Dr. Dorry M. Kenyon, Center for Applied Linguistics, Washington, DC
dkenyon@cal.org

From Test Development to Test Use Consequences: What Roles does the CEFR Play in a Validity Argument?

The development and validation of large-scale language tests is a complicated endeavor, consisting of multiple layers of activity and involving multidisciplinary teams. Within each layer of activity, precedence must be given to the claims that will be made about the defensible interpretations of test scores, appropriate uses of test scores, and evaluation of the consequences of those uses.

Extending the *Assessment Use Argument* (AUA) of Bachman and Palmer (2010) and the *Interpretation/Use Argument* (IUA) of Kane (2013) and integrating it with the tenets of Evidence-Centered Design (ECD) by Mislevy and colleagues (for example, Mislevy, Steinberg and Almond, 2002; Mislevy and Yin, 2012), language testing experts at the Center for Applied Linguistics are developing an integrated validation argument framework. The goal of this framework is to help language testers and their colleagues across disciplines gain a complete picture of the interaction of all aspects of the language testing endeavor.

In this talk I will outline the layers of this integrated framework, illustrating in particular how it helps test developers clarify the role of proficiency level descriptions, such as embodied in the CEFR and other descriptions of developing language proficiency. For example, while standard-setting procedures such as described in the *Manual for Relating Examinations to the CEFR* may be useful to provide evidence to link claims about the interpretation of performances on test to the CEFR, the integrated framework clarifies how linkages can and should be related to many other layers of a test validation argument, beginning with foundational layers of domain analyses and description. In doing so I will illustrate the usefulness of an integrated validation argument framework in conceptualizing test development projects, communicating internally to multidisciplinary teams involved in a language test development project, and in communicating externally to all stakeholders.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice.* Oxford: Oxford University Press.

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

Mislevy, R.J., Steinberg, L.S. & Almond, R.G. (2002). Design and analysis in task-based language assessment. *Language Testing*, 19 (4), 477-496.

Mislevy, R. J., & Yin, C. (2012). Evidence-centered design in language testing. In G. Fulcher & F. Davidson (Eds.), *The routledge handbook of language testing* (pp. 208-222). Abingdon, United Kingdom: Routledge.