# Annotated Reading List for ECD
Bob Mislevy
August 4, 2010

Evidence-Centered Design (ECD) of assessment is a conceptual framework for the design and delivery of educational assessments. This is a reading list that supplements the technical report series from the NSF-supported PADI projects, found at http://padi.sri.com/publications.html and http://ecd.sri.com/publications.html.

**Best place to start for someone new to ECD:**

Mislevy, R.J., Almond, R.G, & Lukas, J. (2004). A brief introduction to evidence-centered design. *CSE Technical Report*. Los Angeles: The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. http://www.cse.ucla.edu/products/reports/r632.pdf
[Readable introduction to the basic ideas and structures of evidence-centered design. Contains an annotated bibliography.]

**Good ones to read re ECD and the Cisco Packet-Tracer simulation assessment:**

*Psychometric and Evidentiary Approaches to Simulation Assessment in Packet Tracer Software.*
http://www.education.umd.edu/EDMS/mislevy/CiscoPapers/PsychometricEvidentiaryPacketTracer_Frezzoetal(2009)_FINAL.pdf
Dennis C. Frezzo, John T. Behrens, Robert J. Mislevy, Patti West, and Kristen E. DiCerbo. Presented at the Fifth International Conference on Autonomic and Autonomous Systems , April 20-25, 2009, Valencia, Spain. [Good discussion of PT tasks in terms of the four-process delivery system.]

*Design Patterns for Learning and Assessment: Facilitating the Introduction of a Complex Simulation-Based Learning Environment into a Community of Instructors.*
http://www.education.umd.edu/EDMS/mislevy/CiscoPapers/DPs_for_Learning.pdf
Dennis Frezzo, John Behrens, and Robert Mislevy. Presented at the Annual Meeting of the American Educational Research Association, Chicago, Il, April, 2007.

**A general presentation aimed at test developers:**

Mislevy, R.J., & Riconscente, M.M. (2006). Evidence-centered assessment design: Layers, concepts, and terminology. In S. Downing & T. Haladyna (Eds.), *Handbook of Test Development* (pp. 61-90). Mahwah, NJ: Erlbaum. Research report version available online as *PADI Technical Report 9*. Menlo Park, CA: SRI International: http://padi.sri.com/downloads/TR9_ECD.pdf [Good discussion of the layers in ECD. Uses the PADI project's ECD representations and some terminology]

**More on applications in Cisco context:**

In the early 1990s, we built the NetPass prototype simulation-based assessment, an ancestor of Panther.  Volume 4, No. 4, 2004 of the *International Journal of Testing* was devoted to this project.  There were four articles:

*Introduction*
http://www.education.umd.edu/EDMS/mislevy/CLI/Introduction.pdf
"Introduction to Evidence Centered Design and Lessons Learned From Its Application in a Global E-learning Program." John T. Behrens, Cisco Systems, Robert J. Mislevy, University of Maryland, Malcolm Bauer and David M. Williamson, Educational Testing Service, and Roy Levy, University of Maryland.

*Design Rationale*
http://www.education.umd.edu/EDMS/mislevy/CLI/DesignRationale.pdf
"Design Rationale for a Complex Performance Assessment." David M. Williamson and Malcolm Bauer, Educational Testing Service, Linda S. Steinberg, University of Pennsylvania, Robert J. Mislevy, University of Maryland, and John T. Behrens, Cisco Systems, Inc.

*Measurement Model*
http://www.education.umd.edu/EDMS/mislevy/CLI/MeasurementModel.pdf
"Specifying and Refining a Measurement Model for a Computer Based Interactive Assessment." Roy Levy and Robert J. Mislevy, University of Maryland.

*Statistical Natural Language Processing*
http://www.education.umd.edu/EDMS/mislevy/CLI/NaturalLanguageProcessing.pdf   "Using Statistical Natural Language Processing for Understanding Complex Responses to Free-Response Tasks ." Sarah F. DeMark and John T. Behrens, Cisco Systems, Inc.

*On the Roles of External Knowledge Representations in Assessment Design.*
http://www.education.umd.edu/EDMS/mislevy/papers/KR_8-22-07.doc
Robert J. Mislevy, John T. Behrens, Randy E. Bennett, Sarah F. Demark, Dennis C. Frezzo, Roy Levy, Daniel H.  Robinson, Daisy Wise Rutstein, Valerie J. Shute, Ken Stanley, Fielding I. Winters.  [The focus is on knowledge representations, but there is a good discussion of the CNS simulation-based tasks from the perspective of ECD as the final example.]

*Activity and Assessment Theory in the Design and Understanding of the Packet Tracer Ecosystem.*  http://ijlm.net/knowinganddoing/10.1162/ijlm.2009.0015   Dennis C. Frezzo, John T. Behrens, and Robert J. Mislevy. *International Journal of Learning and Media,* Spring 2009, Vol. 1, No. 2.

**More technical details:**

Mislevy, R.J., Steinberg, L.S., & Almond, R.G.  (2002). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3-67. Also available as *CSE Technical Report 543*.  Los Angeles:  The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. http://www.cse.ucla.edu/CRESST/Reports/TECH597.pdf .  [A very comprehensive, and rather technical, discussion of evidence centered design, spanning assessment arguments, to design elements, to delivery system architecture, and the connections within and across these levels.]

Almond, R.G., Steinberg, L.S., & Mislevy, R.J. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *Journal of Technology, Learning, and Assessment*, *1*(5). http://www.bc.edu/research/intasc/jtla/journal/v1n5.shtml  Also available as *CSE Technical Report 543*.  Los Angeles:  The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. http://www.cse.ucla.edu/CRESST/Reports/TECH543.pdf. [Extended discussion of the four-process delivery system architecture, including explanation of relationships between the design objects of the conceptual assessment framework and the processes and messages in an assessment delivery system.]

Levy, R., &  Mislevy, R.J. (2003).  Specifying and refining a measurement model for a simulation-based assessment. *CSE Technical Report 619*.  Los Angeles: The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. http://www.cse.ucla.edu/reports/R619.pdf  [Focus on estimation of conditional probability models in the Bayes net psychometric model in the Cisco Learning Institute's NetPASS prototype assessment.  A very technical psychometric paper.]

**Some applications in other domains:**

Mislevy, R.J., Steinberg, L.S., & Almond, R.A. (2002).  Design and analysis in task-based language assessment. *Language Assessment, 19,* 477-496. Also available as *CSE Technical Report 579*.  Los Angeles:  The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA.  http://www.cse.ucla.edu/CRESST/Reports/TR579.pdf. [ECD perspective on designing task-based language assessments.  Fairly readable. Includes examples of Bayes nets for tasks that tap multiple aspects or knowledge and skill.  ]

Mislevy, R.J., Steinberg, L.S., Breyer, F.J., Almond, R.G., & Johnson, L. (2002). Making sense of data from complex assessment.  *Applied Measurement in Education, 15,* 363-378. Also available as *CSE Technical Report 538*.  Los Angeles:  The National Center for Research on Evaluation, Standards, Student Testing (CRESST),

Center for Studies in Education, UCLA.
http://www.cse.ucla.edu/CRESST/Reports/RML%20TR%20538.pdf. [Argument that
the way to design and analyze complex assessments, such as computer-based
simulations, is from the perspective of the evidentiary argument--not from the
perspective of technology. Ideas are illustrated in some detail with the DISC
prototype assessment of problem-solving in dental hygiene.]

Steinberg, L.S., Mislevy, R.J., Almond, R.G., Baird, A.B., Cahallan, C., DiBello, L.V.,
Senturk, D., Yan, D., Chernick, H., & Kindfield, A.C.H. (2003). *Introduction to the
Biomass project: An illustration of evidence-centered assessment design  and delivery
capability*. CSE Technical Report #609. Los Angeles, CA: UCLA Center for the
Study of Evaluation. http://www.cse.ucla.edu/reports/R609.pdf  [Design rationale
for a standards-based, web-delivered assessment of science inquiry, in the areas of
transmission genetics and microevolution. Much discussion of working with experts
and National Science Education Standards, to carry out the ECD design work and
then implement a prototype assessment at the level of secondary science.]