Deane, Paul; Song, Yi

A case study in principled assessment design: Designing assessments to measure and
support the development of argumentative reading and writing skills

Psicología Educativa

A case study in principled assessment design: Designing assessments to measure and support the development of argumentative reading and writing skills

Paul Deane and Yi Song

Educational Testing Service, Princeton, New Jersey, U.S.A.

ABSTRACT

This paper presents a principled approach to assessment design in which major design decisions are structured to support teaching and learning. This approach, developed as part of a long-term research initiative at ETS, Cognitively Based Assessments *of*, *for* and *as* Learning (CBAL), draws upon the learning and cognitive science literatures to create richly-structured assessments that simultaneously measure critical component skills and model effective strategies for applying those skills to complex performance tasks. To illustrate our approach, we focus on an important literacy practice: argumentation. Our model seeks to measure qualitative shifts in the development of critical argumentation skills by postulating argumentation learning progressions informed by the developmental literature. These learning progressions play a critical role in guiding assessment design decisions (selecting targeted skills, developing items to measure those skills, and determining task sequences) and may have the potential to support teachers' instructional decisions that effectively scaffold the development of students' argumentation skills.
© 2014 Colegio Oficial de Psicólogos de Madrid. Production by Elsevier España, S.L. All rights reserved.

Un estudio de casos en el diseño de la evaluación centrada en principios: diseño de evaluaciones para medir e impulsar el desarrollo de la argumentación en habilidades de lecto-escritura

RESUMEN

Este artículo aborda la cuestión de cómo diseñar de manera fundamentada una evaluación donde las principales decisiones se toman con el fin de apoyar el proceso de enseñanza-aprendizaje. Este trabajo ha sido desarrollado como parte de un extenso proyecto de investigación en el ETS –evaluación cognitiva *de*, *por* y *para* el aprendizaje (CBAL en su acrónimo inglés)– y se nutre de la literatura previa sobre cognición y aprendizaje para crear evaluaciones con una estructura muy elaborada que, de forma simultánea, miden habilidades críticas y modelan estrategias eficaces para aplicar esas habilidades a tareas complejas de resolver. Para ilustrar este marco de trabajo, nos centramos en una importante práctica relacionada con la lectura y la escritura: la argumentación. Nuestro modelo trata de medir cambios cualitativos en el desarrollo de habilidades críticas de argumentación, postulando una progresión de aprendizaje para la argumentación tomada de la literatura especializada. Las progresiones de aprendizaje juegan un papel decisivo a la hora de tomar decisiones relativas al diseño de la evaluación (seleccionar las habilidades básicas, elaborar preguntas para medir esas habilidades y determinar la secuencia de las tareas) y pueden también contribuir a que los profesores tomen decisiones relativas a la instrucción que sirvan para estructurar de forma efectiva el desarrollo de la capacidad de argumentar de sus estudiantes.
© 2014 Colegio Oficial de Psicólogos de Madrid. Producido por Elsevier España, S.L. Todos los derechos reservados.

*Correspondence concerning this article should be addressed to Paul Deane.
Educational Testing Service. 660 Rosedale Road, MS 11-R. Princeton, NJ 08541 USA.
E-mail: PDeane@ets.org

People use arguments on a daily basis to accomplish many purposes, including persuasion, negotiation, debate, consultation, and resolving differences of opinion (van Eemeren, Grootendorst, & Henkemans, 1996; Walton, 1992). For example, citizens argue about proposed policies, weighing their benefits and drawbacks; scientists advance hypotheses, support them with experimental evidence, and address alternative hypotheses; students participate in classroom debates about the interpretation of literature. Argumentation plays a critical role in the development of critical thinking and in developing a deep understanding of complex issues and ideas. To become successful professionals and members of a democratic society, students must learn to use arguments appropriately and effectively.

Despite the importance of argumentation, the U.S. educational system does not appear particularly effective at developing the ability to produce or critically evaluate arguments, as evidenced by a variety of large-scale assessments and empirical studies. The literature indicates that even college students and adults may have difficulty recognizing argumentative text structures while reading (Chambliss, 1995; Larson, Britt, & Larson, 2004). Nor are U.S. students particularly effective at producing written arguments (Ferretti, MacArthur, & Dowdy, 2000; Nation's Center for Education Statistics, 2008; Persky, Daane, & Jin, 2003). Students often fail to include critical argumentative elements (thesis, reasons, evidence) or do not present them clearly; supporting evidence may not be sufficiently developed, and students may not recognize or respond to alternative viewpoints (Ferretti et al., 2000).

Moreover, lack of argumentation skill creates major obstacles for students working toward college and career readiness. College level reading material often includes multiple sources that present conflicting ideas and arguments, and most college writing falls within an intellectual tradition of rational discourse: claim, evidence, consideration and rebuttal of potential criticisms, and conclusion, all intended to appeal to the reader's reasonable judgment. Writing in different disciplines may vary in tone and content, but the basic argumentative framework persists across a wide range of academic genres (Butler & Britt, 2011). Consequently, the new Common Core State Standards (CCSS)[1], adopted by more than 40 U.S. states, explicitly emphasize argumentation, especially the skills of building logical arguments and using relevant evidence. Argumentation is a key strand in both the CCSS reading and writing standards (Council of Chief State School Officers & National Governors Association, 2010).

Argumentation is one of the most complex skills taught in school, but has often not been well-supported in U.S. educational practice, which has frequently emphasized basic composition and specific formal templates such as the five-paragraph essay while doing relatively little to develop argumentation and critical thinking (Hillocks, 2002). Furthermore, traditional assessments of argument writing, which typically require students to write an on-demand essay on a single prompt with no source material, may constrain students' ability to develop good arguments because this type of writing task does not provide sufficient background information. Knowledge about the topic, stored in long-term memory, is a key element in effective writing processes (Hayes & Flower, 1980), and thus, shallow knowledge can lead to ineffective argumentation. Perhaps even more importantly, poor performance on an essay test provides relatively little information about *why* students failed to produce strong arguments. Ideally, an assessment of argumentation skill would provide useful information on which an argument could be built, would assess both argument comprehension and argument production, and would be structured to support more effective educational interventions.

Using assessment to inform instruction and learning is a key goal of the CBAL ("Cognitively-Based Assessments *of*, *for*, and *as* Learning") research initiative at Educational Testing Service (Bennett, 2010). CBAL tests are designed, as much as possible, to capture useful information about what students know and can do (assessment *of*

learning), model effective practice (assessment *as* learning) while providing actionable information for teachers (assessment *for* learning). CBAL seeks to build an assessment system that helps teachers make sound educational decisions towards enhancing their students' knowledge and skills. For example, the link between assessment and instruction could be created through assessment tasks corresponding to instructionally appropriate activities that teachers could use to teach the targeted skills or through a series of lead-in tasks that require strategies that teachers could teach to ensure that students perform the final, integrated task successfully.

When viewed in this light, the problem of designing an effective assessment of argumentation becomes an instance of a more general problem: the problem of principled assessment design in which due consideration is given to underlying cognitive processes and the impact of test design on learning and instruction. We approach this problem from the perspective of evidence-centered design (Mislevy, Steinberg, & Almond, 2003; also see Michael Zieky's paper in this volume), a method that builds explicit validity arguments that link test design decisions to inferences about student skill. In particular, we adopt the perspective outlined in Deane (2011), which provides a domain analysis for the English Language Arts – one essential aspect of Evidence-Centered Design under the CBAL research initiative.

One of the key ideas underlying the CBAL approach to assessment development is a focus on *scenario-based assessment*, in which different parts of a test are used not only to assess key skills but to model important steps that a skilled practitioner would follow. O'Reilly and Sheehan (2009), Sheehan and O'Reilly (2011), and O'Reilly and Sabatini (in press) develop arguments motivating the use of scenario-based assessments in the English Language Arts. Essentially, scenario-based assessments are design to combine advantages characteristic of traditional assessment designs (multiple items providing reliable, independent measurement) with advantages characteristic of simulations and performance assessments (such as increasing authenticity, encouraging student engagement, and modeling effective practice).

Another critical element (outlined in Deane, 2011, and made publicly available in draft form at http://elalp.cbalwiki.ets.org/) is an attempt to specify *learning progressions* that can help to inform both instruction and assessment. A CBAL assessment uses principles of evidence-centered design to select items that illustrate how key skills are connected in expert practice, while simultaneously measuring specific levels on targeted learning progressions. To the extent that it succeeds, it provides information about student performance that will help teachers identify what students need to learn to progress to the next level on targeted skills.

As part of the CBAL ELA competency model, we have developed a framework for analyzing argumentation that defines typical argumentation scenarios and identifies the major skills needed at each stage in the process of understanding, building and presenting arguments. In the first part of this paper, we present this framework and describe learning progressions intended to capture major developmental patterns observed in the literature. In the second part of this paper, we present a scenario-based assessment design that focuses on key argument skills. This design is explicitly linked both to a general model of argumentation as a social practice and to argumentation learning progressions designed to measure key argumentation skills.

Domain Analysis: Phases of Argument, Types of Argumentation Skills

Argumentation is best conceived as a rule-governed form of discussion in which various speech acts, including assertions, questions, and explanations, are coordinated in the service of social norms for reasoned discussion, or dialectic (van Eemeren & Grootendorst, 1992). While written arguments may appear to be

independent texts, they function as part of a larger social context in which discussion is distributed across multiple publications in a written discourse. However, in pragmatic terms and in the oral contexts from which argumentation typically emerges, argumentation is a dialogue in which participants may take many different positions and change their minds as it proceeds, rather than as a purely logical exercise. Newell et al. (2011) argue that appropriate social practices that model effective argumentation are critical to the development of argument skill, and suggest that teacher training may have a significant impact on whether students are able to internalize the appropriate values and practices. Reznitskaya et al. (2001) report that such interventions do, in fact, appear to result in more sophisticated argumentation from students, i.e., essays containing more relevant arguments, and more use of sophisticated elements such as counterarguments, rebuttals, and textual citations.

In order to successfully engage in such a dialogue, skilled participants undertake several different types of activity, which can be analyzed as an argumentation cycle with five distinguishable phases (see Figure 1). Each phase focuses on a different set of problems and goals, and therefore requires somewhat different kinds of reading, writing and thinking skills, as represented by the questions associated with each phase in Figure 1. These five phases are outlined below.

1) Understanding the stakes. Effective argument requires the participant to consider the context and the target audience (including their beliefs and values) in order to make effective appeals to persuade the audience. We use the term *appeal building* to describe the skillset required in this phase. Appeal building is a form of social reasoning that focuses on how people are moved to belief and action. In classical terms, appeal building addresses rhetoric rather than logic, and deals with classical rhetorical appeals (Aristotle, trans. 1939): *ethos* (appeals to authority), *pathos* (appeals to emotion), and *logos* (appeals to logic). Effective argumentation requires sensitivity to these issues, both to evaluate arguments presented by others, or to develop arguments of one's own. In a pedagogical setting, a failure to consider rhetorical

issues, including the intended audience, can be problematic. Instead of thinking through who should be the intended audience and employing appropriate rhetorical strategies, students often view their teacher as the only audience simply because their goal is to complete the assigned tasks (Bright, 1995). Lack of authentic audience can prevent students from developing a sense of audience awareness.

2) Exploring the subject. Effective argument requires the participant to have a deep understanding of the topic, which typically develops in tandem with the development of effective arguments about important issues to which the subject is relevant. The unique skillset required during this phase focuses on research and inquiry (e.g., posing important questions, identifying useful sources and evaluating their credibility, and synthesizing information across sources that addresses the research questions). In our analysis, exploring the subject draws upon more general literacy skills that are not unique to argumentation.

3) Considering positions. Well-reasoned argument requires an open mind, in which people examine many possible points of view, thinking through which positions make the most sense and can be easily defended, rather than trying to justify a predetermined position. As a practice, argument depends on confrontations between contrasting points of view (van Eemeren et al., 1996). From an early age, most people can effectively express their own beliefs and opinions. It is far more difficult to understand and properly consider alternative positions, with typical consequences for the kinds of arguments students typically produce in a school context. Various studies indicate that student essays are strongly biased in that they only consider one side of the issue (Knudson, 1992; Leitão, 2003; McCann, 1989; Perkins, Farady, & Bushey, 1991). Part of this deficit may be due to the transition from oral interaction (where alternative viewpoints are represented by other participants) to written interaction (where the reader and writer must simulate multiple viewpoints without direct feedback or interaction). We use the term *taking a position* to name this set of skills, although it is important to remember that skilled practitioners take a position informed by thoughtful consideration of alternative perspectives.
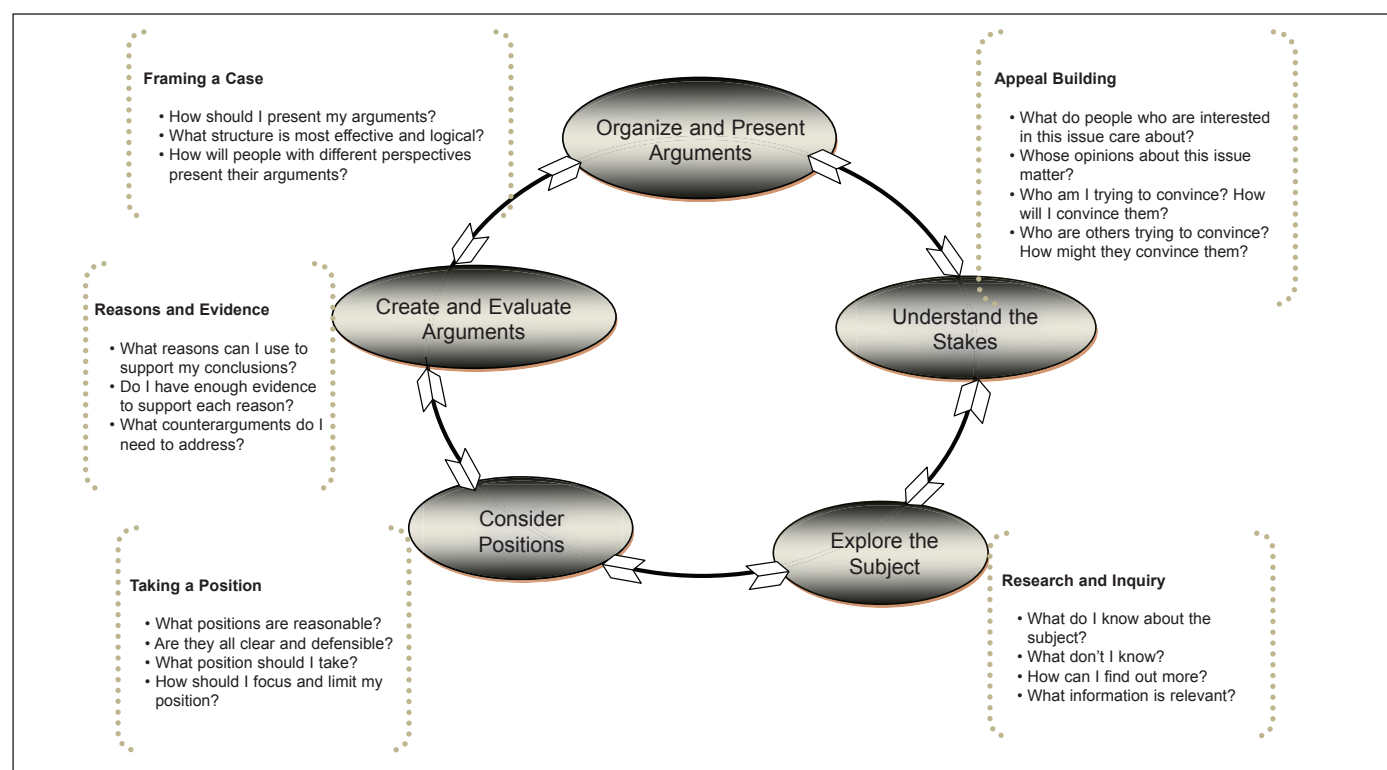

Figure 1. Five phases of participating in argumentative discourse.

4) Creating and evaluating arguments. Effective argument is logically valid and evidentially sound. Effective arguments have been supported with appropriate reasons and evidence, taking counterarguments into account (and where necessary, refuting them.) This kind of skill – for which we will use the term *reasons and evidence* – is the focus of the CCSS argumentation standards, and involves both the ability to create arguments and the capacities necessary to evaluate and critique them (such as identifying logical fallacies). It is a relatively late-developing skill, unless effectively supported through social interactions that develop appropriate social norms and epistemic understandings (Kuhn & Udell, 2003; Kuhn, Zilmer, Crowell, & Zavala, 2013).

5) Organizing and presenting arguments. An effective writers know how to reconstruct others' arguments from textual cues, and can structure and present their own argument in an effective form and format. Students should follow a set of genre conventions that govern how argumentative discourse is organized, in both informal conversation and in written texts. This set of the skills is the ability to *frame a case*. For example, many students learn to write a five-paragraph essay, which is a form of presenting one's arguments. A number of self-regulated strategy development (SRSD) studies have focused on teaching students various strategies for framing a case in argumentative writing (De La Paz & Graham, 1997; Graham & Harris, 1989; Sexton, Harris, & Graham, 1998). Lack of development of other aspects of argument skill, such as reasons and evidence, will complicate instruction focused on organizing and presenting arguments. All writing instruction must provide students appropriate criteria for evaluation (Bracewell, Scardamalia, & Bereiter, 1978; Flower, Hayes, Carey, Schriver, & Stratman, 1986) and requires well-structured protocols for peer interaction (DiPardo & Freedman, 1988; Kinsler, 1990), but in the context of argumentation, the dialogic nature of argument raises the stakes.

Note that in our conception, the five phases need not occur in the order we set them out in the diagram, since the actual process of building an argument is flexible and fluid. People can and do start at different points in the argumentation cycle, and may move back and forth between processes. They may even repeat steps when needed. However, at a high level of abstraction, the five phases capture key processes engaged in by skilled arguers, and represent appropriate goals for instruction and assessment. Note also that the mapping from phases of argument to the skills they require is not entirely one-to-one. Each phase requires some combination of the skills; however, certain skills (such as reasons and evidence) are most salient in one phase (creating and evaluating arguments), and less salient in others (such as understanding the stakes). But an analysis of the phases helps identify skills specially needed to support each phase.

This framework captures the fact that effective argumentation presupposes a combination of social, conceptual, and discourse skills (Deane, 2011). Skilled arguers consider the social context of argument, including the interests and motivations of an audience, build up their knowledge about an issue, anticipate alternative perspectives, support a well-thought-out position by providing appropriate reasons and evidence, and structure their argument to communicate it effectively.

Towards an Evidence Model: Defining Argumentation Learning Progressions

In order to design an assessment that can provide evidence about student acquisition of argumentation skills, it is important to be able to define a series of well-defined evidence targets that we can set as measurement goals. This measurement framework needs to be consistent with the results of empirical investigations into the development of argumentation skills in oral and written contexts.

The general trend of development resembles that for many literacy skills: it begins with context-bound, informal skills, for example, interpersonal argumentation with a familiar addressee (Clark & Delia, 1976; Eisenberg & Garvey, 1981), and it gradually increases with age and moves into less and less familiar contexts (Stein & Miller, 1993). However, many argumentation skills develop before adulthood only for individuals at the upper end of the verbal proficiency scale, with novices often doing little more than expressing a position then elaborating a few supporting reasons (e.g., Kuhn, Katz, & Dean, 2004). The following empirical studies provide important insights about the development of argumentation skills.

It has been found that young children can express their opinions and offer supporting reasons (McCann, 1989), and upper-elementary school students are able to elaborate and provide details in support of their arguments (Ferretti et al., 2000; Ferretti, Lewis, & Andrews-Weckerly, 2009). A more complex skill is the use of evidence, for example, distinguishing evidence and explanation, selecting relevant evidence to support one's arguments, and explaining how the evidence supports an argument (Brem & Rips, 2000; Kuhn, 1991; Kuhn, Shaw, & Felton, 1997). Some argumentation skills are challenging to students and might not develop before adulthood unless instruction or scaffolding is provided. For example, even high school or college students find it difficult to analyze the assumptions behind people's arguments or present arguments from both sides of an issue (Klaczynski, 2000; Kuhn, 1991). When students present counterarguments with weak rebuttals, their standpoint would be undermined, making their essays unpersuasive, but refuting opposing viewpoints is not an easy task, especially in a written context (Ferretti et al., 2000; Knudson, 1992; Leitão, 2003; Nussbaum & Kardash, 2005).

Felton and Kuhn (2001) examined the argumentative discourse of young teens and community college students on the issue of capital punishment. Both groups offered arguments to support their own position, but compared to teens, the young adults demonstrated more sophisticated skills in refuting the opponent's arguments. Their finding is consistent with other studies that students often do not realize that considering and rebutting an opposing side can increase the persuasiveness of their arguments and they do not have strategies to do this successfully. According to Kuhn et al. (2004), the cognitive capacities that support argumentation, involving metacognitive and metalinguistic awareness, may mature relatively late. Some students might have a desire to maintain cognitive consistency, especially when they are overwhelmed by high cognitive load in writing where cues from oral discourse are missing (Coirier, Andriessen, & Chanquoy, 1999; Scardamalia & Bereiter, 1986; Simon & Holyoak, 2002).

We seek to capture the insights of these and related developmental studies by defining *learning progressions*, a series of formal statements that make expected developmental sequences explicit and define the kinds of evidence that would support the claim that a student's achievement is at an earlier or later stage of development. Designing assessment tasks aligned to the learning progressions can support evidence-based inference about student achievement levels, which can be further developed to recommend classroom practices that scaffold students toward the next level of performance.

The concept "learning progression" was first introduced to describe possible levels in student development of a specific skill or concept in the context of science assessment (e.g., see Corcoran, Mosher, & Rogat, 2009, p. 37), but has now been extended to a variety of other domains. The definitions of learning progressions held by researchers vary slightly (Heritage, 2008). The CBAL research initiative emphasizes the idea that learning progressions are qualitative shifts in performance that indicate when students are ready to move on to more challenging tasks within a general skill, and formally defines a learning progression as a description of "qualitative change in a student's level of sophistication for a key concept, process, strategy, practice, or habit of mind" (see the CBAL

Wiki at http://elalp.cbalwiki.ets.org/Outline+of+Provisional+Learning+Progressions).

We have developed a learning progression for each of the four skill types unique to argumentation: *appeal building, taking a position, reasons and evidence*, and *framing a case*. Table 1 presents a general overview of the learning progressions across these four strands. Each learning progression has five developmental levels: for which we have adopted the convenience levels *preliminary, foundational, basic, intermediate,* and *advanced*.[2] A rough mapping to expected grade levels is provided to indicate how the learning progression levels relate to typical patterns of student development. However, since students' skills and prior knowledge vary widely within a grade, these mappings should be interpreted only as an indication of how the expectations for performance typically increase by age under current U.S. educational practice.

In our detailed analysis of the development of argumentation skills, we distinguish three types of cognitive process: interpretation (reading), expression (writing), and deliberation (thinking and strategy use). Our goal is to capture parallels in the development of related skills, since the ability (for instance) to recognize a flaw in an argument is going to be closely connected to the ability to write critiques, or to apply specific strategies to facilitate critical analysis of argument validity. These are expressed by parallel columns in our detailed learning progression descriptions. The detailed learning progressions also indicate not only what we believe students are able to accomplish at each level, but also specify limitations, that is, features that help define how each successive level of development falls short of the next. Table 2 presents an example of this detailed analysis for the Reasons and Evidence strand. In general, the qualitative shifts in the CBAL argumentation learning progressions correspond to developmental milestones, but are often more specific than we can support from the existing literature, reflecting hypotheses extrapolated from the literature. As a result, the CBAL argumentation learning progressions must be viewed as provisional models, subject to empirical verification.

Our intention is to use the CBAL argumentation framework and learning progressions to support assessment design and instruction. How we do this is the focus of the section that follows.

Scenario-Based Assessment Design

In the preceding sections, we introduced two related ideas: *phases of argumentation* (part of a general description of argumentation as a practice important to college and career readiness) and *argumentation learning progressions* (a schematic representation of measurable levels of student growth in argumentation skill). These ideas can be combined to create scenario-based assessment designs. In particular, we will discuss the considerations that govern our design for several 8th-grade assessments of argumentative reading and writing.

As a first step, we map the general model (the five phases of argumentation) onto a specific scenario type, and define what we intend to measure. Table 3 summarizes one such mapping, reflecting a series of decisions about the relationship between the argumentation scenario (and thus the phases of argumentation) and measurement goals. It reflects a series of specific decisions and priorities, which must be given due consideration.

Unlike a typical direct writing assessment, the design includes significant reading work focused on sources that present multiple positions and provide background information about the topic. This integrated literacy approach is grounded in the research on reading and writing connection. On the one hand, skillful writing builds upon prerequisite reading skills (Newell, Beach, Smith, & VanDerHeide, 2011), and on the other hand, writing about texts can enhance students' reading comprehension (Graham & Hebert, 2010). Literacy practices in the classroom also involve both reading source texts, and writing one's interpretations or arguments of the texts. Reading argumentative texts may be an effective way to learn to identify key argumentative elements, including thesis statements, supporting reasons, evidence, and counterpoints. In addition, students are likely to write strong and meaningful arguments about issues they

Table 1
Overview of Argumentation Learning Progressions

| | Social | Conceptual | | Discourse |
	Appeal-building	Taking a position	Reasons and evidence	Framing a case
Preliminary K-2	Understands the idea of trying to convince someone by making some sort of persuasive appeal.	Understands the idea of taking a side in an argument and accepting or rejecting another person's statements as true or false based on how well one thinks it fits the facts.	Understands the idea that positions may need to be supported with reasons that will be convincing to the audience.	Approaches argument as chain of individual turns, and understands and produces such turns in context, such as taking a position or giving a reason.
Foundational Grades 3-6	Transfers the idea of making a persuasive appeal into a written context and does some simple analysis of how oneself or an author might appeal or has appealed to different audiences and interests.	Understands and expresses positions in writing with reasonable attention to what one knows and some ability to focus on what is important in the domain.	Recognizes, generates and elaborates on reasons in writing, with some awareness of the need for evidence, and uses one's own argument to counter others' argument in an engaging, familiar context.	Approaches persuasive text as a coherently organized sequence of reasons supporting a position.
Basic Grades 7-9	Infers rhetorical structure in texts, and builds rhetorical plans of one's own that coordinate multiple appeals and moves into a coherent effort to persuade a target audience.	Understands and expresses positions clearly, capturing their relationships both to similar and contrasting points of view.	Understands use of evidence and clearly grasps the need to provide evidence and reasons that are directly relevant to and support the main point and which are logically sound.	Approaches persuasive text as a logically structured presentation of a case with embedded reasons and evidence.
Intermediate Grades 10-12	Shows flexibility in interpreting and developing rhetorical plans, with sensitivity to differences among audiences with different points of view.	Successfully analyzes unstated assumptions, biases, and other subjective elements in a text and can use that to develop one's own position more clearly.	Understands the role of critique and rebuttal and is able to reason about and respond to counterevidence and critical questions.	Approaches persuasive text as part of a dialog between multiple perspectives with appropriate attention to counterpoint and rebuttal.
Advanced College	Displays a well-developed rhetorical (metacognitive) understanding of persuasion.	Can use others' arguments to develop one's own understanding, and then frame one's own position in terms that exploit the current "state of discussion".	Builds systematic mental models of entire debates, and use that model to frame one's own attempts to build knowledge.	Displays mastery of many different forms of argument, demonstrating flexible understanding and control of genre features.

Table 2
Development of Reasons and Evidence (Levels 1-4)

Level	Interpretation		Expression		Deliberation	
	Achievement	Limitation	Achievement	Limitation	Achievement	Limitation
Preliminary K-2	Reason-1-I-A Identifies reasons people give to support a specific point.	Reason -1-I-L Is limited to atomistic comprehension, showing little global understanding of the overall case.	Reason-1-E-A Generates at least one reason to support a specific point, in sentence form.	Reason-1-E-L May display an unselective and additive approach, generating arguments based primarily on most obvious points of attack.	Reason-1-D-A Can apply template-based argument-generation strategies (such as making lists or filling in a pro-con chart).	Reason-1-D-L May deploy only a narrow range of argument schemes, thus limiting the quality of generated arguments.
Foundational Grades 3-6	Reason-2-I-A Identifies supporting reasons or evidence in a written text and relates them to the point they support.	Reason-2-I-L May have only a minimal understanding of evidence; anything that seems loosely relevant to the point may be viewed as providing support.	Reason-2-E-A Generates multiple reasons to support a point, and uses these reasons to counter others' argument in an engaging, familiar context.	Reason-2-E-L May have only a limited ability to evaluate or provide supporting evidence, exercising it only in response to specific questions, instructions, or objections.	Reason-2-D-A Can apply analytical strategies to identify information needed to support a point, reflecting implicit understanding of common argument schemes).	Reason-2-D-L May have an entirely implicit understanding of argument, with little meta-cognitive control over the process of building or evaluating an argument.
Basic Grades 7-9	Reason-3-I-A Recognizes and explains the relationship between main and supporting points and keeps track of which evidence supports which point.	Reason-3-I-L May not see flaws in elaborated arguments and so find them overly plausible.	Reason--3-E-A Builds logical, hierarchically structured arguments by selecting and arranging reasons and evidence to support main and subsidiary points.	Reason-3-E-L May fail to recognize critical questions which must be addressed in order to make a reasonable argument.	Reason-3-D-A Can evaluate the strength of evidence and distinguish sound and unsound arguments by recognizing common syllogisms and fallacies.	Reason-3-D-L May significantly overestimate the strength of evidence and arguments with which one agrees.
Intermediate Grades 10-12	Reason-4-I-A Identifies specific points in a text that are vulnerable to objections and counterarguments.	Reason-4-I-L May not be adept at selecting the most important questions and objections to address.	Reason-4-E-A Writes simple critiques or rebuttals that critically provide summaries of or responses to other people's arguments.	Reason-4-E-L May miss problems and issues that are highlighted in prior discussions but not in the current situation.	Reason-4-D-A Can apply critical-question strategies for commonly used argumentation schemes to generate counterarguments and determine how to reinforce specific points.	Reason-4-D-L May not have explicit knowledge about argumentation schemes and critical questions to systematically evaluate strength of arguments.

Table 3
An Argumentation Assessment Scenario

Argument phase	Task	Task description	Description of argument skills assessed	Description of other skills assessed
Understand the stakes, explore the subject, consider positions.	1	Evaluate summaries of the argument in an article.	Understand written arguments drawn from source article 1.	Evaluate summary quality.
	2	Write summaries of two more articles.	Understand written arguments drawn from source articles 2 and 3.	Write effective summaries.
Create and evaluate arguments.	3	Evaluate the argument in a letter to the editor.	Recognize faulty argumentation. Write a critique of an argument.	
	4	Classify arguments as for or against a position.	Connect claims with reasons.	
	5	Classify evidence as supporting, weakening or as irrelevant to an argument.	Connect reasons with evidence.	
Organize and present arguments.	6	Write an argument essay	Formulate an effective argument. Present that argument in written form.	Write a coherent, well-organized essay that adheres to standard English conventions.

understand well, which requires some amount of reading beyond a brief statement of the prompt. This decision reflects the importance in our model of the preparatory phases of argumentation – understanding the stakes, exploring the subject, and considering positions.

However, the design is not intended to probe how well students can accomplish each of these early stages independently. Students are provided instead with three source articles to read (one on each side of the issue, and one presenting a balanced view), and are asked to perform summary tasks using these articles as sources. The resulting work provides direct evidence about student summarization skill, and indirect evidence about their ability to understand the stakes, explore the subject, and consider alternate positions using information from multiple source texts.

By contrast, the scenario emphasizes the skills needed to create and evaluate arguments, which are measured directly or indirectly in all the remaining tasks, including the final essay. Instead of assessing

argumentative reading and writing separately, the scenario provides an opportunity to measure not only how skillfully students can read and interpret the arguments presented in text, but also how well they develop their own arguments through writing.

Finally, by modeling effective strategies for combining skills to accomplish a performance task, we hope to facilitate learning in light of Vygotsky's (1978) zone of proximal development. This kind of assessment is intended to be used in populations where many of the individual skills have been mastered, but where the majority of students are not yet able to coordinate subtasks effectively to achieve a more complex integrated performance.

The general framework captured in Table 3 is consistent with a wide range of specific test designs. We flesh this design out further by identifying the specific learning progressions (and levels in those learning progressions) that we intend to measure, and then identifying the specific type(s) of item(s) that will provide the intended evidence. For tasks 3-6 in Table 3, that mapping covers four levels in the argumentation learning progressions, and focuses primarily on one specific progression (*reasons and evidence*), as shown in Table 4. Tasks 1 and 2 are mapped to a different CBAL skill, summarization, which covers more general reading skills needed to be able to read and summarize arguments from sources, and will therefore not be discussed in detail here.

We have developed, piloted, and field-tested multiple assessments built to this design. The companion article in this special issue, written by Peter van Rijn, discusses and evaluates the results. The topics around which individual assessments are built are linked to student life, and are designed to provide an overarching social context within which the sequence of individual assessment tasks will have a clear purpose and function. Example topics focus on questions like the following: Should there be restrictions on advertising to younger children? Should parents limit the amount of time their children spend on social networking sites? Should students be rewarded with money for getting good grades? Should junk food be sold in your school?

Several features of this design are worth emphasizing: (a) the way the scenario is introduced, by providing brief background information and source materials, (b) the four major tasks listed in Table 4 that target argumentation skills of varying levels, and (c) additional techniques used in the assessment design that provide built-in scaffolding, including guidelines, tips, and tools.

The scenario is described to students before the test begins. An opening screen provides a brief overview of the issue about which they are going to have to write, and outlines the tasks they are going to perform. In the first two tasks, students read three short articles that provide information about the issue and the positions and arguments people take on that issue. They read and evaluate other students' summaries of the fist article, and write summaries of the second and third articles. The directions for these tasks are designed to provide an authentic purpose for reading the articles and summarizing their arguments. Although the summarization task is not designed to align to a specific argumentation learning progressions, it plays a critical role in preparing students for the following tasks through setting up the background information, and presenting a variety of arguments on both sides.

In the third task, students critique arguments in a letter to the editor. They must identify and explain problems in the reasoning or use of evidence. For example, in one assessment, the letter to the editor argues that watching advertising is good because families can learn a lot from it and it brings families together. The writer gives an example to support this argument, and the example is a family in his or her neighborhood. Students should point out that the writer is

Table 4

Argument Assessment Design: Mapping to Learning Progression Levels and Item Types

Argument phase	Task	Task description	Specific learning progression	Overall level in the argument learning progression set	Specific descriptor	Items used to measure this skill
Create and evaluate arguments.	3	Evaluate the argument in a letter to the editor.	Reasons and evidence.	4	Writes simple critiques or rebuttals that critically provide summaries of or responses to other people's arguments.	One paragraph-length constructed response item. Stimulus: a letter to the editor that exemplifies several logical fallacies.
	4	Classify arguments as for or against a position.	Taking a position.	1	Infers which side people are taking in an argument based on the reasons or evidence they provide.	Ten binary classifications identifying whether statements drawn from the source articles provide reasons to support the pro or the con side of the issue.
	5	Classify evidence as supporting, weakening or as irrelevant to an argument.	Reasons and evidence.	2	Identifies supporting reasons or evidence in a written text and relates them to the point they support.	Six multiple-choice items in which students decide whether a piece of evidence drawn from the source articles supports, weakens, or is irrelevant to a specific argument.
Organize and present arguments.	6	Write an argument essay.	Reasons and evidence.	3	Builds logical, hierarchically structured arguments by selecting and arranging reasons and evidence to support main and subsidiary points.	One essay-length constructed response item. Stimulus: the three source articles and guidelines for an effective, well-argued essay.

over-generalizing a single observation. We expect that only students who have developed strategies for evaluating arguments can accomplish this task successfully. The task is at the fourth level in our learning progression, a level we believe is typically not addressed pedagogically before high school in U.S. schools. Our field test results suggest that relatively few middle school students can handle this task successfully, but it is part of the test design to measure the highest level of student performance and to communicate the importance of knowing how to evaluate arguments in this kind of argument scenario.

The fourth and fifth tasks focus on assessing important argument reading skills: (1) organizing a list of people's statements by deciding which side each statement supports, and (2) evaluating the relationship between evidence and claims (i.e., whether a piece of evidence supports or weakens a particular claim). Task four is a very simple interpretation task, assessing the skill of inferring which side people are taking in an argument based on the reasons or evidence they provide, which we classify as the earliest level in the argumentation learning progressions. Based on our pilot results to date, if students have difficulty identifying people's positions from the arguments they have made, it is very likely that they will encounter serious problems with the other tasks. Task five targets the skill of identifying supporting reasons or evidence in a written text and relating them to the point they support. Since it focuses on recognizing the relation of evidence to arguments, and understanding whether the evidence supports or weakens the argument, we classify this task slightly higher in the argument learning progression.

After students have completed tasks 1 to 5, they will have had a chance to read and think extensively about the issue. The next logical step is to ask them to write an essay. The essay task necessarily draws upon all aspects of argumentation skill, but focuses on the ability to construct and present logical, hierarchically structured arguments, which corresponds to the third level in the argument learning progression (particularly the basic level descriptors for the progressions dealing with *reasons and evidence* and *framing a case*). This level of performance is a reasonable target for normally-achieving 8th-grade students, though students with more sophisticated argumentation skills may also incorporate alternative perspectives and counterarguments in their writing.

Another important feature in the assessment design is the inclusion of various elements designed to support students and make it easier to obtain their best work when they are asked to perform relatively challenging literacy tasks. For example, in the summarization task, we provide explicit guidelines for summarizing an article, and provide students opportunities to learn and apply the guidelines before they have to write summaries of their own. Similarly, the essay task includes a variety of planning tools (e.g., outline, idea tree, and list) that encourage students to develop a plan for their essays, and a writer's checklist that explicitly details the scoring criteria that will be applied to student essays. Light scaffolding like this may be helpful because many students have difficulty setting appropriate goals in their writing and explicit goals enable students to plan, compose, and revise more effectively (Ferretti et al., 2000; Graham, MacArthur, & Schwartz, 1995; Matsuhashi & Gordon, 1985). The scenario structure, the use of supporting tools, and many of the details of how items are sequenced and presented are designed to work together to create an experience that will help students internalize assessment goals and help teachers interpret test results to support instruction.

Conclusion and Discussion

If literacy assessment is expanded to support additional goals (assessment *for* learning; assessment *as* learning), it has the potential to support innovative approaches to teaching and learning. This paper presents one step in the process necessary to validate

innovative assessments that can be used for such purposes. It illustrates a principled approach to assessment design that draws upon the learning and cognitive science literatures to create richly-structured assessments that simultaneously measure critical component skills and model effective strategies for combining those skills to successfully manage complex performance tasks. We have applied this strategy to the problem of assessing argumentation. Informed by the developmental literature, our model captures not only the critical argumentation skills, but also seeks to describe qualitative shifts in the development of those skills. In particular, argumentation learning progressions play a critical role in guiding the assessment design, and when validated, may be able to support teachers as they make decisions about instruction.

This approach has a number of interesting implications. First of all, the design method helps to link learning sciences research, pedagogy, and assessment in new and interesting ways. For instance, the CBAL argumentation framework identifies five types of argumentation activities in which experts are likely to engage, but when we examine common pedagogical practice, we observe some of these activities may be underemphasized, such as considering alternative points of view (and related skills such as addressing counterarguments). Including such elements in an assessment may help communicate a richer understanding of the construct both to educators and to students.

Second, the assessment design is explicitly structured to identify qualitative shifts in key skills, as measured by levels in the learning progressions. The pattern of performance on individual items is potentially much more interpretable, especially compared with an undifferentiated performance task. Often, for instance, argumentation skills are assessed by having students write an essay, but a holistic score on the essay task may indicate little more than the student's general literacy level. Even though some rubrics provide an analytical score (e.g., organization, reasons, and grammar), the descriptors in the rubric are often relatively vague (for example: "arguments are weak" versus "arguments are somewhat weak" or "arguments are somewhat convincing" versus "arguments are very convincing"). Such differences between scoring levels provide relatively little information about what students have learned and what they need to be taught. The scenario-based assessment design, linked to levels in the learning progressions, has the potential to provide much richer information about the skills students need to acquire to reach higher levels of performance.

Third, this kind of design embodies the idea of integrating reading and writing along lines supported by a large body of educational research (Gorin, O'Reilly, Sabatini, Song, & Deane, 2014; Shanahan, 1990). In a scenario-based assessment design, reading takes place in a context where there is a valid purpose for reading, and writing takes place in a context where students have the opportunity to develop a rich understanding of a topic and are prompted to examine an important issue from multiple perspectives.

Finally, this kind of scenario-based design is an opportunity to model effective strategies for complex, integrated performance. The assessment is designed to illustrate how the skills it measures will be used in effective practice. Such designs are more likely to connect with best instructional practice, especially if they have been developed precisely with the intent of modeling such practices, and if (as in our current work) they are developed in collaboration with teachers and other educational professionals.

Though learning progressions help bridge assessment and instruction, as Shavelson (2009) noted, learning progressions could be vulnerable to data fitting. How successfully we achieve such goals is, of course, an empirical question. Some initial results from field tests of the assessment described in this paper are presented in Peter van Rijn's companion paper in this volume. But in the absence of definitive validation studies, it is important to view the learning progressions that underlie CBAL assessments realistically, as

hypotheses derived from a strong research base. Even though many points are well-established in the research literature, there are gaps in our knowledge that have to be filled in to create a workable framework for assessment design. In other words, the work reported in this paper represents the first stage in a longer-term research program in which patterns of student performance are used to validate both the assessments and the framework underlying individual assessment designs.

Resumen ampliado[3]

Este artículo presenta un sólido marco de trabajo para diseñar una evaluación, teniendo en cuenta los procesos cognitivos subyacentes y el impacto que el diseño del test puede tener en el proceso de enseñanza-aprendizaje. Para ello se propone trabajar con el marco del Diseño Centrado en la Evidencia (DCE) y con una habilidad que juega un papel decisivo en el desarrollo del pensamiento crítico y en la capacidad de llegar al fondo de cuestiones e ideas complejas: la argumentación.

Los autores han desarrollado un marco para analizar la capacidad de argumentar como parte del modelo de competencias de lecto-escritura del proyecto CBAL (acrónimo inglés de 'evaluación cognitiva *de*, *por* y *para* el aprendizaje'). Se trabaja con una evaluación basada en escenarios que pretende combinar las ventajas propias de los diseños tradicionales (un buen número de preguntas que proporcionan medidas fiables e independientes) con las ventajas características de simulaciones y pruebas de evaluación de la actuación (mayor autenticidad, aumento del compromiso del estudiante, modelado de prácticas eficaces). Un segundo elemento crítico del marco de trabajo es que se intenta especificar progresiones de aprendizaje que pueden contribuir tanto a la instrucción como a la evaluación. En el proyecto CBAL se utilizan los principios del DCE para seleccionar preguntas que midan niveles específicos en las correspondientes progresiones de aprendizaje pero que al mismo tiempo muestren cómo se relacionan entre sí determinadas habilidades clave en la práctica de personas expertas. Si se tiene éxito en la empresa, entonces se puede obtener información acerca del rendimiento de los estudiantes que ayudará a los profesores a identificar lo que éstos necesitan aprender para pasar al siguiente nivel en las habilidades consideradas.

En la primera fase del DCE se analiza el dominio de interés, definiendo cinco etapas o actividades en el ciclo de la argumentación: (1) comprender las partes implicadas, esto es, considerar el contexto y la audiencia en cuestión para poder persuadirla con los correspondientes argumentos, (2) examinar el tema para poder llegar al fondo de las cuestiones planteadas, (3) considerar las distintas posiciones, esto es, ser capaz de analizar la cuestión desde distintas perspectivas con una mente abierta, (4) elaborar y evaluar argumentos y (5) organizar y presentar los argumentos en modo y formato apropiados. Estas cinco fases recogen los principales procesos implicados en la argumentación y al mismo tiempo representan objetivos apropiados para la instrucción y para la evaluación. En cada fase hay que movilizar un determinado tipo de habilidades específicas de esta capacidad, con la única excepción de la fase 2 (que requiere habilidades de lecto-escritura más generales, centradas en la investigación y el análisis). En la fase 1 se trata de la capacidad de apelar (a la lógica, a la emoción, a la autoridad), en la fase 3 es la capacidad de adoptar una postura frente a una cuestión, en la fase 4 la habilidad para razonar y establecer evidencias y en la fase 5 para plantear un tema de manera razonada.

Los autores del trabajo han desarrollado una progresión de aprendizaje para estas cuatro habilidades específicas de la capacidad de argumentación, cada una de ellas con cinco niveles de desarrollo que abarcan desde los cursos iniciales de primaria hasta la universidad (véase Tabla 1). Estas progresiones de aprendizaje pueden ser consideradas como modelos provisionales sujetos a comprobación empí-

rica, dado que son más específicas de lo que permite la literatura actualmente disponible.

Seguidamente se presenta un diseño de evaluación basada en escenarios donde se combina la idea de las fases de la argumentación con la progresión en el aprendizaje formulada para esas cuatro habilidades específicas en estudiantes de 8º curso, planteando seis tareas para un mismo escenario de trabajo que recorren las cinco fases con las que el modelo general ha caracterizado la capacidad de argumentar y distintos niveles en la progresión de aprendizaje formulada (véase Tablas 3 y 4). Junto a estas tareas, el diseño de evaluación incluye distintos elementos integrados en el sistema cuyo objetivo es ayudar a los estudiantes facilitando su trabajo con distintas herramientas, guías y consejos sobre cómo proceder en la realización de esas tareas.

Antes de empezar la prueba, se describe el escenario a los estudiantes proporcionando una visión rápida de la cuestión sobre la que van a trabajar y un resumen de las tareas que tendrán que realizar. Para las dos primeras tareas, los estudiantes han de leer tres breves artículos con información acerca de la cuestión y de los argumentos y posturas que adoptan distintas personas sobre la misma; cada artículo presenta una posición diferente y el tercero un punto de vista intermedio. Tienen que leer y evaluar los resúmenes que otros estudiantes han hecho del primer artículo y resumir el segundo y el tercero. En la tercera tarea los estudiantes deben escribir una carta al editor con una argumentación crítica sobre la cuestión, identificando y explicando los problemas de manera razonada y utilizando evidencia para ello. En las dos tareas siguientes se trabaja sobre la conexión de la evidencia proporcionada con las distintas posturas defendidas: los estudiantes deben indicar si los argumentos seleccionados están a favor o en contra de una determinada posición y si la evidencia proporcionada es relevante o no para el argumento en cuestión y en qué sentido. La última tarea es escribir un ensayo que contenga una argumentación razonada de la cuestión de interés.

Son numerosos los tests que se han construido siguiendo este diseño. En el trabajo de van Rijn, Graf y Deane en este mismo número se ofrecen algunos resultados obtenidos al poner a prueba los modelos formulados con esas progresiones de aprendizaje, tras administrar varias pruebas construidas con este diseño a estudiantes de secundaria obligatoria.

En suma, este trabajo presenta un diseño de evaluación basado en escenarios que está organizado y concebido para poder identificar cambios cualitativos en habilidades críticas, tal y como son medidas por los niveles de las progresiones de aprendizaje. Este tipo de diseño puede ofrecer una información mucho más rica y puede constituir una experiencia de aprendizaje que ayude a los estudiantes a internalizar los objetivos de evaluación y a los profesores a utilizar los resultados de los tests como un apoyo en su trabajo en el aula, ya que proporciona la oportunidad de modelar estrategias eficaces (evaluación *de*, *por* y *para* el aprendizaje).

Conflict of Interest

Acknowledgements

Notes

[1]The Common Core State Standards are research- and evidence-based learning goals, which outline what a student should know and be able to do at the end of each grade, in order to succeed in college, career, and life. See James Pellegrino's paper in this volume for detailed discussion on the CCSS.

[2]Note that we use these terms only as a way of capturing parallels between related skills that tend to be acquired at about the same time. We do not intend them to be linked to other uses of some of these terms, such as the definitions of performance levels for annual progress on various actual assessments.

[3]Este resumen ha sido realizado por la editora del número, María José Navas.

References

Aristotle (1939). *Topics* (E. S. Forster, Loeb, Trans.). Classical Library. Cambridge, MA: Harvard University Press.

Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning: A preliminary theory of action for summative and formative assessment. *Measurement: Interdisciplinary Research and Perspectives, 8*, 70-91.

Bracewell, J. R., Scardamalia, M., & Bereiter, C. (1978). The development of audience awareness in writing. *Resources in education*, 154-433.

Brem, S., & Rips, L. (2000). Evidence and explanation in informal argument. *Cognitive Science, 24*, 573-604.

Bright, R. (1995). *Writing instruction in the intermediate grades. What is said, What is done, What is understood*. Newark, DE: International Reading Association.

Butler, J. A., & Britt, A. M. (2011). Investigating instruction for improving revision of argumentative essays. *Written Communication, 28*, 70-96.

Chambliss, M. J. (1995). Text cues and strategies successful readers use to construct the gist of lengthy written arguments. *Reading Research Quarterly, 30*, 778-807.

Clark, R., & Delia, J. (1976). The development of functional persuasive skills in childhood and early adolescence. *Child Develoment, 47*, 1008-1014.

Coirier, P., Andriessen, J. E. B., & Chanquoy, L. (1999). From planning to translating: The specificity of argumentative writing. In P. Coirier & J. Andriessen (Eds.), *Foundations of argumentative text processing* (pp. 1-28). Amsterdam, The Netherlands: Amsterdam University Press

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence-based approach to reform.* NY: Center on Continuous Instructional Improvement, Teachers College-Columbia University.

Council of Chief State School Officers & National Governors Association (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects.* Washington, DC: Author. Retrieved from www.corestandards.org/the-standards/ELA-Literacy

De La Paz, S., & Graham, S. (1997a). Strategy instruction in planning: Effects on the writing performance and behavior of students with learning disabilities. *Exceptional Children, 63*, 167-181.

Deane, P. (2011). *Writing assessment and cognition* (ETS Research Report No. 11-14). Princeton, NJ: ETS.

DiPardo, A. & Freedman, S. W. (1988). Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research, 58*, 119-149.

Eisenberg, A., & Garvey, C. (1981). Children's use of verbal strategies in resolving conflicts. *Discourse Processes, 4*, 149-170.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88, 915-933.*

Ferretti, R. P., Lewis, W. E., & Andrews-Weckerly, S. (2009). Do goals affect the structure of students' argumentative writing strategies? *Journal of Educational Psychology, 101*, 577-589.

Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 93*, 694-702.

Flower, L., Hayes, J. R., Carey, L., Schriver, K. A., & Stratman, J. (1986). Detection, diagnosis, and the strategies of revision. *College Composition and Communication, 37*, 16-55.

Gorin, J. S., O'Reilly, T. P., Sabatini, J. P., Song, Y., & Deane, P. (2014). Measurement: Facilitating the goal of literacy. In B. Miller, P. McCardle, & R. Long (Eds.), *Teaching Reading and Writing: Improving Instruction and Student Achievement* (pp. 119-128). Baltimore, MD: Paul H. Brookes Publishing Co.

Graham, S., & Harris, K. R. (1989). Improving learning disabled students' skills at composing essays: Self-instructional strategy training. *Exceptional Children, 56*, 201-214.

Graham, S., & Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading* (A Carnegie Corporation Time to Act Report). Washington, DC: Alliance for Excellent Education.

Graham, S., MacArthur, C. A., & Schwarts, S. S. (1995). The effects of goal setting and procedural facilitation on the revising behavior and writing performance of students with writing and learning problems. *Journal of Educational Psychology, 87,* 230-240.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive process in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heritage, M. (2008). Learning progressions: Supporting instruction and formative assessment. Retrieved from http://www.k12.wa.us/assessment/ClassroomAssessmentIntegration/pubdocs/FASTLearningProgressions.pdf

Hillocks, G., Jr. (2002). *The testing trap: How state writing assessments control learning.* New York: Teachers College Press.

Kinsler, K. (1990). Structured peer collaboration: Teaching essay revision to college students needing writing remediation. *Cognition and Instruction, 7*, 303-321.

Klaczynski, P. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development, 71*, 1347-1366.

Knudson, R. E. (1992). The development of written argumentation: An analysis and comparison of argumentative writing at four grade levels. *Child Study Journal, 22*, 167-184.

Kuhn, D. (1991). *The skills of argument*. Cambridge, England: Cambridge University Press.

Kuhn, D., Katz, J. B., & Dean, D., Jr. (2004). Developing reason. *Thinking & Reasoning, 10*, 197-219.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction, 15*, 287-315.

Kuhn, D., & Udell, D. (2003). The development of argument skill. *Child Development, 74*, 1245-1260.

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing Norms of Argumentation: Metacognitive, Epistemological, and Social Dimensions of Developing Argumentive Competence. *Cognition and Instruction*, *31*, 456-496.

Larson, M., Britt, M. A., & Larson, A. A. (2004). Disfluencies in comprehending argumentative texts. *Reading Psychology, 25*, 205-224.

Leitão, S. (2003). Evaluating and selecting counterarguments: Studies of children's rhetorical awareness. *Written Communication, 20*, 269-306.

Matsuhashi, A. & Gordon, E. (1985). Revision, addition, and the power of unseen text. In S. Freedman (Ed.), *The acquisition of written language: Response and revision* (pp. 236-249). Norwood, NJ: Ablex.

McCann, T. M. (1989). Student argumentative writing: knowledge and ability at three grade levels. *Research in the Teaching of English, 23*, 62-76.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3-67.

Nation's Center for Education Statistics. (2008). *The nation's report card: Writing 2007* (U.S. Department of Education Publication No. NCES 2008-468). Washington, DC: U.S. Government Printing Office. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2007/2008468.pdf

Newell, G. E., Beach, R., Smith, J., & Van Der Heide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly, 46*, 273-304.

Nussbaum, M. E., & Kardash, C. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology, 97*, 157-169.

O'Reilly, T., & Sabatini, J. (in press). *Reading for Understanding: How performance moderators and scenarios impact assessment design* (ETS Research Report). Princeton, NJ: Educational Testing Service.

O'Reilly, T., & Sheehan, K. M. (2009). *Cognitively Based Assessment of, for and as Learning: A Framework for Assessing Reading Competency* (ETS Research Report RR-09-26). Princeton, NJ: Educational Testing Service.

Pellegrino, J. W. (2014). Assessment as a Positive Influence on 21st Century Teaching and Learning: A Systems Approach to Progress. *Psicología Educativa, 20*, 65-77.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83-106). Hillsdale, NJ: Erlbaum.

Persky, H. R., Daane, M. C, & Jin, Y. (2003). *The nation's report card: Writing 2002* (U.S. Department of Education Publication No. NCES 2003-529). Washington, DC: U.S. Government Printing Office.

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

Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (pp. 778-803). New York: Macmillan.

Sexton, M., Harris, K. R., & Graham, S. (1998). Self-regulated strategy development and the writing process: Effects on essay writing and attributions. *Exceptional Children, 64*, 295-311.

Shanahan, T. (Ed.) (1990). *Reading and writing together: New perspectives for the classroom*. Norwood, MA: Christopher-Gordon.

Shavelson, R. (2009). *Reflections on learning progressions*. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA.

Sheehan, K. M., & O'Reilly, T. (2011). *The CBAL Reading Assessment: An Approach for Balancing Measurement and Learning Goals* (ETS Research Report RR-11-21). Princeton, NJ: Educational Testing Service.

Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology, 6*, 283-294.

Stein, N. L., & Miller, C. A. (1993). A theory of argumentative understanding: Relationships among position preference, judgments of goodness, memory, and reasoning. *Argumentation, 7*, 183-204.

Van Eemeren, F. H., & Grootendorst, R. (1992). *Argumentation, communication, and fallacies: A pragma-dialectical perspective.* Mahwah, NJ: Erlbaum.

Van Eemeren, F. H., Grootendorst, R., & Henkemans, F. S. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments.* Mahwah, NJ: Erlbaum.

Van Rijn, P. W., Graf, E. A., & Deane, P. (2014). Empirical Recovery of Argumentation Learning Progressions in Scenario-Based Assessments of English Language Arts. *Psicología Educativa, 20*, 109-115.

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

Walton, D. (1992). *Plausible argument in everyday conversation*. Albany, NY: State University of New York Press.

Zieky, M. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa, 20*, 79-87.