Micheline Chalhoub-Deville (University of North Carolina at Greensboro)
Context/task-centered interactional second/foreign language testing constructs and validation

A variety of disciplines shape theories and practices in the second/foreign (L2) language testing field. With regard to conceptualizations of the L2 constructs, language testing has largely adapted representations from fields such as psychology, linguistics, and applied linguistics. Language testing constructs have been portrayed in terms of complex abilities, skills, performances, and hierarchies. While representations of the L2 construct have varied over the years, they have remained primarily cognitive in terms of their underpinnings. With a cognitive orientation, the focus is on representing an individualized, generalized performance with a stable core. It is fairly reasonable to state that cognitive-based explorations have resulted in general models of abilities that have served the L2 testing field well. These models have enriched our knowledge base as well as our practices, and have helped advance the profession as a whole. However, research findings with respect to task specificity, test method effect, language variability, and interactional communication prompt alternative conceptualizations of the L2 construct. Emerging scholarship with a sociocognitive orientation pushes language testing professionals to attend to the constructs of L2 performance in particular contexts. It is expected that this type of research will ultimately yield knowledge, which will impact theory formulations and will allow researchers to better control and manipulate interactional parameters as needed for operational testing purposes.

Educational measurement is another major discipline that has shaped knowledge and practices in language testing. The state of knowledge in areas such as validity and validation has played a critical role in directing language testing research and development efforts. Traditionally and in keeping with trends in educational measurement (e.g., the Standards--AERA, APA, & NCME, 1985, 1999, 2014), language testing validation research has underscored psychometrically-driven, cognitively-oriented practices. McNamara and Roever (2006) point out that research available in the language testing field seems to focus more on the technical aspects of validity even in exploration of consequences, where the focus is on technical aspects of fairness, e.g., DIF analyses, as contrasted to sociopolitical and educational impact. Social, educational, and policy explorations, i.e., "The overt social context of testing" (McNamara and Roever, 2006, p. 14), are less endeavored. McNamara and Roever (2006) reason that this orientation to validity is "heavily marked by its origins in the individualist and cognitively oriented field of psychology" (p. 9). Such a characterization is echoed in educational measurement by Haertel (cited in Sireci, 2013), who holds that measurement has its roots in psychology, which focuses on individual differences. These historical views and practices, however, are beginning to alter. Emerging research, which is informed by the push in recent years to use testing as part of a reform agenda and is guided by theory of action (TOA), points to the need to expand the psychometrically- and cognitively-oriented validity research to include socially-based documentation.

The presentation reflects on the theoretical underpinnings of the L2 construct in language testing and invites the profession to consider the performance of tasks from a more sociocognitive orientation. The article also explores how score interpretation and use, i.e., validity could be more socially-grounded.