Measuring online reading comprehension in open networked spaces: Challenges, concerns, and choices

Jill Castek (jcastek@berkeley.edu) University of California Berkeley
Julie Coiro (jcoiro@snet.net) University of Rhode Island

Introduction. The Internet has become increasingly central to our daily lives (Johnson, Levine, Smith, & Smythe, 2009), transforming the ways we access, use, and exchange information. To fully participate in a globally networked society, every student needs to develop strategies for locating, comprehending, and responding to text in ways that exploit the potentials of information and communication technologies (ICTs) (Educational Testing Service [ETS], 2003; International Reading Association [IRA], 2009; National Council of Teachers of English [NCTE], 2007). Making sense of digital information requires skills and strategies that are complex, and in some cases unique, to online reading and writing contexts (Afflerbach & Cho, 2008; Coiro & Dobler, 2007; Transliteracies Project, 2006). Thus, students' proficiencies in the new millennium cannot be determined solely on the basis of their literacy performance in non-digital contexts (Leu, Castek, Hartman, et al., 2005; O'Brien, 2006).

The emergence of the Internet as today's defining technology for literacy and learning (Leu, 2007) and the rapid shift in literacy practices from page to screen (Lankshear & Knobel, 2006), has prompted educators to examine how to assess students' online literacy competencies, explore ways of tracking progress over time, and discuss methods for identifying strengths and weaknesses that can be addressed through instruction.

Purpose. This paper and poster present insights gleaned from four years of work designing, administering, and scoring **o**nline **r**eading **c**omprehension **a**ssessments (ORCAs). These measures were intended for students in upper elementary and middle school classrooms and have proven to be reliable and valid in three defended dissertation studies (Castek, 2008; Coiro, 2007; Henry, 2007) and two federally funded research grants (Leu & Hartman, 2004-05; Leu & Reinking, 2005-08). By drawing attention to research and principles of evidence-centered design, we aim to incite thinking that will prompt additional questions about assessment that will drive future work.

Perspectives and Methods of Inquiry. Two perspectives informed the analysis of our trajectory of assessment design over a period of years. First, from a new literacies perspective of online reading comprehension (Leu, Kinzer, Coiro, & Cammack, 2004), we argue that traditional offline reading comprehension skills are necessary, but not sufficient, to read and learn from information on the Internet. Second, we draw on evidence-centered principles of assessment design (National Research Council, 2001; Mislevy, 1996) to answer three questions: 1) What evidence is needed to represent proficiency in online reading comprehension? 2) What situations or tasks can elicit such evidence? and 3) How can this evidence be interpreted in ways that are meaningful for classroom teachers? This 'assessment triangle' offered a useful framework to analyze insights emerging from our analysis of assessment approach while also identifying challenges to address when designing future ones.

Findings. Analyses of our assessment design efforts over the course of several years have yielded considerations in three areas: 1) cognition; 2) observation; and 3) interpretation. Insights, challenges, and design choices are summarized (by category) in the sections that follow.

I. Cognition: What knowledge and skills are important to measure when assessing students' ability to read for information on the Internet?

Across several studies, (Castek, 2005, 2008; Coiro, 2005, 2008, 2009; Coiro, Castek, Henry, & Malloy, 2007; Coiro, Malloy & Rogers, 2006; Henry, 2008; Leu et al., 2008) we have created and examined measures that look at the ways learners apply online reading comprehension strategies to solve researcher-generated informational problems aligned to content-area subjects.

Insights

1. Our work has yielded an emerging taxonomy of multidimensional online reading comprehension skills and strategies that include questioning, locating, critically evaluating, synthesizing, and communicating (Leu et al., 2005; 2008).
2. These skills progress developmentally in phases. We have discovered how to keep tasks authentic but appropriately leveled by starting with shorter inquiries (2 components) and moving to more complex inquiries that integrate 3 or 4 components with varying degrees of difficulty (Coiro, 2007; Leu et al., 2008).
3. Online reading comprehension tasks appear to prompt gradations of online reading skills – some that are more similar offline reading comprehension and others that are less similar to offline reading comprehension (Coiro, 2007; Leu et al., 2005)
4. A reader's positive dispositions play an important role in online reading comprehension performance. These dispositions include readers' attitudes and beliefs about using the Internet for inquiry and learning and their own levels of self-efficacy in relation to their peers and adults (Coiro, 2007; O'Byrne & McVerry, 2009).
5. In online learning, topic-specific prior knowledge for successful comprehension may vary in relation to one's ability to locate online information (Coiro, 2007).

Challenges

1. Item-dependency. Assessments of reading comprehension assume items are linear, separate, and independent. However, the process of online reading is cyclical, individualized, and highly dependent on a reader's online locating skills (to access texts) and a reader's online communication skills (to share what was learned with different audiences using a range of online tools).
2. Valid measures. Each component of online reading comprehension covers multiple dimensions that, in reality, would require many, many items. Often as assessment designers, we must choose which dimensions are most important for a given task or most developmentally appropriate for a certain age level reader. However, it simplifies what is typically a complex task to create a simplistic representation and has implications for the validity of online reading comprehension assessments.
3. Reliable measures. Online texts and reading contexts change rapidly, often from one day to the next. Designers update their websites and readers post new comments using social networking tools. This rapidly changing reading environment has implications for the reliability of assessments that measure online reading comprehension.

4. Measuring reading comprehension through writing and communicating a response. While reading and writing are interconnected in online contexts, understanding ideas and being able to share what was learned and communicate these ideas requires learners to have well-developed writing processes as much as it does reading comprehension skills.

Design Choices

1. **Unidimensional OR Multidimensional Constructs:** When is it useful to design more items to assess a limited number of dimensions of one online reading component (e.g., searching, critically evaluating, synthesizing, and communicating) and when is it useful to have fewer items that assess more dimensions of online reading comprehension?
    a. **Locate:** This construct includes the ability to: (1) generate keywords; (2) locate relevant links within search results of a global search engine; (3) locate relevant links with search results of a website's local search engine); (4) make inferences from hyperlink menus on a website to locate key ideas; and (5) determine when to initiate a new search.
    b. **Critically evaluate:** This construct includes the ability to judge information's level of: (1) relevance; (2) accuracy; (3) reliability; and (4) perspective. Each involves a process that could be measured in isolation, but is often more effective if readers integrate ideas across multiple dimensions.
    c. **Synthesize:** This construct includes the ability to: (1) identify key ideas across one or more texts; (2) articulate how parts of texts work together, inform one another, or contradict one another; (3) make connections across points of multiple authors; (4) integrates key ideas to meet demands of context; and (5) extends key ideas with a personal connection or interpretation.
    d. **Communicate:** This construct includes the ability to: (1) select an appropriate tool; (2) use specific features of a tool; (3) demonstrate an understanding of audience and appropriate level of formality and tone; (4) demonstrate understanding of purpose; (5) construct a written response to a question or problem; and (6) construct a multimedia response to a question of problem.
2. **Item Independence OR Item Dependence:** When is it best to design items that measure each specific online reading comprehension component separately (which makes it possible to fulfill psychometric assumptions) and when is more appropriate to design items that integrate skill performance into a final product (and address the confounding variables for analysis)?
    a. **Location:** Locating online information appears to be a bottleneck skill that impacts every other component of online information. If students are not able to locate the information needed to answer a question, at what point should they be encouraged to move on and or what supports should be made available?
    b. **Synthesis:** When students successfully synthesize online information from multiple sources, it indicates that locate and evaluate skills are in use. In what ways can assessment information be used to determine when a higher-level skill (such as synthesis) has occurred, and how can this complex process be captured?
    c. **Communicate:** When students communicate what they have learned to an audience, the process of composing a clear and well-supported message using appropriate tools and social practices confounds components of synthesis and communication.

3. **Ever-changing OR Constant reading contexts:** Should an assessment take advantage of ever changing texts common to the Internet environment (which could be more valid but probably less reliable) or one that works around these changing texts to hold them constant (which could be more reliable but probably less valid)?

II. Observation: What specific tasks and situations will illicit responses that reveal students' knowledge and skills related to locating, critically evaluating, synthesizing and communicating information to solve a set of information problems?

We have found that scenario-based online reading comprehension tasks, contextualized as information challenges, related to a content area or societal implication (e.g., homelessness) promote higher level thinking, engagement in learning, and investment in the assessment task.

Insights

1. Authentic Inquiry: Our assessment work suggests that applying content area knowledge to solve broad interdisciplinary problems using online resources involves using the full range of online reading comprehension strategies a learner brings to an assessment task. This approach and can provide a range of assessment data from which to draw insights and implications.
2. Curriculum-Based Connections: We have found that using curriculum-based themes for informal classroom assessments, rather than more generalized topics for large-scale assessments, promotes engagement and investment in assessment tasks.
3. Authentic Communication: Our research suggests that making the tools students use in out-of-school contexts (IM, social networking tools, digital video) and the workplace (email, video-conferencing, support networks) accessible can support collaboration and skilled application of online reading comprehension. However, these online tools don't often appear in school classrooms and many students are unfamiliar with ways to use them to support academic learning.

Challenges

1. Making Assessment Tasks Simultaneously Authentic and Motivating. Designing externally generated problems to solve when online inquiries are most often prompted by personal questions presents a challenge. Personal questions may be more meaningful but including them in assessment tasks makes assessment less uniform and more difficult to score in a reliable way.
    1. Providing an option to follow questions posed by individuals may increase motivation and persistence (more authentic) but produce a wide array of experiences that can not be evaluated in a similar way (causing challenges for scoring)
    2. Involving students in authentic tasks is time consuming. Tasks need to fit within the boundaries of a school day and the time constraints of a school schedule.
    3. An authentic means by which to accomplish a real-word problem is to invite collaboration. However, in schools, assessment tasks are most often individualized and require an assessment of individual contributions.

Design Choices

1. **Open Reading Environments (unbounded and authentic) OR Closed Reading Environments (bounded and simulated)**
   - Open: the advantages: real-time and current, reading choices are unlimited, and distractions/advertising are authentic, but it is difficult to capture real-time data (although possibilities are emerging with software). Online texts constantly change.
   - Closed: it is possible to capture log file data, but closed environments limit navigational choices. It is possible to populate closed interfaces with information that remains constant (standardizing the reading experience for each reader), but utilizing a "simulated" version of the Internet is not an adequate stand in for performance in an authentic online reading (there are fewer navigational choices; less complexity; which may artificially elevate performance).
2. **Capturing isolated skills OR integrated application**
   - Isolated skills: Examining closely students' competency in terms of discrete skills they possess lends itself to the design of lessons that promote greater competency with a given component. However, can skills learned in isolation be readily integrated and used to solve authentic problems in meaningful ways?
   - Integrated application: Examining integrated application offers options for assigning more engaging assignments and projects that utilize a range of skills. However, it is often difficult to determine the precise areas where students need more support.
3. **Capturing Individualized performance OR collaborative performance**
   - Individualized: Tasks are more streamlined, and enable examination of decision-making, time-management, and meaning making on an individual level. However, individualized task completion may not represent the collaboration and idea sharing that most often occurs between readers connected online.
   - Collaborative: It may be more difficult to parse individual contributions during authentic online collaboration since decision-making is undertaken as a joint endeavor. Conversational turns influence navigational choices. Dialogue may either distract from or promote meaning making depending on the situation, but it may represent authentic uses of social networking and collaboration that readily occurs online (both face to face and virtually).
   - Pretend-collaboration: Controlled collaboration (communicating with a researcher as a collaborator) can be considered as an option and may evoke some of the ways authentic collaboration occurs online. However, constraints in terms of how much support to provide and uniformity in the level of support offered each reader must be considered and controlled.

III. Interpretation: What set of methods and tools can classroom teachers use to make sense of student processes and products in ways that inform their instruction?

Across several studies (Castek, 2008; Coiro, 2008; Coiro, Malloy, & Rogers, 2006; Leu et al., 2004-2005; Leu & Reinking, 2006 – 2009; The New Literacies Research Team and Internet Reading Research Group, 2007) online reading comprehension tasks have been designed with the purpose of eliciting, documenting, and examining students' online reading behaviors. Archiving screen captures of students processes have made it possible to highlight reading processes, identify patterns across learners, and help teachers understand the processes involved.

Insights

1. Useful to Researchers and Teachers. Assessment information is useful to both researchers as well as teachers, but each in a different way. While researchers intend to measure growth over time in response to an intervention, teachers use assessment data to drive instructional decision making, to construct assignments that offer valuable experience with essential skills, and to structure engaging learning context where scaffolding (peer and teacher) can take place.
2. Capturing Process and Product. Assessment tasks can be seen as a means of evaluation or can be viewed as instructional tools. Assessment can also be viewed as a space to share strategies and monitor progress over time. In our early work, the purpose of online reading comprehension assessment was to help teachers understand the processes involved. However, we found that by examining these data closely that we could get an idea of more and less effective strategies to inform classroom instruction.
    a. Examining an online reader's process is an important means of understanding the skills that readers control and strategies that are well developed or underdeveloped). These strengths and weaknesses can be used formatively to design meaningful instructional sequences.
3. Assessment of Learning; Assessment for Learning. We have found that similar protocols can be useful for assessments **of learning** and **for learning**. Formative assessment are not scores to be used for evaluation but instead are used to drive instructional decision-making and instructional support. Researchers can use similar assessments to interpret the degree of student learning from pre to post. Depending on the purpose of any given assessment, different kind of output data can be collected and used in a variety of ways.
4. Building Online Reading Skills and Applying them to Content Learning. Student products created over a period of time (written responses, projects) provide a way to assess the application of a range of skills and that may be linked to content aims.
5. The same assessment structure can be used for multiple purposes. The structure of online assessments can be flexible. It is possible to: 1) assess learning in one short test (using a set of items); or 2) within an assignment or inquiry based task completed over the course of a unit. Using a short test will provide information about mastery of unique skills. However, longer, more in-depth inquiry based projects can measure learning over a period of time in and integrated and applied way.

Challenges

1. **Making time within a school day** to administer assessments, score and analyze assessments, and synthesize the results of assessments in ways that will drive instruction.
2. **Interpreting multiple choice items**. These types of items are less open to interpretation, and can be scored quickly in a uniform way. Scoring can be done with a computer and results can be automated and quick. However, they are more difficult to construct (less precise), lack authenticity within the online context (they don't represent the way information is shared online), and may provide an artificial view of a learners true abilities (students who guess correctly).

3. **Interpreting constructed response items**. These types of items require creating specialized rubrics to evaluate responses. Archiving and later scoring items can be time consuming and requires many hours of training. Depending on the evaluation tool used to score these items (levels of correctness) it may be difficult to determine the skills learners control and exactly where additional support may be needed. However, inviting students to share their opinions and explain what they learned (as well as where and how the information was found) may be a more authentic means of communicating in an online environment (sharing resources, extending ideas).
4. **Interpreting integrated, project-oriented tasks.** These sorts of projects invite multiple student responses and are often constructed over a period of days or weeks. They require students to show what they have learned by applying these skills and strategies to the completion of a project that showcases the content learned. However, the processes used may be difficult to infer and the challenges with skills and strategies students faced may be difficult to determine (and support) after the project is complete.

Design Choices

1. **Computerized/Automated OR Human Scoring Systems:** Computerized. Each type of scoring format lends itself to particular affordances and constraints, as outlined in the above section on challenges. Computerized scoring systems are quick and provide simplistic information about knowledge and skills while human scoring systems can prompt more insights into the processes and misconceptions of students as they engage with different dimensions of the online reading tasks.
2. **Using prepared/available assessments of online reading comprehension OR designing your own that are specific to what you want to learn about what your students know and can do.** If instantly available results are what is most needed, opt for using prepared online test of skills that follows a multiple choice format. Because results can be readily available, they can inform instructional decisions immediately. However, if integrated content area learning is to be assessed, then a curriculum based and process based choice may be the best option because it can be customized to any content and adjusted to students' (and teachers') individual choices.
2. **Assessments OF learning OR Assessments FOR learning**. Making this decision requires focus on different kinds of output data and interpretations depending on your purpose. Examining students processes (through screen captures) or examining responses that are constructed after students gather pieces of information to solve a task.

References

Afflerbach, P. & Cho, B.Y. (2008). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. Israel and G. Duffy (Eds.). *Handbook of reading comprehension research*. Mahwah, NJ: Erlbaum Associates.

Castek, J. (2004). Exploring Internet meta-language during assessment tasks: An examination of 7th grade students' new literacies in action. In D. Johnson (Chair) *Bridging instruction in literacy and technology: Applications for new literacies in instructional settings*. Symposium presented at the National Reading Conference (NRC). San Antonio, TX.

Castek, J. (2008). *Examining classroom instruction that integrates the new literacies of online reading comprehension: Exploring the contexts that facilitate acquisition and the learning outcomes that result.* Unpublished doctoral dissertation, University of Connecticut.

Coiro, J. (2005). Making sense of online text. *Educational Leadership, 63*(2), 30-35.

Coiro, J. (2007). *Exploring changes to reading comprehension on the Internet: Paradoxes and possibilities for diverse adolescent readers.* Unpublished doctoral dissertation, University of Connecticut.

Coiro, J. (2009a). Rethinking reading assessment in a digital age: How is reading comprehension different and where do we turn now. *Educational Leadership*, 66, 59-63.

Coiro, J., Castek, J., Henry, L.A., & Malloy, J. (2007, December). A closer look at measures of online reading achievement and school engagement with seventh graders in economically challenged school districts. In D. Reinking (Chair) *Teaching Internet comprehension to adolescents who are at risk to become dropouts*. A symposium presented at the National Reading Conference (NRC). Austin, TX.

Coiro, J., & Dobler, E. (2007). Exploring the comprehension strategies used by sixth-grade skilled readers as they search for and locate information on the Internet. *Reading Research Quarterly*, *42,* 214-257.

Coiro, J., Malloy, J., & Rogers, M. (2006). Patterns of effective strategy use among adolescent online readers. In D.J. Leu, Jr. (Chair) *Developing Internet Reading Comprehension Strategies Among Adolescents At Risk to Become Dropouts*. A structured poster session presented the American Education Research Association (AERA). San Francisco, CA.

Educational Testing Service [ETS]. (2002). *Digital Transformation A Framework for ICT Literacy: A Report of the International ICT Literacy Panel*. Princeton, NJ: ETS.

Henry, L. A. (2007). *Exploring new literacies pedagogy and online reading comprehension among middle school students and teachers: Issues of social equity or social exclusion?* Unpublished doctoral dissertation, University of Connecticut, Storrs, Connecticut.

International Reading Association. (2009). *New literacies and 21st century technologies: A position statement*. Newark, DE: IRA. Retrieved Sept. 16, 2009 from www.reading.org/General/AboutIRA/PositionStatements/21stCenturyLiteracies.aspx

Johnson, L., Levine, A., Smith, R., and Smythe, T. (2009). *The 2009 Horizon Report: K-12 Edition*. Austin, Texas: The New Media Consortium.

Lankshear, C., & Knobel, M. (2006). *New literacies: Everyday practices and classroom learning* (2nd ed.). Buckingham, U.K.: Open University Press.

Leu, D. J. (2007). *What Happened When We Weren't Looking? How Reading Comprehension Has Changed and What We Need To Do About It*. Keynote research address presented at the Annual Meeting of the International Reading Association, Reading Research Conference, Toronto, CA.

Leu, D. J., Castek, J., Hartman, D., Coiro, J., Henry, L., Kulikowich, J., & Lyver, S. (2005). *Evaluating the development of scientific knowledge and new forms of reading comprehension during online learning*. Final report presented to the North Central Regional Educational Laboratory/Learning Point Associates. Retrieved April 1, 2010 from http://www.newliteracies.uconn.edu/ncrel.html.

Leu, D.J., Coiro, J., Castek, J., Hartman, D., Henry, L. A., Reinking, D. (2008). Research on instruction and assessment in the new literacies of online reading comprehension. In K. Collins Block and S. Parris (Eds). *Comprehension instruction: Research-based best practices*. (321-346). New York, NY: Guilford.

Leu, D. J. & Hartman, D. (2004-2005). *Evaluating the Development of Scientific Knowledge and New Forms of Reading Comprehension During Online Reading*. North Central Educational Research Lab.

Leu, D. J., Jr., Kinzer, C.K., Coiro, J., Cammack, D. (2004). Toward a theory of new literacies emerging from the Internet and other ICT. In R.B. Ruddell & N. Unrau (Eds.), *Theoretical Models and Processes of Reading*, Fifth Edition (1568-1611). Newark, DE: International Reading Association.

Leu, D. J. & Reinking, D. (2005-2008). *Developing internet comprehension strategies among poor adolescent student at risk to become dropouts.* Three year federal research grant funded by U.S. Department of Education's Institute for Education Science (IES).

McVerry, J. G., O'Byrne, W. I. & Robbins, K. (2009). *Validating instruments used in the TICA Project.* American Educational Research Association (AERA) Annual Meeting, San Diego, CA.

Mislevy, R. J. (1996). Test theory reconceived. *Journal of Educational Measurement, 33,* 379-416.

National Council of the Teachers of English [NCTE]. (2007.) 21st-century literacies: A policy research brief. Retrieved February 2, 2008 from http://www.ncte.org/library/files/Publications/Newspaper/Chron1107ResearchBrief.pdf

New Literacies Research Team* & Internet Reading Research Group† (2007). Defining online reading comprehension: Using think aloud verbal protocols to refine a preliminary model of Internet reading comprehension processes. In D. Alvermann (Chair) *21st Century Literacy: What is it, How do students get it, and how do we know if they have it?* A Symposium presented at American Education Research Association. Chicago, IL. *Including: Castek, J., Coiro, J., Henry, L.A., & Leu, D.J. at the University of Connecticut. †Including: Carter, A., Malloy, J., Reinking, D., Rogers, A., Robins, K., & Witte, J. at Clemson University. Online: http://docs.google.com/Doc?id=dcbjhrtq_10djqrhz

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.

New Media Consortium. (2007). *Social Networking, The Third Place, and the Evolution of Communication*. Retrieved October 1, 2009 from http://www.nmc.org/evolution-communication

O'Brien, D. G. (2006). "Struggling" adolescents' engagement in multimediating: Countering the institutional construction of incompetence. In D. E. Alvermann, K. A. Hinchman, D. W. Moore, S. F. Phelps, & D. R. Waff (Eds.), *Reconceptualizing the literacies in adolescents' lives* (29-45). Mahwah, NJ: Lawrence Erlbaum.

O'Byrne, W. I., & McVerry, J. G. (2009). Measuring the dispositions of online reading comprehension: A preliminary validation study. In K. Leander, D. Rowe, D. Dickson, M. Hundley, R. Jiménez, & V. Risko (Eds.), *The 57th National Reading Conference Yearbook* (362-375). Oak Creek, WI: National Reading Conference.

Transliteracies Project. (2006). *Research in the technological, social, and cultural practices of online reading*. Retrieved on October 1, 2009 from http://transliteracies.english.ucsb.edu