally occurs at the beginning of an assignment and throughout, as students continued to work with lesson content. The three codes within this dimension include:

1. *Working with text*. This construct refers to tasks in which students work with text(s), individually or collaboratively. While text is defined broadly to include reading a wide range of materials, including graphics, etc., from a wide range of sources, including computer screens, not just connected text and traditional print material, what distinguishes *Working with text* from other *Content Delivery* codes is that it affords the ability to revisit texts and confers responsibility on the reader for making meaning.

2. *Viewing* identifies task content delivered through audio, video/film or other visual presentation that affords a receiver-oriented stance toward content acquisition.

3. *Teacher.* This construct identifies tasks in which content is delivered via teacher lecture, demonstration or PowerPoint, in which teacher has done the work of understanding and organizing material and delivers information to students. May involve some student interaction around presented material, but the primary focus is on teacher delivered content.

The *Opportunity to Learn: Task* dimension captures opportunities to learn from various disciplinary and literacy learning tasks and activities in observed lessons. Coding on this dimension is based on **what the tasks/activities asked students to know and do, rather than on what students themselves are doing**, which is coded under *Student Behavior: Learning*. The *Task* dimension consists of eight codes: *Close reading*, *Argumentation*, *Argumentation: Missed opportunity*, *Disciplinary knowledge building*, *Cross textual analysis*, *Fact acquisition*, *Other task*, and *Writing*. When tasks engage students in multiple learning opportunities, the task is coded to all relevant *Opportunity to learn: Task* categories. Definitions of the eight *Task* codes are presented below:

1. *OTL: Close reading:* This code identifies tasks that ask students to engage in interactive negotiation of meaning at the local and global levels to unearth and evaluate possible meanings, individually or collaboratively. Close reading is characterized by approaching

texts to understand vs. to find information, and reflects the basic understanding and attitude that reading means comprehending, interpreting, analyzing, and critiquing texts (Norris & Phillips, 2003).

2. *OTL: Argumentation* refers to tasks that ask students to make a claim or assertion that is supported by evidence that connects to the claim in a principled way. Argumentation tasks are framed as inquiry into multiple possibilities and/or viewpoints (i.e., asking students to find "evidence" to support a fact is not argumentation). Tasks may or may not be explicitly identified as "argumentation."

3. *OTL: Disciplinary knowledge building* identifies tasks characterized by approaching the discipline and disciplinary knowledge building through overarching frameworks, concepts and themes. Disciplinary knowledge building tasks often ask students to identify or apply disciplinary epistemologies, frameworks, concepts and themes to specific texts, cases, situations or contexts.

4. *OTL: Cross textual analysis* identifies tasks that ask students to synthesize, evaluate, or critique information from multiple texts (e.g., comparing/contrasting across evidence/representations to notice, reconcile agreements/ disagreements).

5. *OTL: Fact acquisition* refers to tasks characterized by testing understanding, recall, or rote learning with little or no opportunity for sense-making. *Fact acquisition* tasks may or may not involve brief instructional exchanges between teacher and students, but the overwhelming focus is on learning facts/information, recall, rote learning and right answers. Learning in these tasks may rely on an outside authoritative source, such as a textbook or teacher.

6. *OTL: Other task* is used for tasks that do not fit other Task categories.

7. *OTL: Writing.* This broad code includes both tasks that involve writing for knowledge building and for knowledge showing.

8. In addition, we coded *Argumentation: Missed opportunity* when a task framed as argumentation did not require students to generate claims supported by evidence related to the claim in a principled way, or when a task with promising elements of argumentation failed to have students generate claims supported by evidence related to the claim in a principled way.

The *Student Behavior: Learning* dimension is used for student learning based on ample visual/aural evidence. Codes for this dimension parallel *Opportunity to Learn* codes and code definitions. However, this dimension captures evidence of **what students themselves are doing, rather than on what the tasks/activities asked students to know and do**:

1. *Student learning: Close reading:* This code identifies engagement in interactive negotiation of meaning at the local and global levels to unearth and evaluate possible meanings, individually or collaboratively. *Student Behavior: Close reading* is characterized by evidence of students

approaching texts to understand vs. to find information, and reflects the basic understanding and attitude that reading means comprehending, interpreting, analyzing, and critiquing texts (Norris & Phillips, 2003).

2. *Student learning: Argumentation* refers to students making claims or assertions supported by evidence that connects to the claim in a principled way. Argumentation is framed as inquiry into multiple possibilities and/or viewpoints (i.e., finding "evidence" to support a fact is not argumentation). Student behavior may or may not be explicitly identified as "argumentation."

3. *Student learning: Disciplinary knowledge building* identifies segments where students approach the discipline and disciplinary knowledge building through overarching frameworks, concepts and themes. Disciplinary knowledge building often involves students identifying or applying disciplinary epistemologies, frameworks, concepts and themes to specific cases, situations or contexts.

4. *Student learning: Cross textual analysis* identifies segments where students synthesize, evaluate, or critique information from multiple texts (e.g., comparing/contrasting across evidence/representations to notice, reconcile agreements/ disagreements).

5. *Student learning: Fact acquisition* refers to segments in which students are involved in recall or rote learning with little or no opportunity for sense-making. Students may rely on an outside authoritative source, such as a textbook or teacher. *Fact acquisition* may or may not involve brief instructional exchanges between teacher and students, but the overwhelming focus is on learning facts/information, recall, rote learning and right answers.

6. *Student learning: Other task* is used for evidence of student engagement in tasks that do not fit on other *Task* categories.

7. *Student learning: Writing.* This broad code is used for student writing in the service of both knowledge building and for knowledge showing.

8. In addition to coding evidence of student engagement in *OTL: Tasks*, the *Student Learning* dimension includes a code for *Nascent argumentation*. The *Nascent argumentation* code is used for evidence of student engagement in nascent argumentation or reason-giving thinking as a byproduct of negotiating meaning.

*Student Behavior: Engagement* describes evidence of student engagement and effort in relation to reading, argumentation and disciplinary knowledge building, including persistence and grappling, student ownership, agency and extended instructionally focused conversation. This dimension comprises three codes:

1. *Student engagement: Unknown/NA* is used for segments characterized by lack of sufficient evidence to rate engagement.

2. *Student engagement: Low* identifies segments with ample evidence of low engagement and effort throughout.

3. *Student engagement: Moderate* is used for segments with ample evidence of variable and/or episodic intellectual engagement and effort among students, or where all students are moderately engaged throughout.

4. Segments are coded as Student *engagement: High* when there is ample evidence that most students are highly intellectually engaged most of the time.

**Unit of Analysis**

Video data can be coded in a variety of ways. While some researchers have used generic episodes as a "simple yet effective framework" for describing and analyzing lessons (Andrews, 2008, p. 5), analysis based on episodes focuses on the number and proportion of episodes in which a coded activity, behavior or interaction occurs, rather than on its duration. Because we were interested in the duration as well as the occurrence of student opportunities to learn, we coded the actual duration of each activity, behavior or interaction independently, rather than within the boundaries of an episode. The majority of our codes capture duration.

However, a few codes were designed primarily to flag the occurrence of an observed activity, interaction, or behavior.[2] "Non-duration" codes were used for activities, behaviors and interactions that are often intermittent, and where we anticipated that video footage would not afford accurate assessment of duration, such as teacher guidance and support or student engagement and learning during small group work. The analysis reported here focuses primarily on duration codes. When a non-duration code is included in a duration analysis, its value reflects the duration of segments that contain occurrences of the coded activity, such as the duration of small group work during which the teacher provided modeling, guidance and support.

*Grouping*, OTL: *Content Delivery* and OTL: *Task*, *Housekeeping/management*, and *Teacher gives instructions* are duration codes. *Teacher modeling, guidance and support*, and all *Student Engagement* and *Student Behavior* dimension codes are non-duration codes.

## V. Method

**Coding Protocol and Training**

To support researchers in coding accurately and reliably we developed a Coding Manual. The Video Segmentation Coding Manual is found in Appendix C.

---

[2] Unlike the broad opportunity to learn codes related to content delivery and tasks, capturing interactions, talk and behavior is limited by the scope and angle of the video. In some cases, the wide-angle lens did not capture the entire classroom and while it provided ample aural and visual evidence of some teacher and student activities and behaviors, it excluded others. Because our ability to accurately determine duration was limited by what was visible and audible on camera, *Student Behavior: Learning* codes and one *Teacher-Student Interaction* code, *Teacher modeling, guidance and support*, are used to flag occurrence of activities, interactions and behaviors within a segment, rather than for duration. For *Student Behavior: Engagement*, coders assigned ratings of 1-3 (low, medium, high engagement) as rough indicators of student engagement and/or changes in engagement during the lesson.

Videotapes from observed lessons were coded by four researchers, including two UIC graduate students with expertise in learning sciences and previous high school teaching experience, and two WestEd research and professional development staff. Training and support for coding was provided by the WestEd research coordinator and a WestEd research associate who was the principle architect of the coding schemes and protocols used for the video segmentation and earlier analysis of Year 1 classroom observation data. There were several potential challenges to achieving a shared understanding across sites and coders. One challenge was associated with the distributed nature of Project READI in the Midwest and on the West Coast. The second stemmed from the range of expertise and background experiences of the coders. In particular, WestEd coders were grounded in theoretical perspectives and instructional practices that shaped the coding scheme, while these were less familiar to the two Chicago coders.

To address these potential challenges, we held a series of four cross-site phone and video-conference trainings that employed a Frame of Reference model, based on practice, discussions between researchers, and feedback from trainers (Melchers, Lienhardt, Von Aarburg, & Kleinmann, 2011). Prior to each training session, researchers independently coded the same video, then shared and discussed reasons for assigning their respective coding, coding discrepancies and questions during the training. This process created a common evaluative standard and led to both a deeper understanding of the coding scheme on the part of all coders, and to fine-tuning the codes and coding protocol itself.

Once we established acceptable inter-rater reliability among the four coders, each video was assigned and coded by a single coder. To insure and maintain good inter-rater reliability, two or three additional videotapes were coded and discussed by intra- and cross-site pairs of coders throughout the coding phase.

**Inter-rater Reliability**

Kappa coefficients were computed using the NVivo coding comparison function to measure inter-rater reliability for the eight videos coded by multiple coders (Cohen, 1960). The kappa statistic is the most commonly used statistic for measuring agreement and disagreement between coders (Viera & Garrett, 2005). Kappa is reported in decimal form, from zero to one. A kappa of 1 indicates perfect agreement, whereas a kappa of 0 indicates agreement equivalent to chance. Negative kappas indicate lower than chance agreement. A limitation of kappa is that it is affected by the prevalence of the occurrence under observation. For rare incidents, behaviors, or interactions, very low values of kappa may not necessarily reflect low rates of overall agreement. Our extensive coding scheme included a large number of low-occurring but theoretically important codes such as nascent argumentation. Thus, to assist with interpretation of inter-rater reliabilities, we also calculated percentage agreement.

Finally, we calculated average kappas across the seven jointly coded videos. In order to provide a more accurate picture of interrater reliability, in calculating kappa averages we omitted kappas of 0 and 1. Kappa yields a coefficient of 0 when a code is used by one coder and not the other(s), regardless of how briefly. Similarly, in the case of low frequency codes that are not used by any coder, the kappa coefficient is 1 (indicating perfect agreement among coders that an incident, behavior or interactions is not present). Most coding schemes involve assigning a single code based on a small number of categories (Stein, Devore & Wojcik, N.d.). In contrast, for each observation, coders in this study could potentially select multiple responses from a large number of categories. Because our coding included frequent 1s (n=65) and 0s (n=27), we were concerned that including

these extreme values in our calculations would result in misleadingly inflated or reduced in kappa averages.

Average inter-rater reliability and percentage agreement for the video segmentation codes based on all seven videos coded by multiple coders are shown in Table 4.

Table 4. Inter-rater Reliability[3]

| Segmentation Code | Average Kappa[4] | Range Kappa | Average Agreement (%)[5] | Range Agreement (%) |
|---|---|---|---|---|
| **Content Delivery** | | | | |
| Teacher | 0.32 | -.05-.62 | 79 | 54-89 |
| Audio/Video[6] | -- | | -- | |
| Working with text | 0.44 | -.12-.94 | 77 | 49-97 |
| **Grouping** | | | | |
| Individual | 0.82 | .73-.99 | 97 | 94-97 |
| Pairs | 0.72 | .42-.88 | 93 | 87-96 |
| Small groups | 0.80 | .54-.90 | 95 | 86-98 |
| Whole class | 0.75 | .45-.94 | 90 | 79-98 |
| **Task** | | | | |
| Argumentation | 0.55 | .55[7] | 85 | 80-94 |
| Close reading | 0.73 | .63-.90 | 88 | 83-95 |
| Cross textual analysis | 0.84 | .84 | 94 | 90-97 |
| Disciplinary knowledge building | 0.34 | -.14-.77 | 76 | 43-89 |
| Fact acquisition | 0.63 | .52-.90 | 88 | 77-98 |
| Missed opportunity: Argumentation | 0.85 | .85 | 84 | 75-94 |
| Other task | 0.39 | -.16-.94 | 83 | 61-99 |
| Writing | 0.76 | .30-.96 | 90 | 68-99 |
| **Teacher-Student Interaction** | | | | |
| Housekeeping/management | 0.82 | .71-.96 | 97 | 95-99 |
| Teacher gives instructions | 0.64 | -.03-.95 | 95 | 87-99 |
| Teacher modeling, guidance, support | 0.60 | .25-.88 | 83 | 63-94 |

[3] Based on coding of seven lessons
[4]

| | | | | |
|---|---|---|---|---|
| **Student Learning** | | | | |
| Student: Argumentation | 0.56 | .56 | 88 | 83-94 |
| Student: Close reading | 0.79 | .67-.98 | 89 | 79-99 |
| Student: Cross textual analysis | -- | -- | 92 | 89-94 |
| Student: Disciplinary knowledge building | 0.25 | -.48-.84 | 67 | 29-92 |
| Student: Fact acquisition | 0.87 | .83-.90 | 88 | 77-98 |
| Student: Missed opportunity argumentation | 0.99 | .99 | 94 | 88 |
| Student: Nascent argumentation | 0.04 | -.10-.23 | 87 | 82-97 |
| Student: Other task | -- | -- | 95 | 91-98 |
| Student: Writing | 0.77 | .45-.99 | 92 | 80-97 |
| **Student Engagement** | | | | |
| Unknown/NA | 0.54 | .02-.84 | 85 | 57-93 |
| Low engagement | 0.11 | .11 | 78 | 74-82 |
| Moderate engagement | 0.48 | .24-.87 | 81 | 63-94 |
| High engagement | 0.53 | .46-.61 | 83 | 78-92 |

While kappa coefficients varied considerably by code, inter-rater reliabilities were acceptable for the majority of segmentation codes. Average kappas indicated almost perfect (0.81-0.99), substantial (0.61 to 0.80) or moderate (0.41-0.60) agreement for 24 of the 31 codes (see Landis & Koch, 1977). For seven codes, inter-rater reliability showed only fair (0.21-0.40) or slight (0.01-0.20) agreement among coders. Percentage agreement was acceptable (75 percent or above) for all but one code.

## IV. Results of Descriptive Analysis of Segmented Video Data

*Opportunity to learn* (OTL) refers to resources, curriculum, and teaching that promote learning. For the remainder of this report, we focus on results from the descriptive analysis of opportunities for students to engage in tasks and activities central to disciplinary learning and argumentation from multiple texts in multiple subject areas and at multiple grade levels, based on the low inference segmentation of video footage. The analysis was limited to OTL and did not include evidence of student learning[8]. In the future, video segmentation will help us target episodes for future transcription and fine grained analysis of classroom talk and interactions.

### Research Questions

Several specific inquiry questions guided this descriptive phase of the classroom observation analysis of opportunity to learn. Specifically our analysis explored:

### Content Delivery and Task Opportunity to Learn

A central question focused on opportunities students have to engage in tasks and activities central to disciplinary reading, learning and argumentation:

- *How frequently do students have opportunities to work with text, in contrast to learning content through teacher lecture or other means?*
- *What opportunities do students have to engage in tasks central to argumentation literacy and disciplinary learning?*
- *How are tasks related to one another?*
- *Which tasks are students asked to do with text, and how are tasks assigned in the context of working with text different from tasks performed when students learn content through teacher lecture or viewing?*

### Teacher-Student Interactions and Grouping

A wide range of research suggests that opportunities for collaborative meaning-making may mediate the development of students' knowledge, skills and dispositions to engage in advanced levels of academic literacy and evidence-based argumentation across disciplines. When individual reading and thinking processes are made public through classroom conversation, the resulting collaborative negotiation of meaning provides opportunities for the development of fundamental literacy, including *interactive argumentation* (Chinn & Anderson, 1998). We therefore examined teacher-student

[8] Unlike the broad OTL codes, evidence of interactions, talk and behavior is constrained by the scope and angle of the video. In some cases, the wide-angle lens was focused on a subset of the classroom and while it provided ample aural and visual evidence of some interactions and behaviors, it excluded others. Because our ability to accurately code teacher and student interactions and behavior was limited by what was visible on camera, all *Student Behavior: Learning* codes and one *Teacher-Student Interaction* code, *Teacher modeling, guidance and support*, flag occurrence of activities, interactions and behaviors within a segment, but do not indicate duration. These codes will be used primarily for subsequent in depth analysis of telling cases.

interaction and grouping structures in observed lessons as an important element of opportunity to learn:

- *How are different grouping structures allocated across observed lessons?*
- *Which grouping structures are associated with particular task and content delivery opportunities to learn?*

Likewise, we looked at broad categories of teacher-student interaction. Because teachers communicate understandings and expectations of a task or activity through tone of voice, word choice, interactions, routines, and explicit instructions and comments, the interpretation of which is beyond a first pass through the data, the research question driving the analysis of *Teacher-Student Interactions* was very general:

- *How much time is allocated to the different categories of teacher-student interaction—housekeeping/management, giving instructions, and teacher modeling, guidance and support?*

**Disciplinary, Site and Grade Level Differences in Opportunity to Learn**

Our analysis also embedded questions about disciplinary, site, and grade level differences related to opportunities for students to engage in tasks and activities central to disciplinary and argumentative literacy:

Reading, thinking and argumentation processes are domain-specific (Lee & Spratley, 2010; Moje, 2007; Reisman, 2011; Shanahan & Shanahan, 2008; Shanahan, Shanahan & Misischia, 2011). In addition, existing literature and our own initial analysis of the classroom observation data reveal differences in instruction across the core subject areas that may not be inherent to the discipline, but are nonetheless deeply embedded in the teaching of these disciplines. We therefore looked at disciplinary differences in opportunity to learn.

Our preliminary analysis suggested that classrooms where teachers participated in Reading Apprenticeship professional development focused on metacognitive conversation, collaboration and support for close reading differed from non-RA classrooms in a number of ways. Given the potentially important contrasts that emerged from our earlier analyses, we also analyzed the data by whether or not teachers participated in Reading Apprenticeship professional development.

Finally, we also looked at differences in opportunity to learn as a function of grade level. For this analysis, we created three grade bands consistent with the developmental bands used in the Common Core State Standards: middle school (grades 6-8), grades 9-10, and grades 11-12.

Below, we present the major findings related to opportunity to learn in three parts: 1. Content Delivery and Tasks, 2. Teacher-Student Interactions and Grouping, and 3. Developmental Differences. Disciplinary and site differences are considered within each of the three sections.

We preface these results by offering two caveats:

1. Findings from the classroom observation study should be viewed in light of the relatively small number of lessons. This is particularly true for interpreting discipline-, site- and grade-level-specific results. Furthermore, because we observed many classes more than once, the analysis is based on only 31 teachers and 44 classes.

2. In addition, our observations of discrete lessons offer a snapshot of teaching and learning at a moment in time. Thus, the video data provides an incomplete picture of instruction that unfolds and builds over time.

**Results**

To describe features of observed lessons, we calculated the number and percentage of observed lessons that included each coded activity, task or interaction. We also calculated duration in minutes for each activity, task or interaction. Because the number and length of observed lessons varied between sites, disciplines and grades, in order to make comparisons about how time was allocated across various activities, interactions, and behaviors in these subgroups, we calculated percentage duration using total duration for each descriptive category as the denominator. For example, for calculating percentage time allocated to argumentation tasks in history, we divided *duration of argumentation tasks* across history lessons by *total duration* of history lessons. (For non-duration codes, duration values reflect duration of segments that contained occurrences of the coded activity, interaction, or behavior.)

Classrooms are complex environments where multiple activities and interactions may occur simultaneously. We coded all activities, interactions and behaviors within a segment. Some segments included multiple codes from a single dimension (e.g., multiple tasks), and other segments did not include any code for a particular dimensions (e.g., housekeeping segments generally did not include task codes). Therefore, percentage time within a dimension does not necessarily add up to 100 percent.

**Content Delivery and Task Opportunity to Learn**

To answer the question, *How frequently do students have opportunities to work with text, in contrast to learning content through teacher lecture or other means?***,** we calculated both *occurrence*—the number and percentage of observed lessons in which each content delivery method occurred—and *duration*—the percentage of observed time that content was presented through each of the three Content Delivery codes—teacher lecture/demonstration/explanation, media (audio or visual), and working with text. We conducted a parallel analysis to answer the question, *What opportunities do students have to engage in tasks and activities central to disciplinary learning and argumentation?,* calculating the number and percentage of observed lessons in which each *Opportunity to Learn: Task* code occurred, as well as the percentage of time allocated to each task across lessons. Because extant literature reveals differences in teaching across the content areas, we also examined Content Delivery and Task codes by discipline.

Tables 5 and 6 present descriptive statistics for occurrence and duration of Content Delivery and Task opportunities to learn for the entire corpus of lessons and by discipline.

Occurrence. For Content Delivery codes, as shown in Table 5, the majority of observed lessons included content delivered both by the teacher and by working with text. Teachers delivered content in over two thirds of all lesson, with teachers delivering content more often in science than in history or literature. Students worked with texts in nearly every observed lesson regardless of discipline. This was not surprising since we had asked to observe typical lessons "in which reading plays a central role." Occurrence of media-delivered content was rare across disciplines.

With respect to Task codes, tasks related to argumentation literacy—argumentation, close reading and cross-textual analysis—showed the greatest variation across disciplines. Elements of argumentation occurred most frequently in history. Half of all history lessons included evidence-based argumentation and close reading tasks, and a third of all history lessons included cross-textual analysis. In contrast, no science lesson incorporated any of these building blocks of argumentation. Literature lessons had the highest occurrence of close reading, but argumentation and cross-textual analysis tasks occurred less frequently in literature than in history. Over 20 percent of literature lessons included missed opportunities for argumentation—tasks with promising elements of argumentation, albeit without requiring students to offer evidence to support their claims. Seventeen percent of history and 15 percent of science lessons likewise included missed opportunities for argumentation.

Disciplinary knowledge-building, fact acquisition and writing tasks were common across disciplines. Disciplinary knowledge building and fact acquisition occurred most frequently in science lessons, with disciplinary knowledge-building tasks observed in every science lesson, and fact acquisition in nearly 70 percent of science lessons.

Duration. We also examined how much time was allocated to each activity. Duration of content delivery activities showed greater disciplinary variation than occurrence *per se*. For example, , while the vast majority of both history (79 percent) and science (92 percent) lessons included teacher-delivered content, nearly twice as much time was allocated to teacher delivery of content in science (36 percent of class time) than in history (20 percent of class time). Similarly, students were asked to work with text in over 90 percent of lessons regardless of discipline. However history lessons involved working with text 68 percent of the time, literature lessons involved working with text 65 percent of the time, and science only 55 percent of the time.

For task types, duration mirrored findings related to occurrence. As in the case of occurrence, tasks related to argumentation literacy showed the greatest variation between disciplines. History lessons allocated 26 percent of class time to argumentation and 14 percent of class time to cross-textual analysis tasks, compared with 11 percent and 6 percent in literature. Both history and literature teachers allocated a third of class time to close reading. As discussed earlier, no science lesson included any argumentation, close reading or cross-textual analysis task—even briefly. Mirroring findings related to occurrence, science lessons were focused on content, allocating the greatest percentage of time to disciplinary knowledge-building and fact acquisition tasks. Nearly two-thirds of all science lessons were spent on disciplinary knowledge-building tasks involving scientific concepts, frameworks or themes. Nearly a quarter of class time in science was spent on fact acquisition.

Table 5: Content Delivery and Task: Occurrence by Discipline

*Table 5 shows the number and percentage of observed lessons in which each content delivery and task type occurred.*

| | All Lessons N=71 | | History N=24 | | Literature N=34 | | Science N=13 | |
|---|---|---|---|---|---|---|---|---|
| | Total number and percentage of lessons in which activity/task occurred | | | | | | | |
| | N | % | N | % | N | % | N | % |
| **Content Delivery** | | | | | | | | |
| Teacher | 54 | 76 | 19 | 79 | 23 | 68 | 12 | 92 |
| Viewing/Listening | 7 | 10 | 3 | 13 | 3 | 9 | 1 | 8 |
| Working with text | 66 | 93 | 23 | 96 | 31 | 91 | 12 | 92 |
| **Task** | | | | | | | | |
| Argumentation | 21 | 30 | 12 | 50 | 9 | 26 | 0 | 0 |
| Close reading | 36 | 51 | 13 | 54 | 23 | 68 | 0 | 0 |
| Cross-textual analysis | 14 | 20 | 8 | 33 | 6 | 18 | 0 | 0 |
| Disciplinary knowledge-building | 61 | 86 | 20 | 83 | 28 | 82 | 13 | 100 |
| Fact acquisition | 39 | 55 | 14 | 58 | 16 | 47 | 9 | 69 |
| Missed opportunity argumentation | 13 | 18 | 4 | 17 | 7 | 21 | 2 | 15 |
| Other task | 30 | 42 | 9 | 38 | 16 | 47 | 5 | 38 |
| Writing | 66 | 93 | 24 | 100 | 30 | 88 | 12 | 92 |

Table 6: Content Delivery and Task: Duration by Discipline

*Table 6 shows the total duration (in minutes) and percentage of total time allocated to each content delivery and task type across observed lessons.*

| | All Lessons (N=71) 3699 minutes | | History (N=24) 1265 minutes | | Literature (N=34) 1820 minutes | | Science (N=13) 615 minutes | |
|---|---|---|---|---|---|---|---|---|
| | Total duration[9] (in minutes) and percentage of time allocated to activity/task across lessons | | | | | | | |
| | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| **Content Delivery** | | | | | | | | |
| Teacher lecture, demonstration, explanation | 774 | 21 | 249 | 20 | 306 | 17 | 219 | 36 |
| Viewing/Listening | 38 | 1 | 16 | 1 | 21 | 1 | 1 | 0 |
| Working with text | 2376 | 64 | 857 | 68 | 1184 | 65 | 336 | 55 |
| **Task** | | | | | | | | |
| Argumentation | 519 | 14 | 325 | 26 | 193 | 11 | 0 | 0 |
| Close reading | 1050 | 28 | 407 | 32 | 643 | 35 | 0 | 0 |
| Cross-textual analysis | 284 | 8 | 174 | 14 | 110 | 6 | 0 | 0 |
| Disciplinary knowledge-building | 2041 | 55 | 699 | 55 | 934 | 51 | 408 | 66 |
| Fact acquisition | 599 | 16 | 205 | 16 | 250 | 14 | 145 | 24 |
| Missed opportunity argumentation | 102 | 3 | 26 | 2 | 66 | 4 | 10 | 2 |
| Other task | 205 | 6 | 58 | 5 | 116 | 6 | 32 | 5 |
| Writing | 1809 | 49 | 736 | 58 | 751 | 41 | 322 | 52 |

[9] Duration and percentage of time across categories of Content Delivery do not add up to 100 percent of total lesson time: students sometimes learned lesson content in ways other than teacher lecture, audio/video materials, or text—e.g., through labs or simulations; time allocated to housekeeping or giving directions did not involve any content delivery; and students occasionally learned content in multiple ways—e.g., working with text while listening to an audio-recording of a play. Likewise, because students were often asked to perform multiple tasks during a single segment, percentage time across Task categories does not add up to 100 percent.

Teacher recruitment and selection resulted in a posttest-only quasi-experimental design with nonequivalent groups (Cook and Campbell, 1979) with potentially different practices related to teaching and learning of argumentation literacy—a "treatment group" of experienced Reading Apprenticeship teachers and a "comparison group" of teachers who had not participated in Reading Apprenticeship professional development but whose instruction was reported to foster disciplinary literacies in history, science, and literature. While RA and non-RA group membership overlapped significantly with site, the non-RA group included three science teachers from the San Francisco Bay Area who had not participated in Reading Apprenticeship professional development. Given the potentially important contrasts between RA and non-RA teachers, we also analyzed occurrence and duration of Content Delivery and Task codes separately for RA and non-RA groups. Because the RA group included only one science lesson, the analysis of "treatment group"[10] differences by discipline includes only history and literature lessons. These results are shown in Tables 7 and 8.

As in the analysis of disciplinary differences in content delivery, the occurrence of the three content delivery methods was generally similar across RA and Non-RA groups. However, treatment differences strongly mediated disciplinary differences in the occurrence of several Task codes. As shown in Table 7, this was particularly apparent for argumentation, close reading and cross-textual analysis. Compared with Non-RA lessons, a higher percentage of RA lessons provided students with opportunities to engage in argumentation, close reading and cross textual analysis. Treatment group effects for argumentation were strongest in literature. Across treatment and comparison groups, 26 percent of literature lessons included argumentation tasks. However, argumentation occurred in 44 percent of RA literature lessons and only 11 percent of Non-RA literature lessons. Indeed, while argumentation occurred nearly twice as often in history (50 percent) as in literature (26 percent) across treatment groups, argumentation rates were nearly identical in RA literature (44 percent) and non-RA history (45 percent) lessons. Treatment group differences related to close reading were equally dramatic. Across groups, 54 percent of all history lessons incorporated close reading; however, close reading occurred in 77 percent of RA history and 27 percent of non-RA history lessons. Similarly, whereas close reading occurred in 68 percent of literature lessons across groups, close reading occurred in 81 percent of RA literature lessons and 56 percent of non-RA literature lessons. Finally, missed opportunity for argumentation occurred more frequently in non-RA lessons in both history and literature.

[10] The term "treatment group" is used loosely here, in the absence of a more appropriate phrase to describe the relationship between the two groups of teachers.

Table 7: Content Delivery and Task: Occurrence by Discipline and Treatment Group

*Table 7 shows the number and percentage of observed lessons in which each content delivery and task type occurred.*

| | **All Lessons (N=71)** 3699 minutes | | | | **History (N=24)** 1265 minutes | | | | **Literature (N=34)** 1820 minutes | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RA (N=30) 1838 min | | Non-RA (N=41) 1862 min | | RA (N=13) 786 min | | Non-RA (N=11) 478 min | | RA (N=16) 997 min | | Non-RA (N=18) 823 min | |
| | Total number and percentage of lessons in which activity/task occurred | | | | | | | | | | | |
| **Content Delivery** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| Teacher lecture, demonstration, explanation | 22 | 73 | 32 | 78 | 11 | 85 | 8 | 73 | 10 | 63 | 13 | 72 |
| Viewing/Listening | 3 | 10 | 4 | 10 | 1 | 8 | 2 | 18 | 2 | 13 | 1 | 6 |
| Working with text | 29 | 97 | 37 | 90 | 12 | 92 | 11 | 100 | 16 | 100 | 15 | 83 |
| **Task** | | | | | | | | | | | | |
| Argumentation | 14 | 47 | 7 | 17 | 7 | 54 | 5 | 45 | 7 | 44 | 2 | 11 |
| Close reading | 23 | 77 | 13 | 32 | 10 | 77 | 3 | 27 | 13 | 81 | 10 | 56 |
| Cross-textual analysis | 10 | 33 | 4 | 10 | 5 | 38 | 3 | 27 | 5 | 31 | 1 | 6 |
| Disciplinary knowledge-building | 27 | 90 | 34 | 83 | 11 | 85 | 9 | 82 | 15 | 94 | 13 | 72 |
| Fact acquisition | 12 | 40 | 27 | 66 | 7 | 54 | 7 | 64 | 4 | 25 | 12 | 67 |
| Missed opportunity argumentation | 3 | 10 | 10 | 24 | 1 | 8 | 3 | 27 | 2 | 13 | 5 | 28 |
| Other task | 11 | 37 | 19 | 46 | 4 | 31 | 5 | 45 | 6 | 38 | 10 | 56 |
| Writing | 28 | 93 | 38 | 93 | 13 | 100 | 11 | 100 | 14 | 88 | 16 | 89 |

Table 8: Content Delivery and Task: Duration by Discipline and Treatment Group

*Table 8 shows the total duration (in minutes) and percentage of total time allocated to each content delivery and task type across observed lessons.*

| | All Lessons (N=71) 3699 minutes | | | | History (N=24) 1265 minutes | | | | Literature (N=34) 1820 minutes | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RA (N=30) 1838 min | | Non-RA (N=41) 1862 min | | RA (N=13) 786 min | | Non-RA (N=11) 478 min | | RA (N=16) 997 min | | Non-RA (N=18) 823 min | |
| | Total duration[11] (in minutes) and percentage of time allocated to activity/task across lessons | | | | | | | | | | | |
| **Content Delivery** | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| Teacher lecture, demonstration, explanation | 173 | 9 | 602 | 32 | 99 | 13 | 150 | 31 | 63 | 6 | 243 | 30 |
| Viewing/Listening | 21 | 1 | 17 | 1 | 7 | 1 | 9 | 2 | 14 | 1 | 7 | 1 |
| Working with text | 1387 | 75 | 989 | 53 | 619 | 79 | 238 | 50 | 744 | 75 | 439 | 53 |
| **Task** | | | | | | | | | | | | |
| Argumentation | 346 | 19 | 173 | 9 | 214 | 27 | 111 | 23 | 132 | 13 | 62 | 7 |
| Close reading | 850 | 46 | 200 | 11 | 387 | 49 | 20 | 4 | 463 | 46 | 180 | 22 |
| Cross-textual analysis | 205 | 11 | 78 | 4 | 99 | 13 | 75 | 16 | 107 | 11 | 3 | 0 |
| Disciplinary knowledge-building | 998 | 54 | 1044 | 56 | 428 | 54 | 271 | 57 | 557 | 56 | 378 | 46 |
| Fact acquisition | 122 | 7 | 478 | 26 | 79 | 10 | 126 | 26 | 20 | 2 | 230 | 28 |
| Missed opportunity argumentation | 8 | 0 | 94 | 5 | 4 | 1 | 22 | 5 | 4 | 0 | 62 | 8 |
| Other task | 75 | 4 | 131 | 7 | 34 | 4 | 24 | 5 | 37 | 4 | 79 | 10 |
| Writing | 910 | 50 | 899 | 48 | 429 | 55 | 307 | 64 | 447 | 45 | 304 | 37 |

[11] Duration and percentage of time across categories of Content Delivery do not add up to 100 percent of total lesson time: students sometimes learned lesson content in ways other than teacher lecture, audio/video materials, or text—e.g., through labs or simulations; time allocated to housekeeping or giving directions did not involve any content delivery; and students occasionally learned content in multiple ways—e.g., working with text while listening to an audio-recording of a play. Likewise, because students were often asked to perform multiple tasks during a single segment, percentage time across Task categories does not add up to 100 percent.

While similar percentages of RA and Non-RA lessons included content delivered by teachers, media and through working with text, as shown in Table 8, we found strong treatment group differences in the *percentage of time* allocated to various content delivery methods. Specifically, Reading Apprenticeship classrooms allocated more time to working with text than Non-RA classrooms, and Non-RA classrooms allocated more time to teacher lecture, demonstration and explanation than RA classrooms. We also found sizable treatment group differences in the percentage of time allocated to tasks central to argumentative literacy, as defined by Project READI. Across disciplines, students in RA classrooms had two times the opportunity to engage in argumentation, four times the opportunity to engage in close reading and three times the opportunity to engage in cross textual analysis as students in Non-RA classrooms. In history, where teachers in treatment and comparison groups allocated comparable amounts of time to evidence-based argumentation and cross-textual analysis, there were dramatic differences in time allocated to working with text and close reading. Reading Apprenticeship history lessons allocated 79 percent of class time to working with text and 49 percent of class time to close reading, compared with Non-RA lessons where students were worked with text 50 percent of the time and engaged in close reading only 4 percent of the time. We also found treatment group differences in literature lessons. Literature students in RA classrooms spent twice as much time on both close reading and evidence-based argumentation compared with their Non-RA counterparts. Furthermore, researchers coded 8 percent of the time in Non-RA literature lessons as missed opportunity for argumentation, indicating tasks and interactions with unfulfilled potential for fostering evidence-based argumentation, while no RA literature lesson was coded for missed opportunity for argumentation. Finally, while teachers in RA literature classrooms asked students to engage in cross-textual analysis tasks over 10 percent of the time (107 minutes), teachers in Non-RA classrooms allocated a total of three minutes to cross-textual analysis across all 18 Non-RA lessons.

Non-RA lessons involved considerably more fact acquisition than RA lessons in both history and literature lessons.

To answer the question ***Which tasks are students asked to do with text, and how are tasks performed in the context of working with text different from tasks performed when students learn content through teacher lecture or viewing?,*** we examined relationships between Content Delivery and Task codes.

Tables 9 and 10 present disciplinary differences in co-occurrence and duration of task opportunities to learn when content was delivered through teacher lecture/demonstration/explanation, media (viewing/listening), and working with text.

For both occurrence and duration, content delivery was highly related to task. In general, relationships were stronger for duration—the percentage of time allocated to tasks as a function of content delivery method—than for occurrence, *per se*. Across observed lessons, argumentation, close reading, and cross textual analysis co-occurred with greater frequency and for a greater percentage of the time when students were working with text than when content was delivered through teacher lecture/demonstration/explanation, or multimedia. These building blocks of argumentation were rare when teachers delivered lesson content. In contrast, disciplinary knowledge-building and writing tasks were common across all categories of content delivery. In science, disciplinary knowledge-building co-occurred a high percentage of the time whether content was delivered by the teacher (73 percent), multimedia (100 percent) or through working with text (83 percent). In history and

literature, disciplinary knowledge-building tasks occurred for a greater percentage of time when content was delivered by working with text (67 and 61 percent, respectively) than when content was delivered by the teacher (31 and 39 percent) or through multimedia (33 and 24 percent). In the case of history, more than twice as much time was allocated to disciplinary knowledge-building when content was delivered by text than through the other two methods.

In general, fact acquisition tasks characterized by recall or rote learning with little or no opportunity for sense-making occurred a high percentage of the time when content was delivered by teachers. However, in science, fact acquisition occurred most frequently when students worked with text, and in history, high levels of fact acquisition were observed when content was delivered by multimedia.

Table 9: Co-Occurrence of Content Delivery and Task: Occurrence by Discipline

*Table 9 shows the number and percentage of observed lessons in which students were asked to perform various tasks when content was delivered through teacher lecture/demonstration/explanation, media (viewing/listening), or working with text.*

| | **Teacher lecture, demonstration, explanation** | | | | | | | | **Media (Viewing/Listening)** | | | | | | | | **Working with Text** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total N=54 (76%) | | History N=19 (79%) | | Literature N=23 (68%) | | Science N=12 (92%) | | Total N=7 (10%) | | History N=3 (13%) | | Literature N=3 (9%) | | Science N=1 (8%) | | Total N=66 (93%) | | History N=23 (96%) | | Literature N=31 (91%) | | Science N=12 (92%) | |
| | Number and percentage[12] of lessons in which students were asked to perform task *given a particular content delivery method* | | | | | | | | | | | | | | | | | | | | | | | |
| **Task** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| Argumentation | 5 | 9 | 4 | 21 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 32 | 12 | 52 | 9 | 29 | 0 | 0 |
| Close reading | 7 | 13 | 1 | 5 | 6 | 26 | 0 | 0 | 1 | 14 | 0 | 0 | 1 | 33 | 0 | 0 | 36 | 55 | 13 | 57 | 23 | 74 | 0 | 0 |
| Cross-textual analysis | 3 | 6 | 2 | 11 | 1 | 4 | 0 | 0 | 1 | 14 | 0 | 0 | 1 | 33 | 0 | 0 | 13 | 20 | 8 | 35 | 5 | 16 | 0 | 0 |
| Disciplinary knowledge-building | 36 | 67 | 11 | 58 | 14 | 61 | 11 | 92 | 3 | 43 | 1 | 33 | 1 | 33 | 1 | 100 | 55 | 83 | 19 | 83 | 25 | 81 | 11 | 92 |
| Fact acquisition | 26 | 48 | 10 | 53 | 9 | 39 | 7 | 58 | 2 | 29 | 2 | 67 | 0 | 0 | 0 | 0 | 25 | 38 | 9 | 39 | 8 | 26 | 8 | 67 |
| Missed opportunity argumentation | 8 | 15 | 3 | 16 | 3 | 13 | 2 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 12 | 2 | 9 | 5 | 16 | 1 | 8 |
| Other task | 14 | 26 | 5 | 26 | 6 | 26 | 3 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 21 | 4 | 17 | 7 | 23 | 3 | 25 |
| Writing | 23 | 43 | 9 | 47 | 8 | 35 | 6 | 50 | 3 | 43 | 1 | 33 | 2 | 67 | 0 | 0 | 58 | 88 | 21 | 91 | 26 | 84 | 11 | 92 |

Table 10: Co-Occurrence of Content Delivery and Task: Duration by Discipline

*Table 10 shows the percentage of time students in each of the three disciplines were asked to perform various tasks when content was delivered through teacher lecture/demonstration/explanation, viewing, or working with text.*

| | Teacher lecture, demonstration, explanation | | | | Media (Viewing/Listening) | | | | Working with Text | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total 774 min | History 249 min | Lit 306 min | Science 219 min | Total 38 min | History 16 min | Lit 21 min | Science 1 min | Total 2376 min | History 857 min | Lit 1184 min | Science 336 min |
| **Task** | Percentage of time[13] allocated to task *given a particular content delivery method* | | | | | | | | | | | |
| Argumentation | 3 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 20 | 35 | 14 | 0 |
| Close reading | 3 | 2 | 6 | 0 | 24 | 0 | 44 | 0 | 41 | 47 | 49 | 0 |
| Cross-textual analysis | 2 | 4 | 1 | 0 | 13 | 0 | 24 | 0 | 11 | 20 | 8 | 0 |
| Disciplinary knowledge-building | 49 | 31 | 39 | 83 | 30 | 33 | 24 | 100 | 65 | 67 | 61 | 73 |
| Fact acquisition | 45 | 48 | 51 | 31 | 29 | 67 | 0 | 0 | 11 | 7 | 6 | 36 |
| Missed opportunity argumentation | 5 | 7 | 5 | 2 | 0 | 0 | 0 | 0 | 2 | 1 | 4 | 1 |
| Other task | 7 | 11 | 6 | 4 | 0 | 0 | 0 | 0 | 4 | 1 | 6 | 6 |
| Writing | 37 | 41 | 26 | 47 | 59 | 45 | 74 | 0 | 56 | 65 | 47 | 63 |

[13] Duration and percentage of time across categories of Content Delivery do not add up to 100 percent of total lesson time: students sometimes learned lesson content in ways other than teacher lecture, audio/video materials, or text—e.g., through labs or simulations; time allocated to housekeeping or giving directions did not involve any content delivery; and students occasionally learned content in multiple ways—e.g., working with text while listening to an audio-recording of a play. Likewise, because students were often asked to perform multiple tasks during a single segment, percentage time across Task categories does not add up to 100 percent.

Tables 11 and 12 present occurrence and duration of task opportunities to learn for the three content delivery methods by treatment group. In both the RA and Non-RA classrooms, relationships between content delivery and tasks mirror findings for the whole sample. Argumentation, close reading, and cross textual analysis generally occurred with greater frequently and for a greater percentage of the time when students were working with text than when content was delivered through teacher lecture/demonstration/explanation, or multimedia, and fact acquisition tasks occurring more frequently and for greater duration when content was delivered by the teacher than when students were asked to work with text. However, the magnitude of these relationships differed by treatment group, especially for argumentation literacy and fact acquisition tasks. In general, argumentation, close reading and cross-textual analysis tasks occurred more frequently and for a greater percentage of the time in RA lessons compared with Non-RA lessons whether content was delivered by teacher lecture, demonstration and explanation, media (viewing/listening), or through working with text. In contrast, compared with RA lessons, missed opportunity for argumentation and fact acquisition occurred more frequently and for a greater percentage of the time in Non-RA lessons across all methods of content delivery. In addition, tasks associated with teacher lecture, demonstration and explanation differed by treatment group. In Non-RA lessons, teacher delivery of content was associated with high levels and duration of fact acquisition and writing (presumably notetaking), while fact acquisition and writing occurred at lower rates and/or for shorter durations when teachers delivered content in the RA classrooms. Finally, viewing and listening appeared to assume somewhat different roles in RA and Non-RA classrooms. While media-delivered content was associated with disciplinary knowledge-building and fact acquisition in both groups, in RA lessons, multimedia was also associated with close reading and cross-textual analysis. Students in Non-RA lessons were never asked to do close reading or cross-textual analysis tasks in the context of viewing and listening. It appears that students in Non-RA classrooms were less frequently assigned any task while watching or listening, while students who watched and listened in RA classrooms frequently engaged in literacy and content acquisition tasks.

Table 11: Co-Occurrence of Content Delivery and Task: Occurrence by Treatment Group

*Table 11 shows the number and percentage of observed lessons in which students were asked to perform various tasks when content was delivered through teacher lecture/demonstration/explanation, media (viewing/listening), or working with text.*

| | **Content Delivery** | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Teacher lecture, demonstration, explanation** N=54 (76%) | | | | | | **Media (Viewing/Listening)** N=7 (10%) | | | | | | **Working with Text** N=66 (92%) | | | | | |
| | All Lessons (N=54, 76%) | | RA (N=22, 73%) | | Non-RA (N=32, 78%) | | All Lessons (N=7, 10%) | | RA (N=3, 10%) | | Non-RA (N=4, 10%) | | All Lessons (N=66, 86%) | | RA (N=29, 97%) | | Non-RA (N=37, 90%) | |
| | Number and percentage[14] of lessons in which students were assigned each task *given a particular content delivery method* | | | | | | | | | | | | | | | | | |
| **Task** | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % |
| Argumentation | 5 | 9 | 3 | 14 | 2 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 32 | 14 | 48 | 7 | 19 |
| Close reading | 7 | 13 | 6 | 27 | 1 | 3 | 1 | 14 | 1 | 33 | 0 | 0 | 36 | 55 | 23 | 79 | 13 | 35 |
| Cross-textual analysis | 3 | 6 | 2 | 9 | 1 | 3 | 1 | 14 | 1 | 33 | 0 | 0 | 13 | 20 | 10 | 34 | 3 | 8 |
| Disciplinary knowledge-building | 36 | 67 | 15 | 68 | 21 | 66 | 3 | 43 | 1 | 33 | 2 | 50 | 55 | 83 | 25 | 86 | 30 | 81 |
| Fact acquisition | 26 | 48 | 6 | 27 | 20 | 63 | 2 | 29 | 1 | 33 | 1 | 25 | 25 | 38 | 9 | 31 | 16 | 43 |
| Missed opportunity argumentation | 8 | 15 | 0 | 0 | 8 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 12 | 2 | 7 | 6 | 16 |
| Other task | 14 | 26 | 6 | 27 | 8 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 21 | 4 | 14 | 10 | 27 |
| Writing | 23 | 43 | 9 | 41 | 14 | 44 | 3 | 43 | 2 | 67 | 1 | 25 | 58 | 88 | 26 | 90 | 32 | 86 |

Table 12: Co-Occurrence of Content Delivery and Task: Duration by Treatment Group

*Table 12 shows the percentage of time students were asked to perform various tasks when content was delivered through teacher lecture/demonstration/explanation, media, or working with text.*

| | Content Delivery | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Teacher lecture, demonstration, explanation | | | Media (Viewing/Listening) | | | Working with Text | | |
| | Total N=54 774 min | RA N=22 172 min | Non-RA N=32 602min | Total N=7 21 min | RA N=3 23 min | Non-RA N=4 1003min | Total N=66 2374 min | RA N=29 1385 min | Non-RA N=37 989 min |
| **Task** | Percentage of time[15] students were assigned each task *given a particular content delivery method* | | | | | | | | |
| Argumentation | 3 | 6 | 2 | 0 | 0 | 0 | 20 | 22 | 16 |
| Close reading | 3 | 11 | 1 | 24 | 43 | 0 | 41 | 58 | 19 |
| Cross-textual analysis | 2 | 6 | 0 | 13 | 23 | 0 | 11 | 14 | 7 |
| Disciplinary knowledge-building | 49 | 50 | 49 | 30 | 23 | 39 | 65 | 62 | 70 |
| Fact acquisition | 45 | 19 | 52 | 29 | 34 | 22 | 11 | 4 | 19 |
| Missed opportunity argumentation | 5 | 0 | 6 | 0 | 0 | 0 | 2 | 0 | 5 |
| Other task | 7 | 17 | 4 | 0 | 0 | 0 | 4 | 1 | 8 |
| Writing | 37 | 20 | 42 | 59 | 75 | 39 | 56 | 56 | 56 |

[15] Duration and percentage of time across categories of Content Delivery do not add up to 100 percent of total lesson time: students sometimes learned lesson content in ways other than teacher lecture, audio/video materials, or text—e.g., through labs or simulations; time allocated to housekeeping and giving directions did not involve any content delivery; and students occasionally learned content in multiple ways—e.g., working with text while listening to an audio-recording of a play. Likewise, because students were often asked to perform multiple tasks during a single segment, percentage time across Task categories does not add up to 100 percent.

Finally, to answer the question ***How are tasks related to one another?,*** we also looked at relationships among tasks. Specifically, our analysis explored overlap among tasks for video segments where students were assigned multiple tasks. Occurrence and percentage of time for task co-occurrence were calculated for each of the eight tasks. Results of this analysis are found in Tables 13 and 14. Because some tasks occurred more frequently and for greater duration than others, relationships are asymmetrical. For example, argumentation was relatively rare, with only 519 minutes across observed lessons allocated to argumentation tasks. Writing, on the other hand, was a relatively common task, with 1792 minutes allocated to writing activities. Consequently, the 269 minute overlap between argumentation and writing represented 52 percent of the 519 minutes allocated to argumentation tasks, but only 15 percent of the 1792 minutes allocated to writing.

As indicated both by lesson number and percentage of time that tasks co-occurred, the building blocks of argumentation literacy—argumentation, close reading, and cross textual analysis—kept company with one another, and with disciplinary knowledge-building and writing. In contrast, these argumentative literacy tasks rarely occurred in the context of fact acquisition. Cross-textual analysis occurred most frequently with argumentation, close reading, disciplinary knowledge building, and writing, although it was only infrequently part of these tasks. While only a third of argumentation tasks involved synthesizing across multiple texts, cross textual analysis occurred more than twice as often—and more than double the percentage of time—in the context of argumentation than with any other task. The frequent co-occurrence of these two tasks is especially noteworthy since both argumentation and cross-textual analysis were observed infrequently (519 and 284 minutes, respectively). Researchers coded missed opportunity for argumentation most frequently in the context of disciplinary knowledge-building and fact acquisition, indicating that tasks focused on content sometimes contained promising elements of argumentation that stopped short of asking students to generate claims or to support claims with evidence. Writing was ubiquitous across tasks. This is not surprising since the code includes writing in all forms for the purpose of both knowledge building and knowledge showing.

Table 13: Co-Occurrence of Tasks: Occurrence

| | Task Co-Occurrence | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Argumentation N=21 | Close reading N=36 | Cross-textual analysis N=14 | Disciplinary knowledge-building N=61 | Fact acquisition N=39 | Missed opportunity argumentation N=13 | Other task N=30 | Writing N=66 |
| **Task** | Number of lessons in which tasks co-occurred | | | | | | | |
| Argumentation N=21 | 21 | 8 | 7 | 14 | 0 | 0 | 0 | 14 |
| Close reading N=36 | 8 | 36 | 4 | 23 | 4 | 2 | 2 | 29 |
| Cross-textual analysis N=14 | 7 | 4 | 14 | 13 | 1 | 0 | 0 | 10 |
| Disciplinary knowledge- building N=61 | 14 | 23 | 13 | 61 | 18 | 6 | 7 | 49 |
| Fact acquisition N=39 | 0 | 4 | 1 | 18 | 39 | 2 | 2 | 19 |
| Missed opportunity argumentation N=13 | 0 | 2 | 0 | 6 | 2 | 13 | 1 | 3 |
| Other task N=30 | 0 | 2 | 0 | 7 | 2 | 1 | 30 | 6 |
| Writing N=66 | 14 | 29 | 10 | 49 | 19 | 3 | 6 | 66 |

Table 14: Co-Occurrence of Tasks: Percentage of Time [16]

| | Percentage Time Task Co-Occurrence | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Argumentation 519 min | Close reading 1050 min | Cross-textual analysis 284 min | Disciplinary knowledge 2041 min | Fact acquisition 599 min | Missed opp argumentation 102 min | Other task 205 min | Writing 1809 min |
| **Task OTL** | Percentage of time that tasks co-occurred | | | | | | | |
| Argumentation 519 min | 100 | 29 | 28 | 67 | 0 | 0 | 0 | 52 |
| Close reading 1050 min | 14 | 100 | 8 | 51 | 4 | 2 | 1 | 65 |
| Cross-textual analysis 284 min | 52 | 31 | 100 | 86 | 3 | 0 | 0 | 67 |
| Disciplinary knowledge- building 2041 min | 17 | 26 | 12 | 100 | 9 | 3 | 2 | 55 |
| Fact acquisition 599 min | 0 | 7 | 2 | 32 | 100 | 4 | 2 | 49 |
| Missed opportunity argumentation 102 min | 0 | 17 | 0 | 51 | 21 | 100 | 4 | 29 |
| Other task 205 min | 0 | 3 | 0 | 16 | 5 | 2 | 100 | 11 |
| Writing 1809 min | 15 | 38 | 11 | 62 | 16 | 2 | 1 | 100 |

[16] Read across the table—e.g., 29 percent of argumentation tasks involved close reading, 28 percent of argumentation tasks involved cross-textual analysis, 67 percent of argumentation tasks involved disciplinary knowledge-building, etc. Because some tasks occurred more frequently and for greater duration than others, relationships are asymmetrical. For example, argumentation was relatively rare, with only 519 minutes across observed lessons allocated to argumentation tasks. Writing, on the other hand, was a relatively common task, with 1792 minutes allocated to writing activities. Consequently, the 269 minute overlap between argumentation and writing represented 52 percent of the 519 minutes allocated to argumentation tasks, but only 15 percent of the 1792 minutes allocated to writing.

## Social Support from Teachers and Peers

### Teacher-Student Interaction

To answer the question, ***How much time is allocated to the different categories of teacher-student interaction?,*** we calculated the number and percentage of observed lessons and the percentage of time allocated to each Teacher-Student Interaction code. Results from this analysis are found in Tables 15 through 18.

We defined *Housekeeping/management* as segments of at least 30 seconds with a procedural focus related to general classroom business or non-instructional activity. Across disciplines and treatment groups, housekeeping activities such as taking roll, passing out and collecting materials and student work, changing groupings, reviewing class assignments and school activities, and dispensing rewards consumed an average of 11 percent of class time. Teacher time spent on non-instructional housekeeping activities ranged from 0 to nearly 28 percent. An additional 11 percent of class time on average was allocated to giving instructions at the onset of tasks[17]. Taken together, teachers spent nearly a quarter of lesson time on housekeeping and giving instructions.

We also examined occurrence and duration of teacher modeling, guidance and support. Because teachers often provided intermittent rather than continuous support for students to do the work of reading and learning—for example, visiting small groups periodically during group work or interjecting brief whole class modeling or guidance during individual work —when this interaction occurred, we coded the presence of teacher modeling, guidance and support across an entire grouping segment. Thus *Teacher modeling, guidance and support* is a measure of occurrence rather than duration. Approximately 60 percent of all lesson segments contained instances of teacher modeling, guidance and support.

[17] The *Teacher gives instructions* code was reserved for instructions that set up a task. Giving additional instructions during a task was coded as *Teacher modeling, guidance and support*. In seven lessons, teachers did not give task instructions. The majority of these lessons focused on factual information delivered by the teacher in a whole class setting.

Table 15: Teacher-Student Interaction and Grouping: Occurrence by Discipline

*Table 15 shows the number and percentage of observed lessons in which each teacher-student interaction and grouping configuration occurred.*

| | All Lessons N=71 | | History N=24 | | Literature N=34 | | Science N=13 | |
|---|---|---|---|---|---|---|---|---|
| | Total number and percentage of lessons in which activity/grouping configuration occurred | | | | | | | |
| | N | % | N | % | N | % | N | % |
| **Teacher-Student Interaction** | | | | | | | | |
| Housekeeping/management | 68 | 96 | 24 | 100 | 32 | 94 | 12 | 92 |
| Teacher gives instructions | 64 | 90 | 23 | 96 | 30 | 88 | 11 | 85 |
| Teacher modeling, guidance, support | 67 | 94 | 24 | 100 | 30 | 88 | 13 | 100 |
| **Grouping** | | | | | | | | |
| Individual | 48 | 68 | 18 | 75 | 19 | 56 | 11 | 85 |
| Pairs | 28 | 39 | 12 | 50 | 12 | 35 | 4 | 31 |
| Small group | 34 | 48 | 15 | 63 | 15 | 44 | 4 | 31 |
| Whole class | 70 | 99 | 24 | 100 | 33 | 97 | 13 | 100 |

Table 16: Teacher-Student Interaction and Grouping: Duration by Discipline

*Table 16 shows the duration (in minutes) and percentage of time allocated to each teacher-student interaction and grouping configuration across observed lessons.*

| | **Total Lessons (N=71)** 3699 minutes | | **History (N=24)** 1265 minutes | | **Literature (N=34)** 1820 minutes | | **Science (N=13)** 615 minutes | |
|---|---|---|---|---|---|---|---|---|
| | Total duration[18] (in minutes) and percentage of segments in which each activity/grouping configuration occurred | | | | | | | |
| | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| **Teacher-Student Interaction** | | | | | | | | |
| Housekeeping/management | 397 | 11 | 134 | 11 | 168 | 9 | 95 | 15 |
| Teacher gives instructions | 394 | 11 | 180 | 14 | 174 | 10 | 41 | 7 |
| Teacher modeling, guidance, support[19] | 2254 | 61 | 789 | 62 | 1104 | 61 | 362 | 59 |
| **Grouping** | | | | | | | | |
| Individual | 620 | 17 | 207 | 16 | 280 | 15 | 133 | 22 |
| Pairs | 444 | 12 | 203 | 16 | 181 | 10 | 59 | 10 |
| Small group | 624 | 17 | 265 | 21 | 252 | 14 | 107 | 17 |
| Whole class | 2053 | 55 | 608 | 48 | 1110 | 61 | 335 | 55 |

[18] Duration and percentage of time across categories of Content Delivery do not add up to 100 percent of total lesson time: students sometimes learned lesson content in ways other than teacher lecture, audio/video materials, or text—e.g., through labs or simulations; time allocated to housekeeping or giving directions did not involve any content delivery; and students occasionally learned content in multiple ways—e.g., working with text while listening to an audio-recording of a play. Likewise, because students were often asked to perform multiple tasks during a single segment, percentage time across Task categories does not add up to 100 percent.

[19] Because teachers often provided intermittent rather than continuous support, we coded the presence of teacher modeling, guidance and support across the entire grouping segment to indicate the presence of teacher support in that segment. Therefore, for teacher modeling, guidance and support, the table indicates the duration of segments and percentage of time allocated to segments in which teachers provided modeling, guidance and support.

Table 17: Teacher-Student Interaction and Grouping: Occurrence by Discipline and Treatment Group

*Table 17 shows the number and percentage of observed lessons in which each teacher-student interaction and grouping configuration occurred.*

| | All Lessons | | | | History | | | | Literature | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RA (N=30) 1838 min | | Non-RA (N=41) 1862 min | | RA (N=13) 786 min | | RA (N=30) 1838 min | | Non-RA (N=41) 1862 min | | RA (N=13) 786 min | |
| | Total number and percentage of lessons in which activity/grouping occurred | | | | | | | | | | | |
| **Teacher-Student Interaction** | N | % | N | % | N | % | N | % | N | % | N | % |
| Housekeeping/management | 30 | 100 | 38 | 93 | 13 | 100 | 11 | 100 | 16 | 100 | 16 | 89 |
| Teacher gives instructions | 30 | 100 | 34 | 83 | 13 | 100 | 10 | 91 | 16 | 100 | 14 | 78 |
| Teacher modeling, guidance, support | 30 | 100 | 37 | 90 | 13 | 100 | 11 | 100 | 16 | 100 | 14 | 78 |
| **Grouping** | | | | | | | | | | | | |
| Individual | 24 | 80 | 24 | 59 | 12 | 92 | 6 | 55 | 11 | 69 | 8 | 44 |
| Pairs | 18 | 60 | 10 | 24 | 8 | 62 | 4 | 36 | 9 | 56 | 3 | 17 |
| Small group | 18 | 60 | 16 | 39 | 7 | 54 | 8 | 73 | 11 | 69 | 4 | 22 |
| Whole class | 29 | 97 | 41 | 100 | 13 | 100 | 11 | 100 | 15 | 94 | 18 | 100 |

Table 18: Teacher-Student Interaction and Grouping: Duration by Discipline and Treatment Group

*Table 18 shows the duration (in minutes) and percentage of time allocated to each teacher-student interaction and grouping configuration across observed lessons.*

| | All Lessons | | | | History | | | | Literature | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RA (N=30) 1838 min | | Non-RA (N=41) 1862 min | | RA (N=13) 786 min | | Non-RA (N=11) 478 min | | RA (N=16) 997 min | | Non-RA (N=18) 823 min | |
| | Number of minutes and percentage time across lessons | | | | | | | | | | | |
| | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| **Teacher-Student Interaction** | | | | | | | | | | | | |
| Housekeeping/management 397 min | 144 | 8 | 253 | 14 | 56 | 7 | 78 | 16 | 85 | 9 | 83 | 10 |
| Teacher gives instructions 394 min | 199 | 11 | 196 | 11 | 69 | 9 | 111 | 23 | 126 | 13 | 48 | 6 |
| Teacher modeling, guidance, support[20] 2254 min | 1183 | 64 | 1071 | 58 | 479 | 61 | 309 | 65 | 668 | 67 | 436 | 53 |
| **Grouping** | | | | | | | | | | | | |
| Individual 620 min | 361 | 20 | 259 | 14 | 151 | 19 | 56 | 12 | 204 | 20 | 76 | 9 |
| Pairs 444 min | 317 | 17 | 127 | 7 | 153 | 19 | 50 | 11 | 136 | 14 | 45 | 5 |
| Small group 624 min | 307 | 17 | 316 | 17 | 103 | 13 | 163 | 34 | 205 | 21 | 47 | 6 |
| Whole class 2053 min | 855 | 47 | 1198 | 64 | 380 | 48 | 228 | 48 | 453 | 45 | 656 | 80 |

[20] Because teachers often provided intermittent rather than continuous support, we coded the presence of teacher modeling, guidance and support across the entire grouping segment to indicate the presence of teacher support in that segment. Therefore, for teacher modeling, guidance and support, the table indicates the duration of segments and percentage of time allocated to segments in which teachers provided modeling, guidance and support.

**Grouping**

To answer the question, ***How are different grouping structures allocated across observed lessons?*** we calculated the number and percentage observed lessons, and the percentage time allocated to each Grouping code. Results from this analysis are also found in Tables 15- 18.

Grouping codes identified how students were assigned to perform a task, not their seating arrangement. As indicated by Table 16, across observed lessons, students spent relatively little time working alone, and considerable time in whole class settings—over 50 percent of class time was allocated to whole class instruction. An intermediate amount of time was spent in student-run partnerships or small groups. In addition to these general patterns, grouping structures also varied by discipline. Students in sciences classrooms spent somewhat more time working individually than students in history or literature. History students spent the most time working with peers. Across observed lessons, history teachers allocated 37 percent of class time to partner or small group work, while literature teachers allocated 24 percent and science teachers allocated 27 percent of lesson time to partners/small groups.

We also found treatment group differences related to grouping structure. As shown in Table 18, RA teachers allocated time more evenly across groupings compared with Non-RA teachers. Students in RA classrooms spent more time working individually and in pairs and less time working as a whole compared with Non-RA students. There were also disciplinary differences in time allocated to the two peer grouping structures—partnerships and small groups. In history, students in RA classrooms more frequently worked with partners, while students in Non-RA classrooms spent more time in small groups. Non-RA literature students spent the lowest percentage of time in partner or group work, and the highest percentage of time in whole class settings.

Next we turned to the question, ***Which grouping structures are associated with particular content delivery and task opportunities to learn?***—i.e., what activities and tasks students were asked to perform in different groupings. To answer this question, we examined relationships between Grouping and Content Delivery and Task codes by calculating the number and percentage of lessons in which each content delivery method and task occurred, given a particular grouping, as well as the number of minutes and percentage time allocated to each content delivery and task code when students were working in the different grouping configurations. Results are shown in Tables 19 and 20.

Content delivery was associated strongly with grouping. When teachers lectured, demonstrated or explained content, they almost always did so in a whole class setting. Likewise, students in observed lessons always viewed and listened to media as a whole class. In contrast, students worked with text in all grouping configurations, although higher percentages of individual, partner and small group time were allocated to working with text compared with whole class settings. Nonetheless, while greater percentages of time were allocated to working with text in individual and peer groupings than during whole class instruction, nearly half of all whole group interactions involved working with text, and in terms of absolute time, students worked with text in a teacher facilitated whole class setting more frequently (961 minutes) than in any other single grouping configuration. Over 40 percent of whole class time was spent working with text.

We also found a strong relationship between task type and grouping. Tasks related to argumentation literacy— evidence-based argumentation, close reading and cross-textual analysis—occurred for greater percentages of time when students worked with partners or small groups than when they worked in a whole class setting. However, in terms of absolute time, students spent more time engaged in close reading and E-BA while working as a whole class than in any other single grouping. In contrast, cross-textual analysis rarely occurred in a whole class setting (or when students were working individually). Nearly two-thirds of cross-textual analysis (66 percent) occurred when students worked with partners or small groups even though those groupings combined comprised less than a third of lesson time (29 percent). Time allocated to content learning also varied by grouping. Disciplinary knowledge building tasks occurred most frequently in partner, small group and whole class settings. Fact acquisition occurred most frequently and for the longest duration while in a whole class setting.

Table 19: Co-Occurrence of Task and Grouping: Occurrence

*Table 19 shows the number and percentage of observed lessons in which each content delivery and task type occurred when students worked in different grouping structures.*

| | Grouping | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Individual N=48** | | **Pairs N= 28** | | **Small Group N= 34** | | **Whole Class N = 70** | |
| | Number and percentage[21] of lessons in which each activity/task occurred *given a particular grouping structure* | | | | | | | |
| | N | % | N | % | N | % | N | % |
| **Content Delivery** | | | | | | | | |
| Teacher lecture, demonstration, explanation | 4 | 8 | 0 | 0 | 1 | 3 | 54 | 77 |
| Viewing/Listening | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 10 |
| Working with text | 40 | 83 | 26 | 93 | 32 | 94 | 51 | 73 |
| **Task** | | | | | | | | |
| Argumentation | 7 | 15 | 9 | 32 | 7 | 21 | 16 | 23 |
| Close reading | 19 | 40 | 15 | 54 | 15 | 44 | 31 | 44 |
| Cross-textual analysis | 4 | 8 | 4 | 14 | 6 | 18 | 10 | 14 |
| Disciplinary knowledge-building | 27 | 56 | 18 | 64 | 26 | 76 | 58 | 83 |
| Fact acquisition | 10 | 21 | 5 | 18 | 2 | 6 | 35 | 50 |
| Missed opportunity argumentation | 2 | 4 | 0 | 0 | 1 | 3 | 12 | 17 |
| Other task | 10 | 21 | 4 | 14 | 5 | 15 | 26 | 37 |
| Writing | 41 | 85 | 20 | 71 | 25 | 74 | 49 | 70 |

Table 20: Co-Occurrence of Task and Grouping: Duration

*Table 20 shows the number of minutes and percentage of time allocated to each content delivery and task type when students worked in different grouping structures.*

| | Grouping | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Individual** 620 minutes | | **Pairs** 444 minutes | | **Small Group** 624 minutes | | **Whole Class** 2053 minutes | |
| | Number of minutes and percentage time allocated to each activity/task *given a particular grouping structure* | | | | | | | |
| | Minutes | % Time | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| **Content Delivery** | | | | | | | | |
| Teacher lecture, demonstration, explanation | 31 | 5 | 0 | 0 | 0 | 0 | 759 | 37 |
| Viewing/Listening | 0 | 0 | 0 | 0 | 0 | 0 | 38 | 2 |
| Working with text | 486 | 78 | 412 | 93 | 541 | 87 | 961 | 47 |
| **Task** | | | | | | | | |
| Argumentation | 71 | 11 | 79 | 18 | 128 | 21 | 261 | 13 |
| Close reading | 227 | 37 | 208 | 47 | 195 | 31 | 416 | 20 |
| Cross-textual analysis | 18 | 3 | 55 | 12 | 131 | 21 | 80 | 4 |
| Disciplinary knowledge-building | 260 | 42 | 222 | 50 | 479 | 77 | 1086 | 53 |
| Fact acquisition | 49 | 8 | 60 | 13 | 15 | 2 | 480 | 23 |
| Missed opportunity argumentation | 6 | 1 | 0 | 0 | 2 | 0 | 94 | 5 |
| Other task | 53 | 9 | 12 | 3 | 17 | 3 | 121 | 6 |
| Writing | 476 | 77 | 311 | 70 | 441 | 71 | 602 | 29 |

*Disciplinary differences*. We also examined disciplinary differences related to content delivery and tasks students were asked to perform in various groupings. Results of this analysis are shown in Tables 21 and 22. We found both similarities and differences related to co-occurrence of activities and groupings as a function of discipline. Across disciplines, individual and whole class instruction were found in the highest number of lessons. Whole class instruction occurred in nearly every lesson. The percentage of lessons with individual instruction was highest for science (85 percent) and lowest for literature (56 percent). History had the greatest diversity of groupings, with individual, partner, small group and whole class instruction each occurring in at least half of all history lessons. Science lessons had the least diversity, with fewer than a third of science lessons incorporating either partner or small group work.

In terms of co-occurrence of content delivery method and grouping, as shown in Table 21, both teacher delivery of content and working with text during whole class instruction occurred in a higher percentage of science than history or literature lessons. These findings are elucidated when we look at *duration* of the co-occurrence of these activities. As shown in Table 22, science teachers spent an appreciably higher percentage of whole class time lecturing, demonstrating or explaining than teachers in the other disciplines—more than twice as much time as literature teachers (64 versus 27 percent of whole class instruction, respectively). In contrast, duration figures suggest that while working with text and whole class instruction co-occurred in a higher percentage of science lessons, the percentage of time allocated to working with text in a whole class setting is comparable across disciplines. Just under half of whole class instruction across disciplines was allocated to working with text.

We also found differences in occurrence and duration of content delivery methods in student-led groups. Science teachers were less likely than history or literature teachers to ask students to work in partnerships or small groups. However, when science teachers did include these groupings, chances were high that they involved working with text—100 percent of partnerships and 75 percent of small groups in science included work with text. Nonetheless, as shown in Table 22, science teachers allocated less time to working with text in peer-led groups compared with history and literature teachers. Disciplinary differences related to time allocated to working with text held up across all grouping configurations with the exception of whole class instruction—although the occurrence of lessons that incorporated working with text in individual, partner and small group configurations was generally comparable across disciplines, the duration of working with text was shorter in science compared with the other two disciplines.

We also examined disciplinary differences in tasks assigned to students in various grouping configurations. As shown in Table 21, there were both similarities and differences in co-occurrence of tasks and groupings related to argumentative literacy[22]--evidence based argumentation, close reading and evidence-based argumentation. In both history and literature, close reading occurred in a substantial percentage of lessons across all four groupings. In terms of duration, Table 22 suggests that literature teachers allocated more time to close reading in small groups than in partnerships, while the opposite tended to be true of history teachers. In literature, evidence-based argumentation was of longer duration in small group and whole class settings. In history, the percentage of time

[22] Close reading, cross-textual analysis or evidence-based argumentation—the elements of argumentation literacy—did not occur in any science lesson.

allocated to argumentation was similar across social settings. Cross textual analysis tasks occurred most frequently with peer support, either with partners (history), or in small groups (literature and history)—perhaps a tacit acknowledgement that cross textual analysis requires social support. Nonetheless, both history and literature teachers assigned a variety of tasks across grouping configurations. In contrast, grouping structures in science appeared to be more limited and specialized. Small groups in science were used exclusively for disciplinary knowledge building and writing tasks (96 and 59 percent respectively). In contrast, partner work in science involved relatively low levels of disciplinary knowledge-building (16 percent) and high levels of fact acquisition and writing—while only 59 minutes were devoted to partner work across all 13 science lessons (only four science classrooms incorporated partner work), nearly three-quarters of that time was allocated to fact acquisition and writing. As shown in Table 21, fact acquisition tasks occurred in a higher percentage of science lessons during individual, small group and whole class instruction. Duration figures suggest that time allocated to fact acquisition was substantially higher during individual and partner work in science, but was comparable to literature and slightly lower than history during whole class instruction. While fact acquisition activities were relatively rare in history and literature lessons in general, they appeared most frequently in whole class settings.

Table 21: Co-Occurrence of Task and Grouping: Occurrence

*Table 21 shows the number and percentage of observed lessons in which each content delivery and task type occurred when students worked in different grouping structures, by discipline.*

| | Grouping | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Individual | | | | | | Pairs | | | | | | Small Group | | | | | | Whole Class | | | | | |
| | History N=18 75% | | Literature N=19 56% | | Science N=11 85% | | History N=12 50% | | Literature N=12 35% | | Science N=4 31% | | History N=15 63% | | Literature N=15 44% | | Science N=4 31% | | History N=24 100% | | Literature N=33 97% | | Science N=13 100% | |
| **Content Delivery** | **Number of lessons and percentage of lessons in which activity/task occurred *given a particular grouping structure*[23]** | | | | | | | | | | | | | | | | | | | | | | | |
| | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % |
| Teacher lecture, demonstration, explanation | 0 | 0 | 1 | 5 | 3 | 27 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 7 | 0 | 0 | 0 | 0 | 19 | 79 | 23 | 70 | 12 | 92 |
| Viewing/Listening | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 13 | 3 | 9 | 1 | 8 |
| Working with text | 15 | 83 | 16 | 84 | 9 | 82 | 10 | 83 | 12 | 100 | 4 | 100 | 15 | 100 | 14 | 93 | 3 | 75 | 15 | 63 | 25 | 76 | 11 | 85 |
| **Task** | | | | | | | | | | | | | | | | | | | | | | | | |
| Argumentation | 3 | 17 | 4 | 21 | 0 | 0 | 7 | 58 | 2 | 17 | 0 | 0 | 5 | 33 | 2 | 13 | 0 | 0 | 8 | 33 | 8 | 24 | 0 | 0 |
| Close reading | 7 | 39 | 12 | 63 | 0 | 0 | 6 | 50 | 9 | 75 | 0 | 0 | 6 | 40 | 9 | 60 | 0 | 0 | 10 | 42 | 21 | 64 | 0 | 0 |
| Cross-textual analysis | 3 | 17 | 1 | 5 | 0 | 0 | 3 | 25 | 1 | 8 | 0 | 0 | 3 | 20 | 3 | 20 | 0 | 0 | 5 | 21 | 5 | 15 | 0 | 0 |
| Disciplinary knowledge-building | 9 | 50 | 9 | 47 | 9 | 82 | 6 | 50 | 10 | 83 | 2 | 50 | 12 | 80 | 10 | 67 | 4 | 100 | 18 | 75 | 28 | 85 | 12 | 92 |
| Fact acquisition | 4 | 22 | 2 | 11 | 4 | 36 | 3 | 25 | 0 | 0 | 2 | 50 | 2 | 13 | 0 | 0 | 0 | 0 | 12 | 50 | 15 | 45 | 8 | 62 |
| Missed opportunity argumentation | 1 | 6 | 0 | 0 | 1 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 7 | 0 | 0 | 4 | 17 | 6 | 18 | 2 | 15 |
| Other task | 2 | 11 | 6 | 32 | 2 | 18 | 3 | 25 | 0 | 0 | 1 | 25 | 3 | 20 | 2 | 13 | 0 | 0 | 7 | 29 | 14 | 42 | 5 | 38 |
| Writing | 14 | 78 | 16 | 84 | 11 | 100 | 10 | 83 | 8 | 67 | 2 | 50 | 12 | 80 | 10 | 67 | 3 | 75 | 20 | 83 | 22 | 67 | 7 | 54 |

Table 22: Co-Occurrence of Task and Grouping: Duration by Discipline

*Table 22 shows the percentage of time allocated to each content delivery and task type when students worked in different grouping structures for each discipline.*

| | Grouping Configuration | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Individual | | | Pairs | | | Small Group | | | Whole Class | | |
| | History 207 min/ 16% | Literature 280 min/ 15% | Science 133 min/ 22% | History 203 min/ 16% | Literature 181 min/ 10% | Science 59 min/ 10% | History 265 min 21% | Literature 252 min 14% | Science 107 min/ 17% | History 608 min 48% | Literature 1110 min/ 61% | Science 335 min/ 55% |
| **Content Delivery** | Percentage of time allocated to each activity/task *given a particular grouping structure* [24] | | | | | | | | | | | |
| Teacher lecture, demonstration, explanation | 0 | 4 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 41 | 27 | 64 |
| Viewing/Listening | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 0 |
| Working with text | 83 | 81 | 66 | 93 | 94 | 89 | 94 | 96 | 47 | 43 | 49 | 45 |
| **Task** | | | | | | | | | | | | |
| Argumentation | 25 | 7 | 0 | 35 | 5 | 0 | 36 | 13 | 0 | 21 | 12 | 0 |
| Close reading | 44 | 49 | 0 | 64 | 43 | 0 | 17 | 59 | 0 | 23 | 25 | 0 |
| Cross-textual analysis | 5 | 3 | 0 | 26 | 2 | 0 | 23 | 28 | 0 | 9 | 3 | 0 |
| Disciplinary knowledge-building | 39 | 31 | 70 | 46 | 65 | 16 | 80 | 65 | 96 | 50 | 51 | 65 |
| Fact acquisition | 5 | 1 | 27 | 8 | 0 | 73 | 6 | 0 | 0 | 27 | 22 | 21 |
| Missed opportunity argumentation | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 4 | 6 | 2 |
| Other task | 1 | 14 | 9 | 3 | 0 | 11 | 4 | 2 | 0 | 6 | 6 | 4 |
| Writing | 90 | 61 | 89 | 73 | 67 | 72 | 69 | 77 | 59 | 38 | 23 | 34 |

Duration across categories of Content Delivery does not add up to 100 percent of total lesson time for several reasons: students sometimes learned lesson content in ways other than teacher lecture, audio/video materials, or text—e.g., through labs or simulations; time allocated to housekeeping and giving directions did not involve any content delivery; and students occasionally learned content in multiple ways—e.g., working with text while listening to an audio-recording of a play. Likewise, because students were often asked to perform multiple tasks during a single segment, percentage time across Task categories does not add up to 100 percent.

*Treatment group differences*. We also examined treatment group differences related to content delivery methods and tasks students were asked to perform in different grouping configurations. Results are shown in Table 23. For content delivery, compared with RA teachers, Non-RA teachers lectured, demonstrated and/or explained for longer duration during both individual and whole class instruction—the only two groupings in which teacher delivery of content occurred. Students in RA classrooms worked with text for longer duration individually, in small groups and as a whole class compared with Non-RA students. Partners in both treatment and comparison groups worked with text for comparable percentages of time; almost all partner work in both treatment and comparison groups involved working with text.

We also found treatment group differences in co-occurrence of grouping structures and tasks. As shown in Table 23, in RA lessons, the percentage of time allocated to evidence-based argumentation was distributed relatively evenly across all four grouping configurations. In Non-RA lessons, considerably more time was allocated to evidence-based argumentation in partnerships and small groups than in individual or whole class settings. Whole class instruction was associated with somewhat different tasks in RA and Non-RA lessons. In RA classrooms, whole class instruction was most strongly associated with evidence-based argumentation, close reading, disciplinary knowledge building and writing. In Non-RA classrooms, whole class instruction was most strongly associated with disciplinary knowledge building, fact acquisition and writing. This suggests that whole class instruction played different roles in the two groups, with RA teachers more frequently using whole class time to support disciplinary reading and argumentation, and Non-RA teachers using whole class instruction to support content learning. Additional evidence of this comes from the greater percentage of time allocated to teacher lecture, demonstration and explanation during whole class instruction in Non-RA lessons—Non-RA teachers spent more than twice as much time delivering content to the whole class compared with RA teachers (49 and 20 percent respectively). In addition, whereas 22 percent of whole class time was allocated to evidence-based argumentation in RA classrooms, in Non-RA classrooms, only 6 percent of whole class instruction was allocated to argumentation. Researchers identified missed opportunities for argumentation 7 percent of the time during whole class instruction in Non-RA classrooms compared with 1 percent of the time in the RA classrooms. In both RA and Non-RA lessons, relatively little time was allocated to writing during whole class instruction compared with other grouping configurations, suggesting that whole class interactions may have emphasized other language modalities—e.g., reading, speaking and/or listening.

Table 23: Co-Occurrence of Task and Grouping: Duration by Treatment Group

*Table 23 shows treatment group differences the percentage of time allocated to each content delivery and task type when students worked in different grouping structures.*

| | **Grouping Configuration** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Individual** | | **Pairs** | | **Small Group** | | **Whole Class** | |
| | RA (N=24/80%) | Non-RA (N=24/59%) | RA (N=18/60%) | Non-RA (N=10/24%) | RA (N=18/60%) | Non-RA (N=16/39%) | RA (N=29/97%) | Non-RA (N=41/100%) |
| **Content Delivery** | Percentage of time allocated to each activity/task *given a particular grouping structure* | | | | | | | |
| Teacher lecture, demonstration, explanation | 0 | 12 | 0 | 0 | 0 | 0 | 20 | 49 |
| Viewing/Listening | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 1 |
| Working with text | 86 | 68 | 92 | 96 | 96 | 78 | 57 | 39 |
| **Task** | | | | | | | | |
| Argumentation | 15 | 6 | 13 | 29 | 19 | 22 | 22 | 6 |
| Close reading | 54 | 12 | 65 | 2 | 50 | 13 | 34 | 10 |
| Cross-textual analysis | 3 | 3 | 9 | 20 | 30 | 12 | 8 | 1 |
| Disciplinary knowledge-building | 36 | 50 | 48 | 54 | 69 | 84 | 58 | 49 |
| Fact acquisition | 3 | 15 | 8 | 27 | 1 | 4 | 10 | 33 |
| Missed opportunity argumentation | 0 | 2 | 0 | 0 | 1 | 0 | 1 | 7 |
| Other task | 4 | 15 | 1 | 6 | 2 | 3 | 6 | 6 |
| Writing | 79 | 74 | 68 | 77 | 65 | 76 | 24 | 33 |

**Grade Band Differences in Opportunity to Learn**

We also considered differences in student opportunity to learn across the grade bands defined in the Common Core State Standards (CCSS, 2010). To answer the question, ***How do Content Delivery and Task opportunity to learn change by grade level?***, we calculated the number and percentage of observed lessons, and percentage of time allocated to each Content Delivery and Task code by grade band. We also examined grade level changes by discipline and treatment group. Results from this analysis are found in Tables 24-32.

*Content delivery and task opportunity to learn***.** As indicated in Table 24, the percentage of lessons in which teachers delivered content declined incrementally as students moved from grades 6-8 to grades 11-12 while the percentage of lessons in which students were asked to work with text increased during this same time period. The percentage of time allocated to different content delivery methods mirrors these findings, as shown in Table 25. Teacher lecturing, demonstrating and explaining declined from 30 percent to 11 percent of lesson time from grades 6-8 to grades 11-12, while the duration of working with text increased from 53 to 75 percent.

Students' opportunity to learn tasks associated with argumentation literacy also increased as they moved up the grade levels. The percentage of lessons that included any evidence-based argumentation, close reading and cross-textual analysis increased dramatically from middle school to grades 11-12, along with the percentage of time allocated to these tasks. Fact acquisition activities declined commensurately.

Table 24: Grade-Related Differences in Content Delivery and Task: Occurrence

*Table 24 shows the number and percentage of observed lessons in which each content delivery and task type occurred for each grade band.*

| | Grade-Level Band | | | | | |
|---|---|---|---|---|---|---|
| | Grades 6-8 (N=21) | | Grades 9-10 (N=28) | | Grades 11-12 (N=22) | |
| | Number and percentage of lessons in which each activity/task occurred | | | | | |
| | N | % | N | % | N | % |
| **Content Delivery** | | | | | | |
| Teacher lecture, demonstration, explanation | 17 | 81 | 23 | 82 | 14 | 64 |
| Viewing/Listening | 2 | 10 | 5 | 18 | 0 | 0 |
| Working with text | 18 | 86 | 27 | 96 | 21 | 95 |
| **Task** | | | | | | |
| Argumentation | 4 | 19 | 8 | 29 | 9 | 41 |
| Close reading | 8 | 38 | 13 | 46 | 15 | 68 |
| Cross-textual analysis | 3 | 14 | 4 | 14 | 7 | 32 |
| Disciplinary knowledge-building | 17 | 81 | 25 | 89 | 19 | 86 |
| Fact acquisition | 13 | 62 | 16 | 57 | 10 | 45 |
| Missed opportunity argumentation | 4 | 19 | 6 | 21 | 3 | 14 |
| Other task | 8 | 38 | 12 | 43 | 10 | 45 |
| Writing | 18 | 86 | 27 | 96 | 21 | 95 |

Table 25: Grade-Related Differences in Content Delivery and Task: Duration

*Table 25 shows the number of minutes and percentage of time allocated to each content delivery and task type for each grade band.*

| | Grade-Level Band | | | | | |
|---|---|---|---|---|---|---|
| | **Grades 6-8 (N=21) 1045 minutes** | | **Grades 9-10 (N=28) 1430 minutes** | | **Grades 11-12 (N=22) 1225 minutes** | |
| | Number of minutes and percentage time allocated to each activity/task for each grade level | | | | | |
| **Content Delivery** | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| Teacher lecture, demonstration, explanation | 313 | 30 | 329 | 23 | 132 | 11 |
| Viewing/Listening | 11 | 1 | 27 | 2 | 0 | 0 |
| Working with text | 552 | 53 | 906 | 63 | 918 | 75 |
| **Task** | | | | | | |
| Argumentation | 91 | 9 | 167 | 12 | 260 | 21 |
| Close reading | 229 | 22 | 337 | 24 | 485 | 40 |
| Cross-textual analysis | 39 | 4 | 50 | 3 | 195 | 16 |
| Disciplinary knowledge-building | 453 | 43 | 906 | 63 | 682 | 56 |
| Fact acquisition | 266 | 25 | 225 | 16 | 108 | 9 |
| Missed opportunity argumentation | 16 | 2 | 54 | 4 | 32 | 3 |
| Other task | 51 | 5 | 68 | 5 | 87 | 7 |
| Writing | 426 | 41 | 768 | 54 | 615 | 50 |

*Disciplinary differences*. Tables 26 and 27 show findings related to grade band differences in content delivery and task by discipline. The trajectory of grade-related changes is especially apparent in relation to the percentage of time allocated to various activities and tasks. For example, while Table 26 indicates that the vast majority of lessons across grade bands and disciplines included some teacher delivery of content, Table 27 shows that time allocated to these activities differed dramatically by grade and discipline. Both history and literature lessons showed incremental patterns of change in methods of content delivery from middle school to grades 11-12. Time allocated to teacher lecture, demonstration and explanation decreased steadily as students moved up the grade levels, while working with text increased from grades 6-8 to grades 11 and 12. In science, in contrast, we saw an increase in teacher delivery of content as students moved from middle school to grades 11-12[25]. Whereas science teachers in middle school science lectured, demonstrated and explained 24 percent of class time, in grades 9-10, the teacher lectured, demonstrated and explained 34 percent of the time and in grades 11-12, the teacher delivered content 81 percent of the time. Time allocated to working with text remained relatively constant from middle school to grades 9-10. Students in the single science lesson in the grade 11-12 grade band worked with text during 87 percent of the lesson.

Grade band changes in task opportunity to learn paralleled trends in content delivery. History classes showed incremental grade level increases in tasks focused on disciplinary reading and learning, particularly tasks related to argumentation literacy. Specifically, the percentage of time allocated to argumentation, close reading, and cross textual analysis increased and fact acquisition activities decreased from middle school to grades 11-12. Literature lessons showed fewer incremental changes in tasks related to argumentation literacy, with the exception of cross-textual analysis, which increased from 2 percent in middle school to 13 percent in grades 11-12. Approximately two-thirds of literature lessons across grade bands offered close reading opportunities, and approximately a third of class time was spent in this activity. Time allocated to evidence-based argumentation was similar in middle school and grades 11-12 (11 and 13 percent respectively), but dipped in grades 9-10. As noted earlier, no science lesson at any grade level incorporated argumentation, close reading or cross-textual analysis. Science opportunities to learn focused on content at all grade levels, with both fact acquisition and disciplinary knowledge building tasks showing incremental increases from middle school to grades 11-12. The increase in fact acquisition tasks in science classrooms between middle and high school was especially dramatic—mirroring the increase in time allocated to teacher lecturing, demonstrating and explaining.

[25] It is important to keep in mind that findings related to grade 11-12 science are based on a single lesson.

Table 26: Grade-Related Differences in Content Delivery and Task: Occurrence by Discipline

*Table 26 shows the number and percentage of observed lessons in which each content delivery and task type occurred for each grade band by discipline.*

| | Grade-Level Band | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Grades 6-8 (N=21) 1045 minutes | | | | | | Grades 9-10 (N=28) 1430 minutes | | | | | | Grades 11-12 (N=22) 1225 minutes | | | | | |
| | History N = 6 | | Literature N = 11 | | Science N = 4 | | History N = 10 | | Literature N = 10 | | Science N = 8 | | History N = 8 | | Literature N = 13 | | Science N = 1 | |
| | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % | N | % |
| **Content Delivery** | Number and percentage of lessons in which each activity/task occurred | | | | | | | | | | | | | | | | | |
| Teacher lecture, demonstration, explanation | 6 | 100 | 8 | 73 | 3 | 75 | 8 | 80 | 7 | 70 | 8 | 100 | 5 | 63 | 8 | 62 | 1 | 100 |
| Viewing/Listening | 2 | 33 | 0 | 0 | 0 | 0 | 1 | 10 | 3 | 30 | 1 | 13 | 0 | 0 | 0 | 0 | 0 | 0 |
| Working with text | 5 | 83 | 9 | 82 | 4 | 100 | 10 | 100 | 10 | 100 | 7 | 88 | 8 | 100 | 12 | 92 | 1 | 100 |
| **Task** | | | | | | | | | | | | | | | | | | |
| Argumentation | 1 | 17 | 3 | 27 | 0 | 0 | 6 | 60 | 2 | 20 | 0 | 0 | 5 | 63 | 4 | 31 | 0 | 0 |
| Close reading | 1 | 17 | 7 | 64 | 0 | 0 | 6 | 60 | 7 | 70 | 0 | 0 | 6 | 75 | 9 | 69 | 0 | 0 |
| Cross-textual analysis | 2 | 33 | 1 | 9 | 0 | 0 | 3 | 30 | 1 | 10 | 0 | 0 | 3 | 38 | 4 | 31 | 0 | 0 |
| Disciplinary knowledge-building | 4 | 67 | 9 | 82 | 4 | 100 | 9 | 90 | 8 | 80 | 8 | 100 | 7 | 88 | 11 | 85 | 1 | 100 |
| Fact acquisition | 5 | 83 | 5 | 45 | 3 | 75 | 6 | 60 | 5 | 50 | 5 | 63 | 3 | 38 | 6 | 46 | 1 | 100 |
| Missed opportunity argumentation | 1 | 17 | 2 | 18 | 1 | 25 | 3 | 30 | 2 | 20 | 1 | 13 | 0 | 0 | 3 | 23 | 0 | 0 |
| Other task | 4 | 67 | 2 | 18 | 2 | 50 | 4 | 40 | 5 | 50 | 3 | 38 | 1 | 13 | 9 | 69 | 0 | 0 |
| Writing | 6 | 100 | 9 | 82 | 3 | 75 | 10 | 100 | 9 | 90 | 8 | 100 | 8 | 100 | 12 | 92 | 1 | 100 |

Table 27: Grade-Related Differences in Content Delivery and Task: Duration by Discipline

*Table 27 shows the percentage of time allocated to each content delivery and task type for each grade band by discipline.*

| | Grade-Level Band | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Grades 6-8 (N=21) 1045 minutes | | | Grades 9-10 (N=28) 1430 minutes | | | Grades 11-12 (N=22) 1225 minutes | | |
| | History N = 6 273 min | Literature N = 11 607 min | Science N = 4 165 min | History N = 10 517 min | Literature N = 10 517 min | Science N = 8 395 min | History N = 8 475 min | Literature N = 13 696 min | Science N = 1 55 min |
| **Content Delivery** | Percentage time allocated to each activity/task | | | | | | | | |
| Teacher lecture, demonstration, explanation | 44 | 25 | 24 | 17 | 20 | 34 | 8 | 7 | 81 |
| Viewing/Listening | 4 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 |
| Working with text | 36 | 60 | 55 | 73 | 64 | 50 | 80 | 70 | 87 |
| **Task** | | | | | | | | | |
| Argumentation | 8 | 11 | 0 | 26 | 6 | 0 | 36 | 13 | 0 |
| Close reading | 9 | 34 | 0 | 27 | 38 | 0 | 51 | 35 | 0 |
| Cross-textual analysis | 10 | 2 | 0 | 7 | 3 | 0 | 23 | 12 | 0 |
| Disciplinary knowledge-building | 49 | 37 | 56 | 61 | 62 | 67 | 52 | 55 | 93 |
| Fact acquisition | 42 | 21 | 14 | 14 | 14 | 21 | 3 | 7 | 75 |
| Missed opportunity argumentation | 2 | 1 | 2 | 4 | 5 | 2 | 0 | 5 | 0 |
| Other task | 10 | 2 | 8 | 5 | 5 | 5 | 1 | 12 | 0 |
| Writing | 58 | 38 | 25 | 52 | 44 | 69 | 66 | 43 | 12 |

*Treatment group differences*. We also examined treatment group differences in grade level changes in content delivery and task opportunities to learn. As indicated by Table 28, RA and Non-RA classrooms showed parallel grade-related trajectories in time allocated to various content delivery methods, although the magnitude of change differed. Whereas teacher delivery of content declined in both treatment and comparison groups, the decline was more dramatic in Non-RA lessons, where time allocated to teacher delivered content declined from a high of 42 percent in grades 6-8 to a low of 19 percent in grades 11-12. In RA classrooms, where teacher lecture, demonstration and explanation was relatively low across grade bands, teacher delivery of content declined from 13 percent in middle school to 7 percent in grades 11-12. Likewise, time allocated to working with text increased incrementally for both groups. RA middle school teachers allocated two-thirds of class time to working with text, while their RA counterparts in grades 11-12 allocated over 80 percent of class time to working with text. Middle school students in Non-RA classrooms worked with text 44 percent of the time, while 11-12th grade students in Non-RA classrooms worked with text 63 percent of the time. Thus, while the percentage of time students worked with text increased incrementally from in both treatment and comparison groups, students' opportunity to work with text in middle school RA classrooms exceeded students' opportunity to work with text in grade 11-12 Non-RA classrooms.

Grade band differences in tasks related to argumentation literacy generally showed consistent incremental increases in RA classrooms, while grade-related changes in Non-RA classrooms were more unpredictable. In RA lessons, time allocated to evidence-based argumentation, close reading and cross textual analysis increased from middle school to grades 11-12, and each of these tasks increased incrementally with the exception of cross textual analysis, which dipped from 9 to 3 percent of class time from middle school to grades 9-10 before increasing to 18 percent in grades 11-12. In addition, time allocated to disciplinary knowledge building increased and fact acquisition decreased steadily in RA lessons from middle school through grades 11-12. In Non-RA lessons, low magnitude grade differences in time allocated to argumentation literacy tasks were less consistent, with the exception of cross textual analysis, which increased incrementally from 0 percent in middle school to 11 percent in grades 11-12. Missed opportunity for evidence-based argumentation also increased incrementally in Non-RA classrooms from 2 percent in middle school to 8 percent in grades 11-12. Duration of close reading was relatively constant across grade levels. Time allocated to disciplinary knowledge building in Non-RA classrooms increased from middle school to high school, but was highest in grades 9-10. Likewise, fact acquisition decreased from middle to high school, but was lowest in grades 9-10. Taken together, these findings suggest that opportunities for argumentation literacy and disciplinary knowledge building increased steadily in RA lessons as students moved up the grade level. While there was a trend in Non-RA lessons toward increased challenge in the move from middle to high school—as indicated by increases in time allocated to disciplinary knowledge building and cross textual analysis and decreases in the duration of fact acquisition tasks—grade-related differences in Non-RA lessons were more erratic.

Table 28: Grade-Related Differences in Content Delivery and Task: Duration by Treatment Group

*Table 28 shows the percentage of time allocated to each content delivery and task type for each grade band by treatment group.*

| | Grade-Level Band | | | | | |
|---|---|---|---|---|---|---|
| | **Grades 6-8 (N=21) 1045 minutes** | | **Grades 9-10 (N=28) 1430 minutes** | | **Grades 11-12 (N=22) 1225 minutes** | |
| | Percentage time allocated to each activity/task | | | | | |
| | RA (N=8) 425 minutes | Non-RA (N=13) 620 minutes | RA (N=9) 583 minutes | Non-RA (N=19) 847 minutes | RA (N=13) 830 minutes | Non-RA (N=9) 395 minutes |
| **Content Delivery** | | | | | | |
| Teacher lecture, demonstration, explanation | 13 | 42 | 10 | 32 | 7 | 19 |
| Viewing/listening | 2 | 1 | 2 | 2 | 0 | 0 |
| Working with text | 66 | 44 | 75 | 55 | 81 | 63 |
| **Task** | | | | | | |
| Argumentation | 7 | 10 | 20 | 6 | 24 | 15 |
| Close reading | 38 | 11 | 44 | 10 | 52 | 13 |
| Cross-textual analysis | 9 | 0 | 3 | 4 | 18 | 11 |
| Disciplinary knowledge-building | 46 | 42 | 54 | 70 | 58 | 50 |
| Fact acquisition | 13 | 34 | 9 | 20 | 1 | 24 |
| Missed opportunity argumentation | 1 | 2 | 0 | 6 | 0 | 8 |
| Other task | 6 | 4 | 6 | 4 | 2 | 17 |
| Writing | 45 | 38 | 50 | 56 | 52 | 47 |

**Teacher-Student Interaction and Grouping by Grade Band**

We also looked at grade level changes in social interaction with teachers and peers. Results are shown in Tables 29-32. As seen in Table 29, the occurrence of all three categories of teacher-student interaction was high for all grade bands. Table 29 indicates small grade-level shifts in time allocated to housekeeping and in segments containing teacher modeling, guidance and support from middle school to grades 11-12.

Grouping structures showed larger incremental shifts as students moved up the grade levels. While whole class instruction was the single largest grouping configuration across all grade bands, time allocated to whole class instruction declined somewhat as students moved from middle to high school. And while teachers allocated comparatively little time to either partners or small groups, 11-12 grade students spent somewhat more time working with partners compared with middle school students (from 9 to 15 percent) and nearly twice as much time in small groups as middle school students (from 11 to 21 percent). These increases added up to substantially more peer-directed learning opportunities when we combine partner and small group work—time allocated to partnerships and small groups increased from 20 percent in middle school to 29 percent in grades 9-10 to 36 percent in grades 11-12 . Time allocated to individual work declined commensurately with the increase in partnerships and small groups. Middle school students spent 20 percent of the time working individually, whereas 11-12 grade students spent 12 percent of lesson time working individually.

Table 29: Grade-Related Differences in Teacher-Student Interaction and Grouping: Occurrence

*Table 29 shows the number and percentage of observed lessons in which each category of social support occurred for each grade band.*

| | Grade-Level Band | | | | | |
|---|---|---|---|---|---|---|
| | **Grades 6-8 (N=21) 1045 minutes** | | **Grades 9-10 (N=28) 1430 minutes** | | **Grades 11-12 (N=22) 1225 minutes** | |
| | Number and percentage of lessons in which each activity/task occurred | | | | | |
| | N | % | N | % | N | % |
| **Teacher-Student Interaction** | | | | | | |
| Housekeeping/management | 20 | 95 | 26 | 93 | 22 | 100 |
| Teacher gives instructions | 17 | 81 | 26 | 93 | 21 | 95 |
| Teacher modeling, guidance, support | 19 | 90 | 28 | 100 | 20 | 91 |
| **Grouping** | | | | | | |
| Individual | 14 | 67 | 19 | 68 | 15 | 68 |
| Pairs | 7 | 33 | 12 | 43 | 9 | 41 |
| Small group | 8 | 38 | 14 | 50 | 12 | 55 |
| Whole class | 21 | 100 | 27 | 96 | 22 | 100 |

Table 30: Grade-Related Differences in Teacher-Student Interaction and Grouping: Duration

*Table 30 shows the number of minutes and percentage of time allocated to each category of social support for each grade band.*

| | Grade-Level Band | | | | | |
|---|---|---|---|---|---|---|
| | **Grades 6-8 (N=21) 1045 minutes** | | **Grades 9-10 (N=28) 1430 minutes** | | **Grades 11-12 (N=22) 1225 minutes** | |
| | Number of minutes and percentage time allocated to each activity/task for each grade level | | | | | |
| | Minutes | % Time | Minutes | % Time | Minutes | % Time |
| **Teacher-Student Interaction** | | | | | | |
| Housekeeping/management | 132 | 13 | 157 | 11 | 108 | 9 |
| Teacher gives instructions | 84 | 8 | 190 | 13 | 120 | 10 |
| Teacher modeling, guidance, support | 651 | 62 | 960 | 67 | 643 | 52 |
| **Grouping** | | | | | | |
| Individual | 214 | 20 | 254 | 18 | 153 | 12 |
| Pairs | 89 | 9 | 176 | 12 | 179 | 15 |
| Small group | 117 | 11 | 248 | 17 | 258 | 21 |
| Whole class | 624 | 60 | 772 | 54 | 657 | 54 |

We also looked at grade-related differences in teacher-student interactions by discipline. As shown in Table 31, both history and science teachers allocated less time to housekeeping/management tasks as students moved up the grade levels. The decrease was especially dramatic in science, where middle school teachers spent 21 percent of class time on housekeeping and management. This declined to 7 percent in grades 11-12, although this finding is based on a single 11-12 grade science class. In history, the greatest decline in housekeeping/management activities occurred as students moved from middle to high school. In literature classrooms, housekeeping/management remained relatively constant across the grade levels. Time spent giving instructions showed a mixed pattern of change across grade bands. In history, the time teachers spent giving instructions doubled from grades 6-8 to grades 9-10, then declined again to its original level in grades 11-12. In science, time allocated to giving instructions declined incrementally from 9 to 3 percent as students moved from middle school to grades 11-12. Teacher modeling, guidance and support also showed a mixed pattern of grade-related change. In history, the percentage of segments in which teachers supported students increased incrementally; in science it decreased incrementally; and in literature, teacher modeling, guidance and support increased from middle school to grades 9-10, then decreased in grades 11-12.

Table 31 also shows grade level changes in time allocated to various grouping structures. In science, whole class activities increased from middle school to grades 11-12, while individual and small group work decreased. These changes paralleled the increase in teacher lecture from middle school to grades 11-12 (see Table 27 above). It is suggestive that in middle school and grade 9-10 science lessons, the duration of coded content delivery (i.e., the sum of teacher lecture, demonstration and explanation, viewing/listening, and working with text) fell short of 100 percent (see Table 25). This suggests that other modes of learning content may occur in science that do not involve teacher lecture, media or working with text. It may be that "hands on" science conducted in small groups in middle school gives way to whole class teacher lecture, demonstration and explanation as students move up the grade levels, for example. However, this finding may well be an artifact of having only one 11-12 grade science lesson in this corpus of data.

In history lessons, in contrast, students appeared to become increasingly independent of teacher direct instruction as they moved up the grade levels. The percentage of time spent as a whole class decreased as students transitioned from middle school to high school, and time allocated to partnerships and small groups combined increased incrementally from middle school to grades 11-12 (from 21 to 38 to 46 percent or lesson time). Literature lessons also showed an increase in time spent in partner/small group work in grade 11-12, from 17 percent in middle school, to 22 percent in grades 9-10, to 32 percent in grades 11-12. In literature lessons, time allocated to whole class activities was high across the grades, and the increase in partner and small group work was carved primarily out of time spent working individually, which decreased as literature students moved from middle to high school.

Table 31: Grade-Related Differences in Teacher-Student Interaction and Grouping: Duration by Discipline

*Table 31 shows the percentage of time allocated to each content delivery and task type for each grade band by discipline.*

| | Grade-Level Ban | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Grades 6-8 (N=21) 1045 minutes** | | | **Grades 9-10 (N=28) 1430 minutes** | | | **Grades 11-12 (N=22) 1225 minutes** | | |
| | History (N = 6) 273 min | Literature (N= 11) 607 min | Science (N = 4) 165 min | History (N = 10) 517 min | Literature (N = 10) 517 min | Science (N = 8) 395 min | History (N = 8) 475 min | Literature (N = 13) 696 min | Science (N = 1) 55 min |
| **Teacher-Student Interaction** | Percentage time allocated to each activity/task | | | | | | | | |
| Housekeeping/management 397 min | 16 | 9 | 21 | 9 | 11 | 14 | 9 | 9 | 7 |
| Teacher gives instructions 394 min | 10 | 7 | 9 | 21 | 11 | 6 | 9 | 11 | 3 |
| Teacher modeling, guidance, support 2254 min | 79 | 60 | 42 | 68 | 70 | 62 | 46 | 54 | 86 |
| **Grouping** | | | | | | | | | |
| Individual 620 min | 10 | 23 | 28 | 22 | 11 | 21 | 14 | 11 | 9 |
| Pairs 444 min | 8 | 11 | 2 | 12 | 11 | 14 | 25 | 9 | 0 |
| Small group 624 min | 13 | 6 | 28 | 26 | 11 | 15 | 21 | 23 | 0 |
| Whole class 2053 min | 70 | 60 | 42 | 40 | 67 | 55 | 44 | 57 | 91 |

Finally, we examined grade-related differences in teacher-student interaction and grouping by treatment group. These results are shown in Table 32. Findings suggest that for Non-RA lessons, the greatest grade-related changes in both teacher-student interaction and grouping occurred in the transition from middle to high school. There were peaks at grades 9-10 in time allocated to both giving instruction and providing support. Partner and group work were also highest in grades 9-10 in Non-RA lessons. In contrast, when we saw grade-related changes in RA lessons the greatest magnitude of change generally occurred *during* high school, between grades 9-10 and 11-12. For example, teacher modeling, guidance and support remained steady in grades 6-8 and 9-10, and then dropped substantially in grades 11-12. Time allocated to partner and small group work combined remained constant during middle school and grades 9-10, and then increased significantly in grades 11-12. The exceptions to this were an incremental reduction in individual work from middle school to grades 11-12, and a small peak in whole class instruction in grades 9-10.

Table 32: Grade-Related Differences in Social Support: Duration by Treatment Group

*Table 32 shows the percentage of time allocated to each content category of social support by grade band and treatment group.*

| | Grouping | | | | | |
|---|---|---|---|---|---|---|
| | **Grades 6-8 (N=21) 1045 minutes** | | **Grades 9-10 (N=28) 1430 minutes** | | **Grades 11-12 (N=22) 1225 minutes** | |
| | RA (N = 8) 425 minutes | Non-RA (N = 13) 620 minutes | RA (N =9) 583 minutes | Non-RA (N = 19) 847 minutes | RA (N =13) 830 minutes | Non-RA (N = 9) 395 minutes |
| **Teacher-Student Interaction** | Percentage time allocated to each category of social support | | | | | |
| Housekeeping/management | 8 | 15 | 8 | 13 | 7 | 12 |
| Teacher gives instructions | 11 | 6 | 10 | 15 | 11 | 7 |
| Teacher modeling, guidance, support | 75 | 53 | 78 | 60 | 49 | 59 |
| **Grouping** | | | | | | |
| Individual | 31 | 13 | 24 | 14 | 11 | 16 |
| Pairs | 20 | 1 | 11 | 13 | 20 | 2 |
| Small group | 8 | 13 | 15 | 19 | 23 | 18 |
| Whole class | 41 | 73 | 51 | 56 | 46 | 69 |

## Conclusions and Suggestions for Further Research

In addition to describing opportunities to learn key elements of argumentation literacy in 71 secondary history, literature and science classrooms, findings from this study raise important questions related to supporting the development of disciplinary reading and argumentation. Many of these issues surfaced from affordances of our posttest-only quasi-experimental design with nonequivalent groups. These two groups, with potentially different practices related to teaching and learning of argumentation literacy, comprised a "treatment group" of experienced Reading Apprenticeship teachers focused on supporting close disciplinary reading and a "comparison group" of teachers who had not participated in Reading Apprenticeship professional development but whose instruction was reported to foster disciplinary literacies in history, science, and literature. The following questions obtain their substance from this design.

*The role of text and reading in argumentation*. Project READI defines reading comprehension as the ability to engage in evidence-based argumentation, yet the practice of using text as the foundation for argumentation is uncommon in schools (Schwarz and Asterhan, 2012). While our analysis of video data from 71 history, literature and science lessons found that evidence-based argumentation often kept company with close reading and cross textual analysis, more than two-thirds of evidence-based argumentation occurred absent these "building blocks" of argumentation. This was especially true for comparison classrooms, where argumentation tasks rarely occurred in the context of close reading.

In a recent review of argumentation and reasoning in the International Handbook of Psychology in Education, Schwarz and Asterhan (2012) argue that studying when and how texts are incorporated in argumentative talk is an important issue for understanding argumentation literacy. Our preliminary analysis of field notes, lesson artifacts and teacher interviews revealed a reciprocal relationship between reading and argumentation. In classrooms where teachers provided time and support for close reading and cross textual analysis, students showed high levels of engagement and perseverance in text-based problem solving and evidence-based argumentation with challenging texts. Likewise, extant research suggests that while providing texts at the beginning of discussion between students may be too complex when reading was unstructured, collaborative close reading and meaning-making results in productive argumentation and learning (Baker, 2003; Schwarz, 2003). Well established close reading routines and supports for sense-making around challenging texts thus appear key to fostering text-based disciplinary argumentation with the potential for learning. The finding that argumentation only infrequently occurred in the context of close reading in comparison classrooms begs the question of what the characteristics and affordances might be of argumentation tasks that occur in the virtual absence of close reading opportunities, versus those accompanied by close reading.

*Teacher-student interactions*. Teachers in our classroom observations spent nearly a quarter of lesson time on housekeeping and giving instructions. Our preliminary analysis based on field notes, lesson artifacts and teacher interviews suggested that housekeeping routines (or their absence) and teacher instructions often played an important role in positioning tasks and texts as either procedural display or inquiry (Jiménez-Aleixandre, Rodríguez, & Duschl, 2000). Furthermore, the preliminary analysis suggested that organizational routines, participation structures and discourse routines embedded in *Housekeeping/management* and *Teacher gives instructions* often exerted a powerful influence on classroom climate. Given the significant time devoted to these activities in the study, it would be

worthwhile to know more about how housekeeping, management and giving instructions shape epistemological framing and other features of classroom climate.

*Questions about grouping structures and E-BA*. Our analysis of what activities and tasks students were asked to perform in different groupings revealed that evidence-based argumentation occurred across all four grouping configurations in RA lessons, whereas in Non-RA classrooms, E-BA occurred primarily in partnerships and small groups. Furthermore, whole class instruction was most strongly associated with evidence-based argumentation, close reading, disciplinary knowledge building and writing in RA classrooms, while in Non-RA classrooms, whole class instruction was most strongly associated with disciplinary knowledge building, fact acquisition and writing. Taken together, it appears that students in comparison classroom may receive less support for argumentation literacy compared with RA classrooms where argumentation tasks were distributed across individual, teacher- and peer-supported settings. An important issue for understanding the development of argumentation literacy is to identify which groupings, cycles or progressions of groupings might best support students in learning and practicing different elements of evidence-based argumentation.

*Questions about grade-related differences in opportunity to learn*. Our findings suggest that treatment and comparison groups may have different grade-related trajectories for academic challenge and support, and that these differences may have the greatest import for middle school. In Non-RA classrooms, grade-related changes in task opportunity to learn and social support from teachers and peers were greatest in the transition from middle to high school. The move from middle school to grade 9-10 was characterized by decreases in teacher lecture, demonstration and explanation and in fact acquisition tasks, and increases in disciplinary knowledge building and participation in peer-directed groupings. Middle school RA teachers allocated more time to higher level literacy tasks, provided higher levels of teacher modeling, guidance and support and had students work more frequently in peer-directed settings than did Non-RA teachers. These findings suggest that middle school students in RA classrooms may experience both greater autonomy and greater support than their Non-RA counterparts. Given the well-documented mismatch between the developmental needs and capabilities of middle school students and normative middle school curriculum and instruction, this finding begs the question of the degree to which focusing curriculum and support on disciplinary reading and argumentation literacy in middle school might contribute to better alignment between the developmental capabilities and needs of middle school students and their school experiences (see Farrington et al., 2012, for a discussion of the mismatch between developmental needs and instruction in middle school, and its impact on the development of academic mindsets).

*Questions about opportunities to learn and student engagement*. One approach to establishing predictive validity is to assess the extent to which the practices we observed are associated with desired student outcomes in earlier studies. For example, an extensive body of research on discourse-based instructional approaches is based on the assumption that opportunities for student talk increase and support student reading comprehension and learning (Applebee et al., 2003; Beck et al., 1997; Brown et al., 1993; Goldenberg, 1992; Hiebert &Wearne, 1993; Rosenshine & Meister, 1994). Nystrand & Gamoran based their long program of research on classroom discourse on the assumption that "some tasks, and some patterns of interaction are inherently more substantively engaging. For these reasons, we study student engagement by examining the classroom activities in which students are involved" (Nystrand, Gamoran, and NCESS, 1990, p. 4). While data capture limitations prevented

us from directly exploring links between teacher practices to student outcomes in this data set, we might likewise draw on extant knowledge of tasks and patterns of interaction associated with high and low engagement to make inferences about the quality of student engagement in the classrooms we observed.

According to the High School Survey of Student Engagement (HSSSE), teacher lecture is the least preferred type of teaching, with 44.2 percent of students liking teacher lecture "not at all," while only 6 percent like it "very much"[26] (Yazzie-Mintz and McCormick 2012). In our study, teachers in comparison classrooms lectured, demonstrated and explained more than three times as often as teachers in RA classrooms.

The HSSSE also suggests that students in RA classrooms found tasks more exciting and engaging. According to HSSSE results, boredom was most frequently related to uninteresting or irrelevant material (81.3 and 41.6 percent). A third of HSSSE respondents also related boredom in class to lack of challenge, while only a quarter reported that they were bored because "work was too difficult." To the extent that it is challenging and relevant, argumentation would likely generate greater engagement than less challenging and relevant tasks. In fact, over 28 percent of students taking the HSSSE reported that discussion and debate excited and/or engaged them "very much," while 15.8 percent reported liking these activities "not very much." RA teachers in our study allocated more than twice as much time to argumentation tasks than teachers in Non-RA classrooms. In addition, RA teachers allocated much more time to argumentation-related tasks—close reading and cross textual analysis. Non-RA teachers allocated nearly four times the percentage of class time to fact acquisition tasks compared with RA teachers (26 versus 7 percent). Fact acquisition tasks are characterized by little or no opportunity for sense-making—likely contributing to both low challenge and low interest.

The HSSSE also assessed the degree to which individual reading excited or engaged students. Over 33 percent of HSSSE respondents reported liking individual reading "not at all," while only 10.4 percent reported liking it "very much." While teachers in our study allocated a relatively small percentage of time to individual work, when students did work individually, teachers in RA classrooms allocated more than four times as much time to close reading than Non-RA teachers (54 versus 12 percent). Treatment and comparison teachers were closer in the percentage of time they asked individual students to work with text, however (86 versus 68 percent). Thus, while individual work in both RA and Non-RA classrooms frequently involved working with text, RA students had substantially more opportunity for supported close reading than Non-RA students. Interestingly, while more students reported boredom from lack of challenge than from work being too difficult, 35 percent of the 21% of students who considered dropping out did so because "the work was too hard." The very important question remains whether support for close reading might mitigate student dislike for individual reading and the "slow process of disengagement" that contributes to students dropping out of school.

Finally, HSSSE results inform our understanding of the role of social support and student engagement. Engagement on the HSSSE was associated with both teacher and peer interaction. Over a third of HSSSE respondents reported boredom as a result of no interaction with teacher. Our

[26] HSSSE items asked students, "To what degree does each of the following types of work in class excite and/or engage you?"

research found somewhat more teacher modeling, guidance and support in RA classrooms compared with Non-RA classrooms. Interaction with peers was generally associated with high engagement among HSSSE respondents. Nearly a quarter of student respondents on the HSSSE reported that group projects excited and/or engaged them "very much," while 16.4 percent reported that they liked this type of peer interaction "not at all." Students in RA classrooms spent considerably more time in partnerships and small groups than students in Non-RA classrooms, suggesting more frequent excitement or engagement among RA students. However, the substantial percentage of HSSSE respondents who did not like group projects suggests that group work *per se* does not guarantee high engagement and begs the question of what features of tasks and classroom climate might mediate the relationship between peer interaction and student engagement.

These issues of engagement may have special import for middle school students. As mentioned earlier, in Non-RA classrooms, teacher lecture, demonstration and explanation and fact acquisition activities were highest and peer-directed group work lowest in middle school. Middle school RA teachers lectured less than a third of the time, and middle school RA students received considerably more teacher support and spent more than twice as much time in peer-directed groups than their Non-RA counterparts. Given findings from the HSSSE survey, the limited challenge, autonomy and support characteristic of middle school teaching and reflected in findings from Non-RA middle school classrooms may place students in Non-RA classrooms at greater risk than students in RA classrooms of the "slow process of disengagement" that potentially ends in dropping out of school (Bridgeland, DiIulio and Morison (2006).

*Building blocks of argumentation*. While some research suggests that disciplinary knowledge acts primarily as a constraint on argumentation (von Aufschnaiter, et al., 2008), other studies have found that argumentation potentially builds content knowledge (Reisman, 2011). One difference between these studies is the role of reading in the argumentation task. Additional research is needed to elucidate the relationship between argumentation, close reading and content knowledge. Our preliminary analysis suggests that argumentation builds disciplinary knowledge when close reading and rereading of texts/data is a feature of the argumentation task. Additional research is needed to explore features of argumentation tasks, text use, instruction and classroom culture that contribute to reciprocity among these three elements of literacy—reading, argumentation and disciplinary knowledge—and that support and undermine student investment in sense-making with texts as a foundation for evidence based argumentation.

*Instantiation of argumentation*. Previous analysis of the classroom observation data and extant literature suggests that particular features of argumentation tasks themselves may affect engagement and learning. In our first pass analysis of classroom observation video data reported here, we did not identify subcategories of argumentation tasks. Fine grained analysis is needed to identify features of argumentation tasks—e.g., nascent/interactive argumentation where students engage in reason-giving thinking as a byproduct of negotiating meaning, or more formal generation of claims and evidence (Cavagnetto, 2010); complexity of "argument context" (Berland & McNeill, 2009); and whether the question that framed the argument was generated by the teacher, or arose from students' own reading and inquiry (Fulkerson, 1996; Hillocks, 2010); etc. –to consider how the acquisition of argumentative literacy is mediated by task features and purposes. Similarly, additional research is needed to examine groupings, teacher-student and student-student interactions associated with argumentation

tasks to explore how argumentation is mediated by social support and other social and linguistic factors.

*Multiple sources*. Our preliminary analysis corroborated findings from text processing and disciplinary reading research suggesting that learning from multiple texts requires reading and thinking processes beyond those required to comprehend single texts (Goldman, 2009; Wineburg, 1994). In that analysis, we found that students who successfully comprehended single texts often floundered when the task required them to synthesize multiple sources. At the same time, our data suggested that teachers largely failed to appreciate the additional demands on readers to process, analyze, evaluate and synthesize material from multiple sources. Although we documented tasks that drew on multiple sources, support for reading was largely focused on comprehending single texts. When teachers did provide support for learning from multiple sources, this tended to be in the form of a common notetaker for texts read sequentially. In other words, tasks and instruction framed cross textual analysis as a product of close reading of single texts that did not require additional explicit support. The current analysis of video data added to the complexity of teaching and learning E-BA with multiple sources by identifying tasks that asked student to engage in cross textual analysis absent close reading. For example, in Non-Reading Apprenticeship history classrooms, teachers allocated 16 percent of class time to cross textual analysis but only four percent of class time to close reading. In these cases it appears that teachers bypassed the step of having students comprehend single sources as a prerequisite for synthesizing material from multiple sources. Additional research is needed to understand how to best support evidence-based argumentation from multiple sources, including sequencing, instructional support and social support for reading, analyzing, evaluating and synthesizing multiple sources.

*Methodological issues*. This study also has implications for methods of coding and analysis of video data. In a study of international mathematics instruction, Andrews (2008) advocated a generic episodes approach to coding as "as a simple yet effective framework for describing and analyzing lessons." In this approach, the analytic unit—an episode—is a segment of any length in which the teacher's didactic intent remains constant. Researchers identify each code as either present or absent in an episode, and the analysis is based on the percentage of episodes in which a code is observed. Yet our analysis suggests that *duration* of tasks, activities and interactions is a more sensitive measure of classroom activity than *occurrence* and better able to distinguish between disciplines, grade bands and treatment groups.

Finally, our findings have implications for how to promote the kind of high level literacy instruction envisioned by current educational reforms such as the Common Core State Standards (2010). Our revised analysis of differences between Reading Apprenticeship and Non-Reading Apprenticeship lessons found pervasive treatment group differences in student opportunity to learn associated with long term professional development focused on supporting disciplinary reading and inquiry. This suggests that the kind of high level literacy instruction for all students envisioned by current educational reforms is malleable and that high quality professional development focused on supporting students to engage in disciplinary reading and inquiry may help lead the way in reshaping curriculum and instruction. Additional research is needed to determine how to best support teachers in providing the kind of literacy instruction envisioned in the standards to all students.

**References**


Berland, L. K. & McNeill, K. (2009). Using a learning progression to inform scientific argumentation in talk and writing. Paper presented at the Learning Progressions in Science (LeaPS) Conference, June 2009, Iowa City, IA

Bridgeland, J.M., DiIulio, J.J., and Morison,K.B. (2006). The silent epidemic: Perspectives of high school dropouts. Washington, D.C.: Civic Enterprises.

Cavagnetto , A. R. (2010). Argument to Foster Scientific Literacy: A Review of Argument Interventions in K −12 Science Contexts. *Review of Educational Research, 80*, 336-371.

Chinn & Anderson, 1998. The structures of discussion that support reasoning. Teachers College Record Volume 100, Number 2, Winter 1998, pp. 315-368.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37-46.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company. Farrington et al. (2013).

Farrington, C.A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T.S., Johnson, D.W., & Beechum, N.O. (2012). Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review. Chicago: University of Chicago Consortium on Chicago School Research.

Fulkerson, R. (1996). The Toulmin model of argument and the teaching of composition. In Emmel, B., Resch, P. & Tenney, D. (Eds.), *Argument revisited, argument refined*. Thousand Oaks, CA: Sage Publishing, 45-72.

Hillocks, G. (2010). Teaching argumentation for critical thinking and writing. *English Journal,* 99 (6), 24–32.

Jimenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics, *Science Education, 84*, (6), 689–799.

Landis JR, Koch GG. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-74.

Melchers, K. G., Lienhardt, N., Von Aarburg, A., and Kleinmann, M. (2011). Is more structure really better? A comparison of frame-of-reference training and descriptively anchored rating scales to improve interviewers' rating quality, *Personnel Psychology, 64*, 53–87.

Miles, m.B., & Huberman, A. M. (1994). *Qualitative Data Analysis* (2nd Ed.). Thousand Oaks, CA: Sage Publishing.

Moje, E.B. (2007). Developing socially just subject-matter instruction: A review of the literature on disciplinary literacy teaching. *Review of Research in Education, 31,* 1-44.

Norris & Phillips (2003). How literacy in its fundamental sense is central to scientific literacy. Science Education, 87, (2), 224–240.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Authors.

Nystrand, M., Gamoran, A., & National Center on Effective Secondary Schools, M. I. (1990). Student Engagement: When Recitation Becomes Conversation.

Reisman, A. (2011). *Reading like a historian*: *a document-based history curriculum intervention in urban high schools*. Dissertation submitted to the School of Education. Thesis (Ph.D.)--Stanford University, 2011. Online at: http://purl.stanford.edu/vv771bw4976

Shanahan, T. & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, *78*(1), 40-59.

Shanahan, T., Shanahan, C., & Misischia, C. (2011). Analysis of expert readers in three disciplines: History, mathematics and shemistry. Journal of Literacy Research, 43 (4), 393-429.

Viera, A. J. & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, *37*( 5), 360-363.