

The early to mid-1990s saw assessment designers increasingly focused on developing performance assessment tasks that measured both content knowledge and inquiry skills as they are situated in authentic, "real-world" contexts. At the same time, many assessment experts, including Samuel Messick, became increasingly concerned that the focus on building authenticity into performance tasks could introduce demands for extraneous knowledge or skills as well, making it more difficult to draw valid inferences about the focal knowledge and skills. In Messick's 1994 article, *The interplay of evidence and consequences in the validation of performance assessments*, he addresses this issue by arguing for a more deliberative approach to assessment design that included a specific focus on three main components:

- What complex of knowledge, skills, or other attributes should be assessed?
- What behaviors or performances should reveal those constructs?
- What tasks or situations should elicit those behaviors?

Reference

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