
COMMENTARIES

A Commentary on *Clarifying Inputs and Outputs of Cognitive Assessments*

Fen-Lan Tseng
*Research Center for Psychological and Educational Testing
National Taiwan Normal University*

I would like to thank Professor Schafer (2023) for composing this thoughtful and timely piece, which explores various aspects of cognitive assessments and provides valuable insights into the challenges and potential solutions related to their communication with different stakeholders. It effectively addresses the evolving issues associated with standardized testing in the United States and emphasizes the need to enhance public understanding of these assessments.

Schafer (2023) begins by acknowledging the major issues associated with standardized testing in the United States. Historically, a significant point of contention regarding standardized testing has been the excessive focus on rote memorization and narrowing the curriculum to include only concepts and procedures directly assessed. This critique arises from concerns that such an approach may lead to students receiving a shallow and unbalanced educational experience. The author rightly recognizes that while this issue is important, it

is just one aspect of the larger discussion about cognitive assessments.

Recently, there has been a growing concern about perceived biases in assessment outcomes. Schafer has appropriately highlighted this concern as it underscores the importance of fair testing practices. In an increasingly diverse society, assessments must accurately measure various abilities and skills without introducing biases based on factors such as race, gender, economic status, or cultural background. This awareness of bias represents a developmental aspect of the discussion related to cognitive assessments.

The central argument of Schafer (2023) revolves around the idea that improving public understanding of cognitive assessments and their broader impact is crucial. The author advocates for transparency as the key to achieving this understanding and presents two major strategies to enhance transparency.

The first strategy focuses on clarity regarding the domains of major standardized

tests in relation to the curriculum. This includes clear and concise descriptions of the content covered by these tests, effectively eliminating ambiguity and promoting a deeper understanding of the assessment purposes and relevance. This suggestion aligns with the broader transparency movement in the education sector, where stakeholders demand a clear connection between educational/curriculum content and test content. By establishing this connection, educational institutions can ensure that their assessments are meaningful and relevant to the curriculum.

The author suggested an approach which is known as a heuristic in Schafer and Moody (2004) as the lowest level of specificity in a test map and also provided criteria to achieve a balance between too much specificity and too much generality in heuristics (Schafer & Moody, 2019). Most importantly, the link between heuristics and test items can be used to document whether the test does or does not carefully represent the domain/curriculum that is intended. For a high-stakes standardized test, if instructors teach in a well-designed domain of heuristics, they will be teaching the intended curriculum. In this situation, a positive washback effect (Madaus & Kellaghan, 1992; Messick, 1996) of educational accountability testing would happen as expected. Just as mentioned by the author, "Teaching to the test domain is a good thing as long as the domain is worth teaching to."

Though providing clear and concise descriptions of the test to establish a direct link with the curriculum sounds like a reasonable proposition, it may not be practical and easy in many cases. In some big countries, educational content may vary from one region to another and even from one school district to another. Creating domain descriptions that universally align with all curricula is a monumental task, even with the establishment of the Common Core State Standards Initiative, as in the States (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Moreover, school education

is never static but dynamic, with curricula evolving over time. Maintaining such alignment would require continuous updates, which could be resource-intensive.

The second proposal introduces the concept of a web-based tool that allows users to customize the background information for assessment results. This approach provides flexibility, enabling different stakeholders to interpret data according to their specific needs and perspectives. Whether users prefer to use norm-referenced or criterion-referenced information, this tool allows them to explore test results in the way that makes the most sense to them. This proposal represents the significant leap in the universal availability and feasibility in the provision of assessment data for multiple stakeholders.

Furthermore, Schafer (2023) provides a comprehensive view of the practical applications of this tool for various stakeholders. Students, parents, teachers, administrators, and policymakers can all benefit from the ability to customize and interpret assessment data. For students, this means a clearer understanding of their strengths and areas for improvement. Parents can use this information effectively to support their children's education. Teachers gain valuable insights into student performance, enabling them to adjust their teaching strategies accordingly. Administrators can make informed decisions about resource allocation, and policymakers can use the data to guide the future direction of educational policies.

Schafer's emphasis on transparency and clarity is a much-welcomed shift from historical opacity in the assessment process. It underscores the importance of test developers adopting tools and strategies that elucidate and document the assessment process. This kind of transparency is not just a public relations issue; it is critical for ensuring the continued and expanded use of cognitive assessments in education and the professional field.

While the concept of a web-based tool for customized interpretation of assessment results

is intriguing, its practicality and effectiveness remain questionable. Effective use of such a tool would require a certain level of data literacy and understanding of statistics from the users. This could be a significant barrier for many parents, students, and even some educators. Moreover, the utility of the tool depends heavily on the accuracy and reliability of the underlying data. If the assessments carry flaws in themselves, no amount of customization will make the results more meaningful.

The concerns expressed by a portion of the general public regarding public examinations may be challenging to completely eliminate. Therefore, I would like to propose that we focus on encouraging the public to place greater emphasis on a higher-level cognitive assessment. Research has demonstrated that formative classroom assessment has not only immediate, visible impacts on students' learning but also long-term, intangible effects, including motivation, learning strategies, and attitudes (Crooks, 1988; Earl, 2013). It appears to me that by aligning various classroom assessments and large-scale high-stakes assessments with the official or common curriculum, we can address the concerns raised by Schafer (2023) regarding standardized, large-scale assessments' impact on teaching and learning, if not entirely mitigate them.

To accomplish this goal, a crucial approach is to enhance the assessment literacy of individuals involved in pre-service or in-service teacher training programs. Every teacher should recognize that well-designed assessments can foster assessment for learning. Classroom assessment should be supportive and encourage students to take the initiative in their learning. Students should not feel compelled to prepare for exams solely to achieve high scores. Instead, they should be taught to use assessment as a tool for their own advancement. The primary purpose of assessment, whether in the classroom or on a larger scale, is to align with the curriculum and aid students in learning and improving through the assessment process and its outcomes rather than merely

comparing students' abilities. In other words, we should avoid confining the assessment's purpose to determining students' final grades and acknowledge its potential to facilitate more active and effective learning (Black & Wiliam, 1998; Earl, 2013).

In conclusion, Schafer (2023) provides some practical solutions to address various challenges related to cognitive assessments. It recognizes that standardized testing is not a single-dimensional issue but a complex field with evolving problems. By advocating for transparency, clarity, and flexibility, Schafer has offered a hopeful blueprint for the future of standardized cognitive assessments. However, it still requires further in-depth research and implementation plans to ensure the transparency and clarity of assessment results. For instance, issues related to test validity, reliability, and the potential misinterpretation and misuse of test scores are equally critical. Schafer's narrow focus may inadvertently downplay these issues, leaving them under-addressed. A more comprehensive examination of standardized testing would provide a more accurate picture to guarantee the accuracy and effectiveness of assessment outcomes.

References

- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481. <https://doi.org/10.2307/1170281>
- Earl, L. M. (2013). Assessment for learning; assessment as learning: Changing practices means changing beliefs. *Assessment and Learning*, 2, 1–5.
- Madaus, G. F., & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P. W. Jackson (Ed.), *Handbook of research on curriculum: A project of the American Educational Research Association* (pp. 119–

154). Macmillan.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*.

Schafer, W. D. (2023). Clarifying Inputs and Outputs of Cognitive Assessments. *Journal of Applied Measurement*, 24(1/2), 1–8.

Schafer, W. D., & Moody, M. (2004). Designing accountability assessments for teaching. *Practical Assessment, Research, and Evaluation*, 9, Article 14. <https://scholarworks.umass.edu/pare/vol9/iss1/14/>

Schafer, W. D., & Moody, M. (2019). Designing accountability assessments for teaching. *Practical Assessment, Research, and Evaluation*, 9(1). <https://doi.org/10.7275/ppdm-ng82>