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.

ARTICLE INFORMATION

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.

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.

*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