CHAPTER 5

INFORMATION LITERACY IN DIGITAL ENVIRONMENTS: CONSTRUCT MEDIATION, CONSTRUCT MODELING, AND VALIDATION PROCESSES

Irvin R. Katz
Educational Testing Service

Norbert Elliot
New Jersey Institute of Technology

Information literacy (IL) is a 21st century skill most often conceptualized and measured through 20th century assessment practices. Designed by the Association of College and Research Libraries (ACRL) in 1999 and approved by its board in 2000, the *Information Literacy Competency Standards for Higher Education* (*IL Standards*) is a *fin de siècle* statement. A construct caught between the print-based world of the 20th century and the digitally based networks of the 21st, IL is framed as a skill that, once mastered, will allow those who possess it to stem the tide of the "uncertain quality and expanding quantity of information" that will surely "pose large challenges for society" (ACRL, 2000, p. 2). Those who have this skill will be able to locate, evaluate, and use information effectively to sort through the "abundance of information" that "will not in itself create a more informed citizenry" (p. 2). With the advent of Web 2.0 over the intervening 13 years—its social media, virtual communities, blogs, wikis, podcasts, folksonomies, and mashups—the *IL Standards* look as if they had been cast by Gutenberg.

In response, in 2013 the ACRL chartered a task force to begin the process of updating the *IL Standards*. Noting the limits of the "competency and skill-based approach," the revision promises a "conceptual approach" that provides not just a detailed listing of skills but, rather, "a set of archetypal of core abilities that will serve students well in a continuously changing information environment" (Jacobson & Gibson, 2013, p. 1). The new *Framework for*

Information Literacy for Higher Education (*Framework for IL*) (ACRL, 2015) advances six threshold concepts, transformative perspectives discussed by Barry Maid and Barbara J. D'Angelo (Chapter 2, this collection). Defined as "a spectrum of abilities, practices, and habits of mind that extends and deepens learning through engagement with the information ecosystem" (p. 1), IL is advanced as these threshold concepts are advanced: Authority Is Constructed and Contextual; information is created through intentional acts; information has value; research is a process of inquiry; scholarship is conversational; and searching is strategic. Key to the *Framework for IL* in its present form is the recognition that IL is a contextualized, complex experience in which the cognitive, affective, and social dimensions of the searcher have distinct roles.

The emphasis on underlying cognitive abilities as they are identified, studied, and used in digital environments is the subject of this chapter. To identify important conceptual and validation practices associated with the construct of IL, in this chapter we focus on our experience with iSkills™, a digitally based assessment that uses real-time, scenario-based tasks to measure a defined IL construct.

iSkills: INFORMATION LITERACY IN A DIGITAL ENVIRONMENT

In 2002, Educational Testing Service (ETS) recognized that digital environments made public since the early 1990s presented a unique opportunity for test delivery (ETS, 2002). Instead of simply digitizing the bubble and booklet test format that had been used since its charter in 1947, ETS researchers recognized that emerging information and communication technologies held the potential for new forms of assessment. In digital environments, items could be presented that required a series of actions—not simply the identification of a single correct answer. If a student were to be examined on ability to access information, that student could be required to enact a process of decision-making within rhetorical contexts sensitive to aim, genre, and audience. The student could be placed in a context—as an employee, for instance, seated before a computer at a fictitious company named Restoration Demolition in which a request had been made from a customer to find stained glass panels in good condition. Faced with a choice of databases from which information may be drawn, increasingly precise searches might yield exactly those results needed by the customer (Figure 5.1).Performing a simulation of information retrieval within databases, the candidate demonstrates search capability and then selects information to meet the customer request.

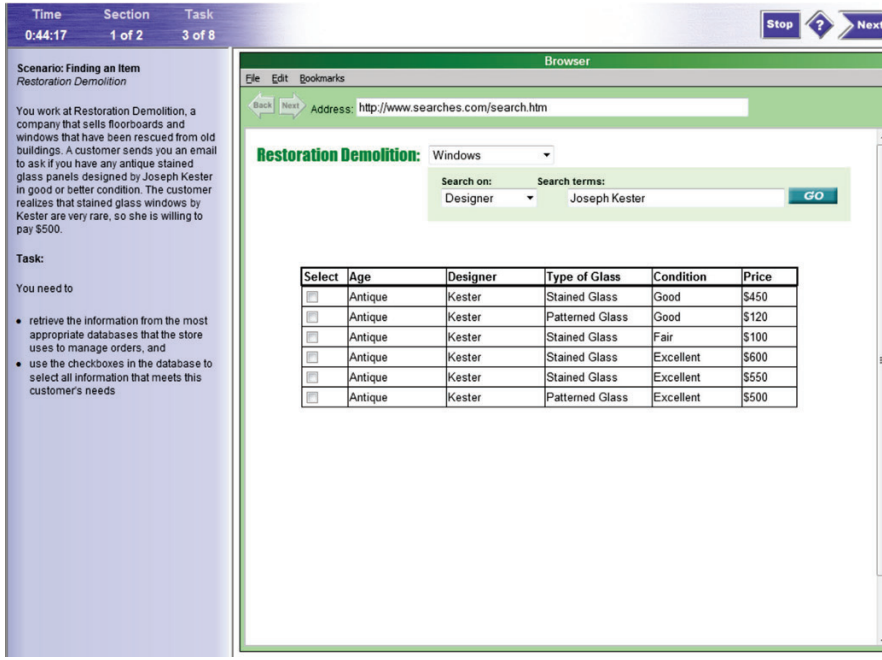

Figure 5.1. Restoration Demolition: iSkills assessment task. © 2015, Educational Testing Service. All rights reserved. Used with Permission.

In a print-based environment, the candidate would select a single correct answer from a list; in this digital environment, the candidate selects the appropriate database from a list of alternatives, types search terms, reviews results and may try an alternative search strategy, and then selects the responses from the results that answers the customer's query—while the machine records database selection, search terms, the number and characteristics of the potentially multiple searches conducted, and the relevance and appropriateness of reviewed and selected results. Blending process and product, the resulting system, iSkills, represents a type of next-generation assessment in which real-life applications track the learning process as it occurs in the digital environments (Tucker, 2009) that exist within the ecologies of interacting information sources identified by Kathleen Blake Yancey (Chapter 4, this collection). From the design of tasks to the delivery of results, digital environments provide a new way of thinking about large-scale assessment programs.

For those responsible for the selection of assessment instruments in specific institutional sites, answers to three questions will be helpful when assessment of IL occurs in digital environments:

1. Is the construct of IL itself mediated by the digital environments in which it is assessed?
2. How might the construct of IL be modeled in digital assessment environments?
3. What do present theories of validation tell us about how institutional assessments might be used to help students become more information literate?

CONSTRUCT MEDIATION IN DIGITAL ENVIRONMENTS

For purpose of exposition, it is useful to begin by returning to the 1955 definition of a construct as proposed by the psychometrician and philosopher team of Lee J. Cronbach and Paul E. Meehl (1955): "A construct is some postulated attribute of people, assumed to be reflected in test performance" (p. 283). Broadly speaking, all constructs are mediated—that is, following the Medieval Latin origin of the word, the construct is divided in the middle, with part postulated (or targeted) and part performed (and measured). Viewed in this fashion, all environments mediate constructs, and the extent of the mediation depends on standpoint. To use a well-known example, writing mediates knowledge, as Walter Ong (1982) famously demonstrated in his study of orality and literacy. Writing restructures consciousness, he boldly proposed, and backed it up by calling to mind distinctions between orally based thought and literacy-based practices: Oral cultures are empathetic, he claimed, while writing establishes objectivity. A list of such dualisms, his psychodynamics of orality illustrate that all constructs exist in domains that are mediated by—and reflected in—performance. For Ong, the mediation of language by writing was of paramount importance; for his critics (Scribner & Cole, 1981), the generalized cognitive effects of the technology of literacy could not be substantiated.

By the mid-1990s, it was becoming clear that the digital world—a new environment—was emerging as the latest in a series of contextual shifts that had begun 5,300 years ago with the writing on the clay tablets of Uruk (Woods, 2010). Whether the technology was the visible language created by stylus or pixel, all mediation, Jay David Bolter and Richard Grusin (1996) claimed, is remediation (p. 346). As a theory of media, remediation is a concept that allows us to investigate the promise of digital representation: the way digital environments import and refashion other media into digital space; the ways these environments suggest reality itself with image, sound, and haptic technology; and the ways they allow participants to reform reality as they synchronously participate in events. So powerful are these digital environments today that it is difficult to imagine a context in which a user could avoid mediation when

engaging the IL construct. The question thus becomes one of agency: How is the IL construct mediated?

Here Herbert A. Simon's parable of the ant is useful. In *The Sciences of the Artificial* (1996), Simon offered the following narrative and its interpretation:

> We watch an ant make his laborious way across a wind- and wave-molded beach. He moves ahead, angles to the right to ease his climb up a steep dune, detours around a pebble, stops for a moment to exchange information with a compatriot. Thus he makes his weaving, halting way back to his home. (p. 51)

In coming to terms with the journey of the ant, we wonder at the irregular, complex series of traced and retraced steps and realize that the wandering is due to the encountered obstacles of pebble and path. Simon offers an hypothesis: "An ant, viewed as a behaving system, is quite simple. The apparent complexity of its behavior over time is largely a reflection of the complexity of the environment in which it finds itself" (p. 52).

In the case at hand, let's imagine that Simon's ant parable is about the IL construct and how it reveals itself through behavior. Let's narrow the construct to the variable of information access as the employee of Restoration Demolition engages it. Following Simon, the seeming elementalist (bubble and booklet) or complex (constructed response) behaviors comes not from different constructs but from the complexity of the environments in which the construct is assessed. As such, print and digitally based samples of the construct used in a given assessment reflect a different beach and therefore lead to different behavior. Effectively, each measures a different construct—the measured construct requiring, let's say, its own smaller circle of behavior—although the constructs may certainly be related, if only because they derive from the same underlying domain.

So, to answer the first question—is the construct of IL mediated in digital assessment environments?—we offer the following answer: Measured constructs are indeed mediated by the way the assessment designers sample the construct. As such, depending on how the assessment designers view the digital scene of action, there may be differences in what is being measured.

CONSTRUCT MODELING IN DIGITAL ENVIRONMENTS

The concept of mediation is extremely helpful in allowing us to reflect on the impact of digital environments on constructs. However, that concept alone is insufficient if we are to examine assessment of learning in digital environments. Required is an additional concept: modeling.

Susan E. Embretson (1983) recognized that the impact of the information processing perspective described by Simon led to a shift from "explaining antecedent/consequent relationships to explaining performance from the systems and subsystems of underlying processes." "As a paradigm shift," she continued, "the information-processing view entails changes not only in the questions that are asked but also in the type of data that are deemed relevant" (p. 179). Because construct modeling was, in fact, the equivalent of theory building, Embretson proposed that a sound construct model must account for individual performance, allow for comparison of alterative models, yield quantification of the constructs in the model, and provide information about individual differences in performance. In the study of writing—a field familiar to readers of this volume—the most significant modeling work has been that of John R. Hayes (2012) who has been modeling the writing construct for over three decades (Hayes & Flower, 1980). Delineation of cognitive processes—writer's control of task, writing processes, composition environment, and resource level—has transformed our concept of writing. Because of the work of Hayes and his colleagues, we now know that writing is not a series of mechanically executed displays of knowledge of conventions but, rather, a complex socio-cognitive process.

While the concept of mediation is one of scene (where IL occurs), the concept of modeling is one of agency (how the construct is modeled). While the digital environment of iSkills involves pebbles on a new beach, the differences in performance we see are due to differences in the way the construct is sampled by iSkills. Ultimately, the targeted construct—information access, for example—is nevertheless identical to those for print communication, inasmuch as the digital and print assessments both intend to assess IL. Even if the assessment-makers' intentions are to assess IL, differences in performance are artifacts of the assessment environment and may result in different measured constructs.

Here is the key: the assessment environment of iSkills introduces nontrivial performance differences within the constructed-response task (Bennett, 1993). Knowing how to avoid pebbles and navigate the paths is essential to the performance of the student and, in turn, to an institution's assessment of that candidate's level of IL. Just because two assessments are labeled "information literacy" by their respective developers does not mean that the mediated construct of IL (the measured construct) is the same—and, thus, may be modeled differently.

Two examples are in order to bring this theoretical discussion into their practical application.

Comparison of two tests—the print-based Standardized Assessment of Information Literacy Skills (SAILS) test (O'Connor, Radcliff & Gedeon, 2002) and the digitally based iSkills (Katz, 2007b)—demonstrates the distinction between construct representation in print and digital environments. (Although

SAILS is delivered now exclusively online, the test maintains similar organization and formatting as when it was delivered on paper.) Based on a strategy of identifying the correct answer, a SAILS item asks the candidate, for instance, to identify the best source of information about an event that took place two days ago by asking the candidate to fill in the bubble next to the term "newspaper." In identifying the answer, the candidate internalizes a correct response by an analytic process of exclusion of incorrect answers. The response, distanced and non-situated, is executed by a single action of identification. In contrast, in the digital environment of iSkills the student is examined on ability to access information experiences, both in a realistically simulated context and in a robust constructed response environment. The task begins with just a description of the customer need and an empty search screen. As the student selects databases, keystrokes search terms, and reviews results over potentially multiple search-and-review cycles, that student engages one aspect of the IL construct in continuous process of mediation (the original approach) and re-mediation (the original approach restructured by the constructed response task). Immersed, the student adopts the persona of an employee of Restoration Demolition (Figure 5.1), a digitally created reality.

In essence, the IL experience is transformed by that created environment. Both the SAILS item and the iSkills constructed response task tap the IL variable of information access but in distinctly different ways. Indeed, the distinction between print and digital environments is also carried into the function of test scoring: SAILS allows only one correct answer; iSkills yields a competency score based on levels of ability encompassing both efficacy of process and correct answer identification.

Contrast of two assessments—the print-based IL assessment at New Jersey Institute of Technology (NJIT, 2012; Scharf, Elliot, Huey, Briller & Joshi, 2007) and iSkills (Katz, 2007b)—demonstrates the distinction between construct assessment in print and digital environments. In 2005, NJIT researchers conducted a study of the relationship between a model of writing informed by Hayes and a model of IL informed by ACRL, as both constructs were represented in a sample of 100 portfolios of senior undergraduate students enrolled in humanities courses. Similar to the curricular project reported by Beth Bensen, Hong Wu, Denise Woetzel, and Ghazala Hashmi (Chapter 19, this collection), the research was the result of collaboration between English faculty and librarians. The overall score on the writing model correlated with the overall score on the IL model at 0.50 ($p < 0.01$), evidence that the two models were related. Based on this print-based system of assessment, NJIT and ETS researchers then collaborated to investigate the relationship between holistic portfolio scores (designed to capture both writing and IL skills) and iSkills scores of students

enrolled in humanities courses (Katz, Elliot, et al., 2008). After controlling for students' overall ability as measured by SAT scores, analysis revealed near zero correlations between the portfolio scores and iSkills scores of first-year students as well as upper-division students. At the time of the study, we concluded that the constructs were related yet distinct. In the print-based environment of the portfolios, students had been asked to read novels, search databases for peer-reviewed articles, and integrate those articles to develop various interpretations of elements within the novels. One of the iSkills digitally based constructed response tasks had asked students to compare advertisements from competing vendors' websites by summarizing information into a table, or students had been asked to present results from a sporting event into a spreadsheet to clarify standings and decide the need for playoffs. From task to scoring, the two assessments could not have been more different. Although the target construct domain of IL was the same, it was mediated by the respective assessments, resulting in different measured constructs.

Based on these two studies, our answer to the second question—how is the construct of IL modeled in digital assessment environments?—is as follows: While identification of the construct of IL, viewed as a system of behavior, can be made in straightforward terms, the observed complexity of the behavior of students as they perform in print and digital environments is a reflection of the complexity of the environment in the assessment. While the core variables—those postulated attributes of Cronbach and Meehl—exist in the larger domain, their representation in the digital environment of the measured construct of assessment systems such as iSkills is unique. The extent to which the construct is digitally mediated depends on the extent to which the assessment leverages, in Bolter and Grusin's terms, the delivery of other media, realistic simulation, and participation in that created reality. And, while the target construct (the intended construct domain) is shared among environments, in digitally mediated environments there may indeed be differences in exactly what is being measured, as the NJIT study demonstrated.

While that somewhat long-winded answer is conceptually useful, institutional assessment staff and their instructional colleagues who are deciding on how to assess IL must consider practical strategies of providing a clearer picture of the underlying construct and its representation in the test at hand. Through validation—the process of gathering evidence for the interpretation and use of an assessment—institutions can begin to make such evaluations.

As we will now show, such evaluations are as nuanced as is the representation of the construct in unique forms of assessment. Through validation, assessment stakeholders can make important decisions leading to claims about the IL of their students. We turn now to present theories of validation and how they can

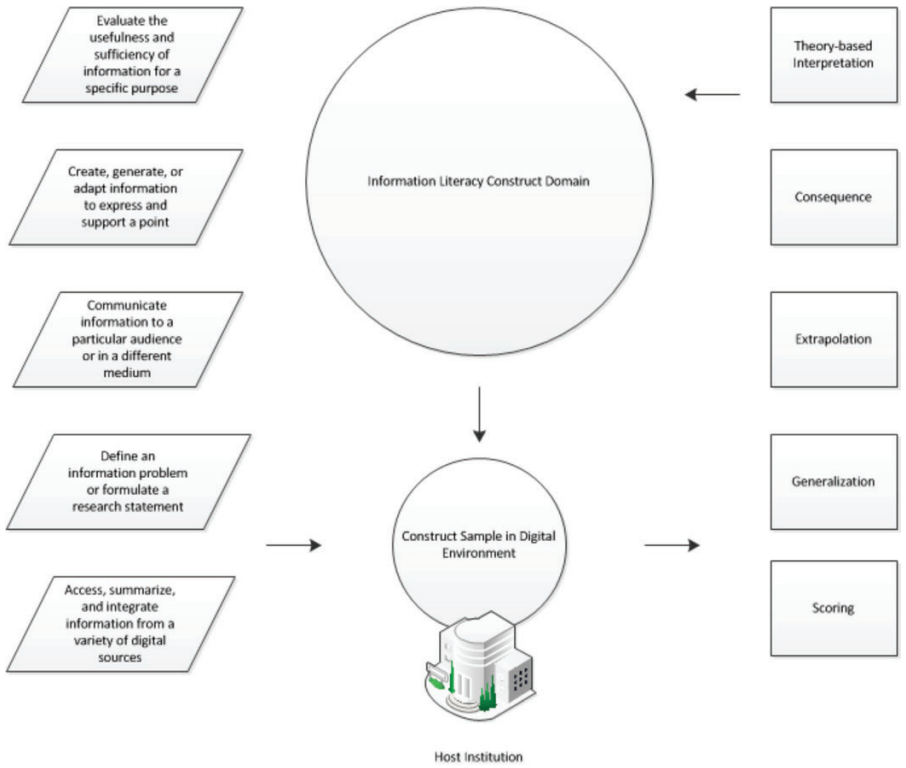

Figure 5.2. Information literacy: A validation model.

support institutional assessment efforts to help students become more information literate.

CONSTRUCT VALIDATION IN DIGITAL ENVIRONMENTS

Michael T. Kane (2013) proposed an Interpretation/Use Argument (IUA) as the most current form of validation. In presenting his case for an evidentiary system of gathering validity evidence, he treated the concept of construct validity in some detail. Tracing the idea of construct representation offered by Cronbach and Meehl (1955), Kane identifies three legacies of their model: test-score interpretations cannot be taken as self-evident; validation is a process that is dependent upon claims made in a defined interpretative framework; and critical inquiry is the appropriate result of validation.

Application of Kane's concepts of representation and validation to the study of IL in digital environments is shown in Figure 5.2.

The large circle represents the construct *domain*—the postulated attribute of developed knowledge, intended to be reflected in test performance—of IL. This circle—the targeted construct—represents the full scope of IL: the not-directly observable knowledge and skills in a student's mind that drive observable performance on any IL-related task, whether real-world activities or assessment tasks. Here we find the work of all who have attempted to define this full construct of IL. Within this circle exists the initial concept of IL found in the *IL Standards*, as well as the *Framework for IL.* Here, too, stand particular institutional interpretations of IL, such as the one established at NJIT, with its emphasis on traits scored in portfolio assessments: citation; evidence of independent research; appropriateness; and integration (Scharf et al., 2007). In the large circle we also find all forms of print and digital assessment—those that have been examined, those that are emerging, and those that can be imagined. Put another way, all the concepts of IL offered in this book fit in the large circle.

The smaller circle represents a single sample of the IL construct as it is reflected in a digital environment. That is, the smaller circle represents the measured construct—those knowledge and skill elements of the range of IL measureable by a digitally based assessment. Note that the smaller circle is drawn from the larger domain; although we would like the measured construct to be a perfect subset, the reality is that the assessed construct might also include knowledge and skills not described in the larger domain. In the case of iSkills, the construct sample includes the five variables of IL that drive the test: access, summarize, and integrate information from a variety of digital sources; define an information problem or formulate a research statement; communicate information to a particular audience or in a different medium; create, generate, or adapt information to express and support a point; and evaluate the usefulness and sufficiency of information for a specific purpose (adapted from ETS, 2002).

The host institution represents the specific site in which the assessment takes place. It is there that the construct takes meaning for users of the assessment. As we will demonstrate, this context shifts the validity framework from that of the assessment designer to that of the assessment user.

Institutional researchers who want to measure and guide improvement of students' IL skills should create their own IUAs that will guide decisions about sampling plans, use of scores, and needed curricular changes. However, because the use of scores for a system such as iSkills—or any test, for that matter—is not simply an up or down vote regarding validity, Kane (2013) offers a process of validation attentive to sources of evidence that guide interpretation and score use. We have illustrated these five sources on the right of Figure 5.2:

1. *Scoring inferences* take us from the observed performances on a test to an observed score. These inferences include evidence about the appropriateness of the scoring criteria and the principles for combining scores. These inferences provide evidence that test scores reflect test performance. This inference might not be as obvious as it might initially seem, as we illustrate below.
2. *Generalization inferences* take us from the observed sample of performances (as reflected in the test score) to claims about expected performance in the construct sample (e.g., that the test score reflects expected performance not only on the current digital IL assessment tasks, but on similar digital IL assessment tasks).
3. *Extrapolation inferences* extend the interpretation into the full construct domain, and are likely among the most common assumption made about test scores—that they reflect actual, real-world ability in the domain of interest.
4. *Consequence inferences* extend the interpretation into the larger assessment environment, thereby strengthening the IUA. (Of course, unintended consequences that threaten validity should be considered and, when possible, avoided.)
5. *Theory-based inferences* extend the interpretation even further, into hypothesized relationships between the construct domain and other areas of interest.

Kane proposed these five categories of evidence as a way to validate the interpretation and use of test scores. We propose that instructors and administrators may use these categories to design a program of research that will yield information about the IL abilities of their students. As we show in the following examples, the extension is not hard to make.

Scoring inferences conceptually refer to the idea that test scores reflect students' performance on the test. Although, traditionally, evidence for this inference includes technical issues such as scoring procedures, in an accreditation context a key factor is motivation: Are students trying their best on an assessment that might have no direct consequences for them? And, if they are not, then how meaningful are the scores themselves? Because more motivated students perform better on such tests (Liu, Bridgeman & Adler, 2012), one type of evidence for sufficient motivation is to investigate the reception of a test by students. Table 5.1 presents a feedback survey (N = 1823) gathered by ETS during early field trials of iSkills. As the responses indicate, the students gave the test their best effort, found it innovative and challenging, and realized that success required both technical and critical thinking skills. The software—the

Table 5.1. iSkills feedback survey: Percentage of responses

Considering the test overall, please indicate how much you agree or disagree with each of the following statements:	N	Agree	Somewhat Agree	Somewhat Disagree	Disagree
I gave this test my best effort.	1823	59%	32%	6%	3%
I have never taken a test like this one before.	1813	77%	15%	4%	4%
This test was appropriately challenging.	1810	53%	35%	9%	3%
The unfamiliar software made it difficult for me to do well on this test.	1804	21%	35%	26%	18%
To perform well on this test requires thinking skills as well as technical skills.	1794	62%	31%	5%	2%
I found the overall testing interface easy to use (even if the tasks themselves might have been difficult).	1800	38%	40%	15%	7%
I enjoyed taking this test.	1804	18%	34%	23%	26%
My performance on this test accurately reflects my ability to solve problems using computers and the Internet.	1801	17%	40%	26%	17%
The tasks reflect activities I have done at school, work, or home.	1803	32%	46%	14%	9%
I encountered a lot of system glitches while taking this test (e.g., system freeze, long time for tasks to load).	1548	23%	25%	19%	34%

digital environment—nevertheless presented problems that may have resulted in interference with construct measurement. Using such information allows a more complete representation of the meaning of the scores themselves.

Evidence for *generalization inferences* would include information about student performance on iSkills in relation to the level of test performance expected to be considered "information literate." A large scale study (N = 1,442) of 14 tasks covering the five variables in Figure 5.2 tells us a good deal about student performance on iSkills.

Figure 5.3 shows the distribution of scores from 1,442 college students and high school seniors. The mean score on this sample was 260, with a standard

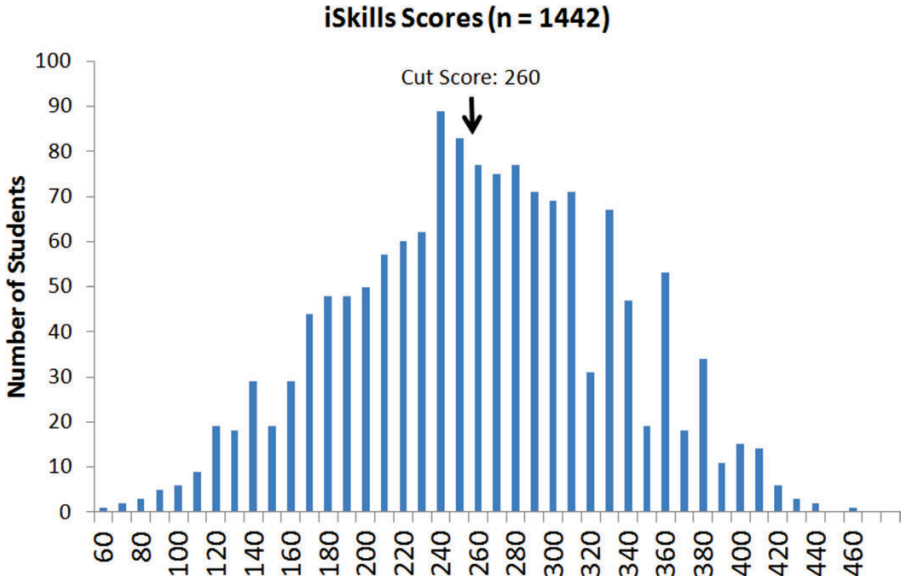

Figure 5.3. iSkills passing rate, April 2011 through March 2012 (N = 1,442).

deviation of 72; scores were approximately normally distributed across the possible score range of 0–500. Approximately 50% of students achieved a level of foundational ICT literacy expected of entering college freshmen. (See Tannenbaum & Katz, 2008, for details on the definition and establishment of the foundational level.) Generalization inferences are supported because the distribution of scores is consistent with what is expected on the distribution of IL. While many librarians would not be surprised at this number, that only half the students "passed" the exam suggests that the so-called digital natives are not uniformly proficient at the effective use of technology (see Katz, 2007a, for more descriptions of strengths and weaknesses of these test takers). At the same time, this percentage is higher than what was observed in previous research (e.g., 27–40% on earlier versions of the assessment; Tannenbaum & Katz, 2008). While we cannot definitively attribute a cause of this rise in passing rates among iSkills test takers, we would hope that the increase is due to increased attention to IL and ICT literacy skills by accreditation agencies as well as by colleges and universities who place these skills in the general education curriculum, such as the work of Alison S. Gregory and Betty L. McCall (Chapter 18, this collection) and Lori Baker and Pam Gladis (Chapter 16, this collection).

However, generalization inferences need not rely on large-scale sampling plans or inferential statistics alone, as a study of 88 undergraduates who responded

to iSkills assessment tasks and to more open-ended "naturalistic" tasks revealed (Snow & Katz, 2009). In that study, analysis of student interviews revealed the response process used to answer both standardized tests and locally developed classroom assignments. In learning more about student response process, instructors can design a curriculum that will allow success on both test types.

Extrapolation inferences would include observations of relationships between performance on iSkills and the performances in a related domain. As noted above, the NJIT portfolio study (Scharf et al., 2007) documented the relationship between writing scores and IL scores. However, the writing scores correlated more strongly with curriculum-related scores (current course grade and overall GPA) than did the IL portfolio scores. The extrapolation inference from these criterion measures allowed instructors to realize that IL was not yet fully integrated into the curriculum and to design a way for librarians to help increase the intensity of coursework in that area. A related study (Katz, Haras & Blaszczynski, 2010) found that iSkills predicted grades in a business writing course, demonstrating the type of connection between IL and business skills explored by Dale Cyphert and Stanley P. Lyle (Chapter 3, this collection). Thus, extrapolation inferences become a significant part of the validation process as IL becomes an enabling construct suggesting across-the-curriculum expansion.

Stemming from the use of iSkills at NJIT and the impact of the assessment on the institution and the assessment itself, identification of *consequential inferences* strengthen the IUA (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014; Haertel, 2013). For NJIT, our studies contrasting the IL construct as measured through print (Scharf et al., 2007) and as measured through a digital environment (Katz, Elliot, et al., 2008) revealed shortcomings in the institutions' view of the IL construct. Accordingly, the iSkills assessment was made part of the NJIT suite of assessments, a decision that strongly reinforced an information-literacy-across-the-curriculum framework that librarians had been building since 2009. This integration had two consequences: digitally based IL became part of the core curriculum for student learning adopted at NJIT; and iSkills served as a key assessment component of the institution's 2012 successful re-accreditation by the Middle States Commission on Higher Education (NJIT, 2012). To bolster student motivation on assessments, the institution is now examining how Certificates of Achievement—awards for predefined performance levels on iSkills—might provide additional motivation for students to try their best. Because enhancing motivation strengthens the validity argument (Liu et al., 2012), case studies from institutions such as NJIT transformed the environment of the assessment itself at ETS and led to certificates awarded for levels of ICT literacy (ETS, 2014).

Theory-based inferences are also of great importance to instructors and administrators as they help make explicit the connection between the construct domain and strategies for curricular change. In the case at hand, a theory of IL postulates the underlying framework that drove Simon's little ant, as well as our students, in certain ways when encountering, respectively, pebbles and constructed response tasks.

An example drives home the importance of theory-building. To investigate the nature of IL, as measured by iSkills, ETS researchers (Katz, Attali & Rijmen, 2008) used factor analysis to identify patterns in a set of items and establish which combinations of items tend to be highly correlated. In the case of iSkills, there were two primary ideal models to consider. First, a seven-factor model arranged the items into the groups corresponding to postulated IL subskills (define, access, evaluate, manage, integrate, create, and communicate; see Katz, 2007b). The iSkills assessment was originally designed with these seven skills, with each task (and the items within a task) corresponding to one of the skills. This model postulates that a student could do well on, say, finding information (access) tasks but do poorly on tasks that require adapting materials to an audience (communicate). Second, the one-factor model took the view that all of the items in the iSkills assessment together measure a single, integrated construct of IL: Students are strong or weak at IL generally, with all of the items on the test being highly correlated (e.g., high performance on one type of item implies high performance on all types of items). Both exploratory and confirmatory factor analyses suggest that IL, as measured by iSkills, consists of a single factor (Katz, Attali, et al., 2008). That is, based on data from a sample of more than 600 test takers, exploratory factor analyses suggested that the entire set of iSkills tasks measure a single, integrated construct: students might have greater or lesser IL, but there was no evidence that the seven IL skill areas were distinct from one another. Similarly, in confirmatory factor analyses, the one-factor ideal model fit the data much better than did the seven-factor ideal model.

What does this research mean for instruction? Interpretatively, it appears that IL is an integrated skill: improving one's IL is a matter of holistic, comprehensive instruction, rather than piecemeal training on component skills. IL appears to be a truly significant threshold concept (Towsend, Brunetti & Hofer, 2011). Such an integrated outlook on IL might reflect either a stronger, more sophisticated view of information generally or a weaker, simplistic view. Of course, instruction cannot ignore the various activities that make up IL skill, as outlined in such documents as the *Framework for IL* or the particular ways that NJIT humanities instructors teach and assess the construct. However, focusing on those foundational skills alone might not be the quickest (or best) path to IL. Instead, a balanced approach that points out the usefulness of more

sophisticated attitudes toward IL might help students recognize the value in, say, trying to figure out alternative descriptors for information (which, in turn, should lead to better search results).

Evidence that IL, as measured by iSkills, is a unified construct impacts how that assessment should be administered. Institutional researchers should explicate theory-based inferences about IL that postulate characteristics of students and their experiences that lead to stronger or weaker IL. Are students who complete a particular set of courses, compared with those that do not, more information literate? Are transfer students entering with weaker IL skills, leading them to struggle in programs compared with students who, from freshmen year, benefit from the university's core curriculum in IL? Which majors tend to have the most information literate students, and is that a function of students who tend to go into that major or a function of the courses in that major? These are just examples of theory-based inferences that could be investigated using an appropriate sampling plan in the administration of IL assessments such as iSkills. They directly tie assessment results to the institutional improvement plans.

Returning to Figure 5.2, we note that the arrows indicate that the interpretation/use argument, and associated evidence, should be used by institutional instructors and administrators to help them reconsider and redefine, as needed, the construct domain itself and the elements of it that are most relevant to their admitted students. Without that feedback loop, the gathered information will only result in reports completed and papers published; with it, stakeholders can work to ensure that the results of the assessment are used to improve learning.

And so we conclude by answering our third question: What do present theories of validation tell us about how institutional assessments might be used to help students become more information literate? Present theories such as IUA reveal the vital importance of a carefully planned program of research, based at the institution, when complex constructs are under examination. In the field of writing studies, such calls for contextualization have been well developed and may serve as basis for IL research (Condon, 2013; Huot, 1996; White, Elliot & Peckham, 2015). In similar fashion, each of the sources of evidence identified by Kane suggests distinct programs of research focusing on areas of validation. Research in these areas provides the level of detail necessary to identify ways to help students improve their IL performance.

Nevertheless, it is an error to conclude with triumphalism because so very much remains to be done. Valuable as it is, the IUA perspective is that of an assessment designer, not an assessment user. For those stakeholders at the host institution shown in Figure 5.2, for example, motivation is of enormous importance. While the assessment designer will justifiably be concerned with technical issues such as scoring procedures, making sure that students are willing to engage

the construct sample of iSkills is of paramount importance. Indeed, cultivating student motivation is one aspect of the assessment over which institutional stakeholders have great influence. As Mariëlle Leijten, Luuk Van Waes, Karen Schriver, and John R. Hayes (2014) have observed of the writing model, however, educators have not adequately learned how to combine motivation with cognitive processes in our construct models in both academic and workplace communication settings. If we follow the recommendations of the National Research Council (2012) and attend to the broad spectrum of cognitive, intrapersonal, and interpersonal domains—as the *Framework for IL* has proposed in its emphasis on cognitive, affective, and social dimensions—we come to realize that we must continue to broaden our investigation of the IL construct mediation and its domain. And, in doing so, we must also continue to conceptualize the IUA perspective in terms of all those who will be influenced by its use: advisory boards, administration, faculty and instructional staff, parents, students, and the public. Depending on audience, the IUA for an assessment may have to be refashioned if it is to have meaning. When perspective is added, we realize that we are only just beginning to understand our parables.

REFERENCES

American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Association for College and Research Libraries (2000). *Information literacy competency standards for higher education.* Chicago, IL: American Library Association.

Association for Colleges and Research Libraries (2015). *Framework for information literacy for higher education.* Chicago, IL: American Library Association.

Baker, L. & Gladis, P. (2016). Moving ahead by looking back: Crafting a framework for sustainable, institutional information literacy. In B. J. D'Angelo, S. Jamieson, B. Maid & J. R. Walker (Eds.), *Information literacy: Research and collaboration across disciplines*. Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.

Bennett, R. E. (1993). On the meanings of constructed response. In R. E. Bennett & W. C. Ward (Eds.), *Construction vs. choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (pp. 1–27). Hillsdale, NJ: Erlbaum.

Bensen, B., Woetzel, D., Wu, H. & Hashmi, G. (2016). Impacting information literacy through alignment, resources, and assessment. In B. J. D'Angelo, S. Jamieson, B. Maid & J. R. Walker (Eds.), *Information literacy: Research and collaboration across disciplines*. Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.

Bolter, J. D. & Grusin, R. (1996). Remediation. *Configurations, 4,* 133–358.

Condon, W. (2013). Large-scale assessment, locally-developed measures, and automated scoring of essays: Fishing for red herrings? *Assessing Writing, 18*, 100–108.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281–302.

Cyphert, D. & Lyle, S. P. (2016). Employer expectations of information literacy: Identifying the skills gap. In B. J. D'Angelo, S. Jamieson, B. Maid & J. R. Walker (Eds.), *Information literacy: Research and collaboration across disciplines.* Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.

Educational Testing Service (2002). *Digital transformation: A framework for ICT literacy.* Princeton, NJ: Author. Retrieved from http://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf.

Educational Testing Service (2014). *iSkills™ Certificates of achievement.* Retrieved from http://www.ets.org/iskills/scores_reports/certificates.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179–197.

Gregory, A. S. & McCall, B. L. (2016). Building critical researchers and writers incrementally: Vital partnerships between faculty and librarians. In B. J. D'Angelo, S. Jamieson, B. Maid & J. R. Walker (Eds.), *Information literacy: Research and collaboration across disciplines*. Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.

Haertel, E. (2013). How is testing supposed to improve schooling? *Measurement: Interdisciplinary Research and Perspectives, 11,* 1–18.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29,* 369–388.

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

Huot, B. (1996). Toward a new theory of writing assessment. *College Composition and Communication, 47*, 549–566.

Jacobson, T. & Gibson, C. (2013). *ACRL information literacy competency standards for higher education task force.* Retrieved from http://connect.ala.org/node/205100.

Kane, M. T. (2013). Validating the interpretation and uses of test scores. *Journal of Educational Measurement, 50,* 1–73.

Katz, I. R. (2007a). ETS research finds college students fall short in demonstrating ICT literacy: National Policy Council to create national standards. *College and Research Libraries News, 68,* 35–37. Retrieved from http://crln.acrl.org/content/68/1/35.full.pdf.

Katz, I. R. (2007b). Testing information literacy in digital environments: ETS's *iSkills* assessment. *Information Technology and Libraries, 26,* 3–12.

Katz, I. R., Attali, Y. & Rijmen, F. (2008, April). *ETS's iSkills™ assessment: Measurement of information and communication technology literacy.* Paper presented at the meeting of the Society for Industrial/Organizational Psychology, San Francisco, CA.

Katz, I. R., Elliot, N., Attali, Y., Scharf, D., Powers, D., Huey, H., . . . Briller, V. (2008). *The assessment of information literacy: A case study* (Research Report RR-08-03). Princeton, NJ: Educational Testing Service.

Katz, I. R., Haras, C. M. & Blaszczynski, C. (2010). Does business writing require information literacy? *Business Communication Quarterly, 73,* 135–149.

Leijten, M., Van Waes L., Schriver, K. & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research, 5,* 285–337.

Liu, O. L., Bridgeman, B. & Adler, R. M. (2012). Measuring outcomes in higher education: Motivation matters. *Educational Researcher, 41,* 352–362.

Maid, B. & D'Angelo, B. J. (2016). Threshold concepts: Integrating and applying information literacy and writing instruction. In B. J. D'Angelo, S. Jamieson, B. Maid & J. R. Walker (Eds.), *Information literacy: Research and collaboration across disciplines.* Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.

National Research Council (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century.* Washington, DC: National Academies Press.

New Jersey Institute of Technology (2012). *A science and technology research university for the 21st Century: Final report.* Retrieved from http://www.njit.edu/middlestates/docs/2012/NJIT_Middle_States_Self_Study_FinalReport.pdf.

O'Connor, L. G., Radcliff, C. J. & Gedeon, J. A. (2002). Applying systems design and item response theory to the problem of measuring information literacy skills. *College and Research Libraries, 63,* 528–543.

Ong, Walter. (1982). *Orality and literacy: The technologizing of the word.* London, UK: Methuen.

Scharf, D., Elliot, N., Huey, H., Briller, V. & Joshi, K. (2007). Direct assessment of information literacy using writing portfolios. *Journal of Academic Librarianship, 33,* 462–477.

Scribner, S. & Cole, M. (1981). *The psychology of literacy.* Cambridge, MA: Harvard University Press.

Simon, H. A. (1996). *The sciences of the artificial* (3rd. ed.). Cambridge, MA: MIT Press.

Snow, E. & Katz, I. R. (2009). Using cognitive interviews to validate an interpretive argument for the ETS™ assessment. *Communications in Information Literacy, 3,* 99–127.

Tannenbaum, R. J. & Katz, I. R. (2008). *Setting standards on the core and advanced iSkills™ assessments* (Research Memorandum RM-08-04). Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/RM-08-04.pdf.

Townsend, L., Brunetti, K. & Hofer, A. R. (2011). Threshold concepts and information literacy. *portal: Libraries and the Academy, 11*, 853–869.

Tucker, B. (2009). The next generation of testing. *Educational Leadership, 67,* 48–53.

White, E. M., Elliot, N. & Peckham, I. (2015). *Very like a whale: The assessment of writing programs.* Logan, UT: Utah State University Press.

Woods, C. (2010). *Visible language.* Chicago, IL: University of Chicago Press.

Yancey, K. B. (2016). Creating and exploring new worlds: Web 2.0, information literacy, and the ways we know. In B. J. D'Angelo, S. Jamieson, B. Maid & J. R. Walker (Eds.), *Information literacy: Research and collaboration across disciplines.* Fort Collins, CO: WAC Clearinghouse and University Press of Colorado.