**Integration of Evidence-Centered Design and Universal Design Principles in Developing Information and Communication Literacy Assessment**

***Dr. Mark Zelman*** *World Bank consultant, United States (marknjpr@yahoo.com)*

***Dr. Svetlana Avdeeva*** *National Training Foundation (NTF), Russian Federation (Avdeeva@ntf.ru)*

***Dr. Tigran Shmis*** *World Bank, Russian Federation (tshmis@worldbank.org)*

***Dr. Kirill Vasiliev*** *World Bank, Russian Federation (*[*kvasiliev@worldbank.org*](mailto:kvasiliev@worldbank.org)*)*

**Keywords**: ICT Literacy, Information Literacy, Psychometrics, International comparison

* **ABSTRACT**
* The paper is based on the innovative information and communication
* literacy measure (ICT) that that was developed within READ[[1]](#footnote-1) program in Russia with a support from the World Bank and the Center for International Cooperation in Education Development ([www.ciced.org](http://www.ciced.org)). The ICT literacy tool is an attempt to develop contemporary problem-based and scenario-based approach to ICT literacy assessment for multinational secondary school environment.

One of several research questions the ICT literacy assessment sought to provide answers to is what variations exist between participating countries, in student ICT literacy? The answer to this research question requires that students with disabilities be included in country’s assessments. However, the inclusion of students with disabilities has given rise to a number of issues related to fairness and validity. Recently explorations of "universal design" have began to find whether tests can be developed to be more accessible and valid for a wider range of test- takers.

This paper introduces evidence-centered design (ECD) and the key principles of “universal design”. The paper then shows how the “universal design” principles were paired with the ECD for ICT literacy assessment.

***1.0 Introduction***

Under the ICT Literacy project we are studying the use of universal design paired with an approach termed "evidence-centered design" (ECD) to develop further and refine ICT literacy test tasks that can more accurately evaluate the information literacy of all students on international assessments. The target population for this study is 9th grade students, but if successful the approach can be applied to other age ranges. This paper introduces background on evidence-centered design, an ECD designed ICT assessment, and the key principles of universal design. The paper then illustrates how the universal design principles could be incorporated into the ICT assessment assessment system *design patterns*.

***2.0 Evidence-Centered Design (ECD)***

The foundations of ECD stem from validity theory (Messick, 1989),  psychometrics (Mislevy, 1994), philosophy (Toulmin, 1958), and jurisprudence (Wigmore, 1937). ECD also builds on developments in fields of expert systems (Breese, Goldman, & Wellman, 1994), software design (Gamma, Helm, Johnson, & Vlissides, 1994), and legal argumentation (Tillers & Schum, 1991). Robert Mislevy, Linda Steinberg, and Russell Almond (2003) at Educational Testing Service operationalized ECD and introduced tools for building assessment arguments that could be used both in designing new assessments and understanding familiar ones. ECD could be described firstly by looking at assessment as an argument from the evidentiary point of view and making explicit the claims (Hansen & Mislevy, 2007). Secondly it could be looked at from the point of view that distinguishes layers at which activities and structures appear in the assessment, showing an assessment argument in operational processes. By making the evidentiary argument more explicit, the ECD makes operational elements more amenable to reviewing and revising. Making the evidentiary argument more explicit also helps test developers meet a range of assessment needs caused by changing technological, social, and legal environments (Hansen & Mislevy, 2007).

ICT Literacy project uses ECD that is expressed in layers design. The ECD layers design center around four key questions:

1. *Domain Analysis* layer*:* Research and experience about the domains and skills of interest are gathered—information about the skills, abilities and the ways test takers acquire KSAs. This lays the foundations and goals of the assessment’s architecture.

2. *Domain Modeling* layer*:* information from *Domain Analysis layer* is organized to form the assessment argument. What KSA’s of test takers do we want to measure to make appropriate claims from the assessment? This layer defines the constructs we are measuring in the ICT assessment and their interrelationships. It is narrative in structure and serves as a ground to the more technical models in the next Conceptual Assessment Framework (CAF) layer.

*3. CAF layer:* Comprised of the Proficiency Model, Evidence Model (where observations of behavior are considered as evidence of proficiency), and Task Model (where we designing situations that will elicit the observable evidence needed). Additionally, the Assembly Model determines how ICT tasks are assembled into tests. Figure 1 below is a schematic representation of the conceptual assessment framework (CAF) for ECD used in designing ICT literacy assessment.



Figure 1. Conceptual assessment framework (CAF).

The CAF is quite general and can be used to describe assessments developed with a number of different methodologies. In ICT literacy assessment design we have used Bayesian networks (Mislevy, 1995), though the fusion model (Hartz, et al., 2002) also fits into this framework.

The described ECD layer approach encourages test developers to design ICT assessment that has:

* *Clarity of purpose* – representation of assessment goals and the relevance of design decisions to those goals
* *Interrelated design* – modeling the interactions of design decisions and how changes in one aspect of design affect other design elements.
* *Evidentiary requirements* – explication of what constitutes relevant evidence of ability and how such evidence bears on assessment-based decision making.
* *Validity* – a documented chain of reasoning and rationale underlying design decisions and their relevance to the criterion of interest.
* *Innovation* – a guide for developing assessments targeting elusive domain constructs or using emerging technologies and new item types.

***2.1 ICT Literacy Assessment System***

Information Literacy Assessment design is a project supported by the World Bank READ project. The ICT literacy project has developed a design framework for assessment tasks, based on the evidence-centered design (ECD) framework. ICT literacy assessment system was developed as a system for designing blueprints for ICT assessment tasks, with a particular concentration toward ICT literacy tasks, high order thinking and reading skills, and building and using task models.

ECD used in ICT literacy assessment seeks to integrate the processes of assessment design, authoring, delivery, scoring, and reporting. Thus work within ICT Literacy Assessment is focused on design layers that generally lie outside of the level of specific environments for task authoring and test delivery.

ICT assessment is also based on assessment *design patterns* analogous to those in architecture and software engineering. It captures capture design rationale in a reproductive form in the *Domain Modeling* layer of assessment.

In a designpattern for ICT literacy assessment*,* key attributes, namely Focal KSAs, Characteristic Features, and Variable Features, are used for building the assessment argument for test takers. As described by Hansen and Mislevy (2007), these key attributes are:

1. **Focal and Additional KSAs** consist of the primary knowledge/skills/abilities of test takers that are addressed by assessment (Mislevy et al., 2003). It helps to addressing Comparability of scores between individuals with different language and cultural background, for example.

3. **Characteristic Features.** Characteristic Features of the assessment consist of the feature that must be present in a situation in order to evoke the desired evidence about the Focal KSAs (Mislevy et al., 2003).

4. **Variable Features.** Variable Features are described as features that can be varied to shift the difficulty or focus of tasks (Mislevy et al., 2003). Variable Features in ICT literacy assessment have a particular significant role with respect to test taker sub-populations (e.g., speakers of different languages).

***3.0 Universal Design***

Universal Design (UD) helps us to meet the challenge of diversity population by suggesting flexible assessment materials, techniques, and strategies (Dolan, Rose, Burling, Harris & Way, 2007). Listed below are the three primary principles that guide UD and provide structure for the infusion of UD into the ECD *design patterns:*

*Principle I. Provide Multiple Means of Representation*

Students differ in the ways that they perceive and comprehend information in the ICT literacy test that is presented to them. For example, some test takers may simply comprehend information better through visual or auditory means rather than printed text. Also there are language or cultural differences, and so forth may all require different ways of approaching content. In reality, there is probably no single way of representation that will be optimal for all students; providing options in representation is essential.

*Principle II: Provide Multiple Means of Expression.*

Test takers differ in the ways that they can navigate a learning environment and express what they know. For example, some may be able to express themselves well in writing text but not oral speech, and vice versa.

*Principle III: Provide Multiple Means of Engagement*

Students differ markedly in the ways in which they can be engaged and or motivated to learn and to do well on a test. Some test takers are highly engaged during testing by spontaneity and novelty while other are frightened and disengaged.

***4.0 Design Patterns for ICT Literacy Assessment***

In the framework of the READ, ICT project we have attempted to introduce new Universal Design-infused ICT *design patterns.*

The Proficiency Model in CAF serves as the basis for developing such *design patterns.* The basis for exemplar design patternswere those that (1) would be useful and important for a particular target population; (2) raised questions such as model-based reasoning; and (3) showed connections with the ICT standards of a participating country.

***5.0 Infusing UD into ECD Design Patterns***

Based on this analysis of ECD design of ICT literacy test and UD principles, the categories derived from UD’s three principals will be used to categorize types of construct-irrelevant solutions that are likely to influence student performance. Definitions of UD categories are provided in Table 1.

The test development team will be adding additional UD-based Variable Task Features into ICT literacy test *design patterns* that were motivated by these categories of additional KSAs within each of the UD categories. Once it will be determined which of the categories are be challenging for certain test takers, test developers will be supporting these different categories through a variety of Variable Task Features. For example, if vocabulary is a Non-Focal KSA, the test developer will draw from the list of Variable Task Features for Language and Symbols in order to find strategies to support vocabulary (e.g., embedded support for key terms through the use of a technical glossary, footnotes to definitions, etc.). (See Figure 2 below).

|  |  |  |  |
| --- | --- | --- | --- |
| ***Principle 1: Multiple Representations*** | | |  |
| ***UD category*** | ***Definition*** | ***Variable Features*** | ***ICT Design Pattern Title*** |
| ***Perceptual*** | Multiple representations of information to ensure that the information is accessible to all test takers: 1) providing the same information through different sensory modalities; 2) providing information in a format that will allow for adjustability by the user (e.g., text that can be enlarged, sounds that can be amplified). | Representational Format   * Flexible size of text and images * Flexible amplitude of speech or sound * Adjustable contrast * Flexible colors   (2): Auditory Information   * Text equivalents (e.g. captions, automated speech to text) * Visual graphics or outlines * Virtual manipulatives, video animation    Verbal descriptions  (3): Visual Information   * Spoken equivalents for text and images * Automatic text to speech | Critiquing the work of others with topic states of matter:  The test taker evaluates the work of others to determine evidence that supports or contradicts the results, identifying faulty reasoning or conclusions that go beyond the evidence and/or are not supported by data. |
| ***Language*** | Students vary in their ability to work with both linguistic and non-linguistic. Thus an important assessment strategy is to ensure that alternative representations are provided for accessibility, clarity and comprehensibility of language for all test takers. | (1): Supports for Vocabulary and Symbols   * Embedded support for key terms (e.g. technical glossary, hyperlinks/ footnotes to definitions, illustrations, background knowledge) * Embedded support for non-technical terms (e.g. non-technical glossary, hyperlinks/ * Embedded alternatives for unfamiliar references (e.g. domain specific notation, jargon, figurative language, etc.)   (2): Supports for Syntactic Skills and Underlying Structure   * Alternate syntactic levels (simplified text) * Grammar aids * Highlighted syntactical elements (e.g. subjects, predicates, noun-verb agreement, etc.) * Highlight structural relations or make them more explicit   (3): Supports for the Target Language   * All key information in the dominant language is also available in prevalent first languages for second language learners * Key vocabulary words have links to both dominant and non-dominant definitions * Domain-specific vocabulary is translated for both special and common meanings * Electronic translation tools, multi-lingual glossaries   (4): Supports for Decoding and Fluency   * Digital text with automatic text to speech |
| ***Cognitive*** | Test takers differ greatly in their information processing skills. Proper design and presentation of information is needed to ensure that assessments accurately measure student ICT literacy skills. | (1): Options that Guide Information Processing   * Explicit prompts for each step in a sequential process * Interactive models that guide exploration and inspection * Graduated scaffolds that support information processing strategies * Multiple entry points and optional pathways through content * Chunking information into smaller elements, progressive release of information, sequential highlighting * Complexity of the scenario * Cognitive complexity * If selected response, distractors based on misconceptions/typical errors vs. non-misconceptions |  |

Figure 2. Infusing UD into ICT Assessment Design Patterns

***6.0 Conclusions***

We showed a systematic assessment design process that is leading us to an assessment that targets ICT literacy in a way believed to be predictive not only of grade 9 students, for whom assessment was initially design, but also for workplace. By focusing on critical thinking -- rather than simple procedures -- the ICT literacy assessment seeks to identify individuals that can adapt to ever-changing technological environments.

***References***

Breese, J. S., Goldman, R. P., & Wellman, M. P. (1994). Introduction to the special section on knowledge- based construction of probabilistic and decision models. *IEEE Transactions on Systems, Man, and Cybernetics, 24,* 1577-1579.

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

Dolan, R. P., Rose, D. H., Burling, K., Harms, M., & Way, D. (April, 2007). The Universal Design for Computer-Based Testing Framework: A Structure for Developing Guidelines for Constructing Innovative Computer-Administered Tests. Paper presented at the National Council on Measurement in Education Annual Meeting, Chicago, IL.

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). *Design patterns.* Reading, MA: Addison- Wesley.

Hansen, E. G., & Mislevy, R. J. (2007). Design Pattern for Improving Accessibility for Test Takers With Disabilities. Draft.

Kane, M. (1992). An argument-based approach to validation. *Psychological Bulletin, 112,* 527-535. Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 32(2),* 13-23.

Mislevy, R. J., & Haertel, G. (2006). Implications for evidence-centered design for educational assessment. *Educational Measurement: Issues and Practice, 4,* 6-20.

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

1. The World Bank and Russia are jointly implementing READ, a program to support improvements in quality of education. The READ Program includes (i) the READ Trust Fund for eight countries, and (ii) the READ capacity-building services program for Russia. [↑](#footnote-ref-1)