The Standards for Educational and Psychological Testing and Deaf Individuals

NDC
National Deaf Center
on Postsecondary Outcomes

FAIRNESS IN TESTING

The 2014 *Standards for Educational and Psychological Testing* significantly elevated the importance of fairness in testing.[1] The standards emphasize that fairness is an essential part of the validity of score interpretations. Thus, alongside validity and reliability, fairness gained status as one of three primary tenets of best practices in test design, development, implementation, and score interpretation. Furthermore, the standards emphasize the importance of considering the characteristics of all possible test takers throughout the entire testing process. Fairness can no longer be an afterthought.

The standards' emphasis on fairness as an essential component of validity is especially important for deaf[2] individuals. Deaf individuals are diverse with varying needs for access to assessments that allow them to demonstrate their knowledge and skills. High-stakes assessment is infused throughout education and employment, particularly during transition periods.

The focus of this document is on testing fairness considerations for young deaf adults transitioning from secondary education systems to higher education or to postsecondary jobs and training.

Deaf individuals often have a different linguistic experience from their hearing peers.

- Deaf children enter school with widely varying language skills, which include phonological awareness, vocabulary, and English syntax. Most deaf individuals (more than 90%) are born to hearing parents, a majority of whom do not use sign language. Lack of a shared language between deaf children and their family can lead to delays in critical steps of language acquisition.

- For a deaf student in a classroom with an interpreter, quality of instruction is dependent not only on the skill and approach of the teacher, but also on the degree to which instruction allows for full use of the interpreter.

- Deaf students who use an interpreter or assistive listening devices[3] do not have access to language in the classroom equal to that of hearing students. Interpreters and assistive listening devices typically cover direct instruction by teachers or primary educators but rarely capture all the dialogue that occurs inside and outside a classroom setting.

For more information, see our accompanying brief, *Why Deaf Individuals May Need Accommodations When Taking Tests*.

The standards document identifies 20 fairness standards within four clusters.[1] The major concepts encompassed in the four clusters support test developers and administrators as they work to create a fair testing process for deaf individuals. The four clusters focus on the following areas:

- Designing, developing, and administering tests
- Ensuring test validity and reliability
- Providing needed accommodations
- Safeguarding against inappropriate score interpretations

In the remainder of this brief, we identify important fairness concepts, based on the four clusters in the standards, for test developers and professionals to consider as they work to support an equitable testing environment for deaf test takers.

Design, Develop, and Administer Tests Recognizing the Needs of Deaf Individuals

Five primary concepts are included in the standards related to test design, development, and administration. Three of those concepts are presented below.

1. **Include individuals knowledgeable of the deaf population during the test design and development process.** It is important to consider the access needs of deaf test takers throughout the assessment design and development process. This work typically involves explicitly defining (a) the construct being measured and (b) the access skills needed to demonstrate knowledge on that construct. Some test items require access skills that may create a barrier to accurately measuring the intended construct for deaf test takers. Access skills may include, for example, being able to hear directions or being able to read or write English at an advanced level, when those skills are not being assessed. Experts in the field can help to identify and rectify issues related to access skills that are not relevant to the target construct.

2. **Ensure that test developers know the characteristics of deaf individuals when developing item templates.** The characteristics of item templates to be included in an assessment should be informed by the characteristics of deaf individuals. For example, the level of language knowledge required should not be more complex than required by what is being measured in the test. Also, if sophisticated video or audio components are included, they need to be captioned, at a minimum. Accessibility considerations also need to include how possible accommodations will interact with item templates. Documentation of these considerations and decisions is important for future reference.

3. **Identify ways to collect information on items and test forms during pilot and field testing.** During the test development process, collect appropriate data to document considerations for how deaf individuals access test content. Specifically, the data should illustrate item and test form performance for deaf individuals. When sample sizes are too small for group-level data analyses, test developers can use a single-case approach, or conduct cognitive labs, focus groups, or interviews to obtain evidence of item and test form performance by deaf individuals.

Cognitive labs (also known as "think-alouds") are a way to study cognitive processes during a task.[4, 5] Currently, researchers[6, 7, 8] and test developers use cognitive labs. This method typically involves having subjects "think aloud" as they complete a task, such as a test item, and then following up with questions for the subjects to answer after the task is completed.

Document Procedures to Ensure Valid and Reliable Test Scores for Deaf Individuals

Three primary concepts in the standards relate to score validity and reliability. Each focuses on documenting fairness in testing. In addition to documentation on overall test validity, the documentation process should specify actions specifically for deaf individuals.

1. **Explain how deaf individuals were considered in each step of test development to ensure the validity and reliability of test scores.** Deaf individuals and professionals knowledgeable of the characteristics and needs of deaf individuals should be involved in the design of a test, including determining how deaf individuals will be involved in cognitive labs, pilot tests, and field testing.

   It is important to consider all types of tasks. Constructed response items may unintentionally assess advanced language skills, for example, when they are not the objective of a task. Cognitive labs and other procedures such as focus groups or interviews help to identify when problems occur, but it is better to identify these issues during test design and item development.

   The principles of Universal Design and evidence-centered design should be adhered to during test design. These principles can take many forms, but in essence, tests should be designed to be accessible by the widest range of individuals. Several resources are available to support Universal Design and evidence-centered design approaches inclusive of deaf individuals.

   > According to the standards, Universal Design refers to an approach to assessment development that maximizes the accessibility of a test for all intended test takers.
   >
   > > When using Universal Design, test developers differentiate what is both relevant and irrelevant to the intended construct and test purpose. Test items and tasks can then be purposively designed and developed to address the construct(s) to be measured and to minimize construct-irrelevant features that might otherwise impede the performance of intended test taker groups—for example, individuals with disabilities, those from diverse linguistic or cultural groups, older adults, or young children.[1]
   >
   > For more information, see the Resources section of this brief.

2. **Explain how deaf individuals were included in each step of test development and implementation.** Test development and implementation include steps to support valid interpretations and uses of test scores, including assessment design, pilot testing, field testing, analysis of results, and operational testing. Because the number of deaf individuals included in the pilot and field testing processes is often too small for typical analytical procedures, it is important to consider targeted procedures such as cognitive labs or focus groups during each step of the development process to ensure that the needs of deaf individuals are met.

3. **Document unsuccessful efforts to include deaf individuals in the test development and design process, including possible validity and reliability implications.** Given the importance of including deaf individuals in every step of design and development, it is important to document circumstances when desired levels of participation did not occur. This documentation should also include perceived implications for score validity and reliability due to the lack of desired participation. Attention should be given to any indications of differential validity for deaf individuals compared to hearing peers.

Provide Accommodations That Increase Access for Deaf Individuals

The standards include two general concepts on the topic of accommodations. In addition to the points below, we have developed an accompanying brief for the general public on accommodations for deaf individuals during high-stakes testing titled *Why Deaf Individuals May Need Accommodations When Taking Tests*.

1. **Consider accommodations used by deaf individuals to access test content and document why those accommodations may or may not be appropriate for specific test constructs.** Test developers who know the intended constructs to be assessed should identify the appropriateness of the accommodations that deaf individuals commonly use (see Common Accommodations for Deaf Individuals sidebar below). Accommodation decisions may need to include justification for different accommodations for different sections of an assessment. Developing construct-based rationales for these decisions can inform those who make decisions about specific accommodations for each deaf individual. Ideally, such rationales for accommodations should be accompanied by evidence of the effectiveness of accommodations for deaf individuals with similar characteristics (e.g., language modality and background, prior testing experience). The evidence should also provide an indication that the accommodations do not modify the construct being assessed. Because the number of deaf individuals included often is too small to use traditional data analytical procedures (i.e., differential item functioning), other types of evidence may need to be included. Cognitive labs with deaf individuals and expert opinions are respected ways to provide this evidence.

Common Assessment Accommodations for Deaf Individuals

- Assistive listening device
- Captioned media
- Extended time
- Glossary or dictionary
- Individual administration
- Interpreted test materials in sign language (American Sign Language is most common)
- Scribe

2. Gather and document evidence that supports the use of accommodations by a deaf individual during the assessment. Those who make accommodations decisions should consider the relationship between assessment accommodations and accommodations used during nontesting contexts (e.g., during instruction), striving for consistency unless the construct measured is compromised. Consideration should also be given to the implementation of standardized procedures for administering accommodations, particularly for choosing the interpreter and defining the interpreter's role. For example, guidelines will be needed as to whether interpreters answer questions about test item directions, translate specific words in an item, or translate entire item components. For tests developed with prerecorded translations of test items into sign language, careful consideration should be given to how item translations are implemented and match the range of test taker language use (see Common Accommodations for Deaf Individuals sidebar above).

NOTES AND REFERENCES

1   American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

2   NDC uses the term *deaf* in an all-inclusive manner to include individuals who identify as Deaf, deaf, deafblind, deafdisabled, hard of hearing, late-deafened, and hearing impaired.

3   Assistive listening technology (either preinstalled or portable) allows a listener to tune in directly to a speaker's voice while reducing extraneous sound barriers. Many technologies can also work with a listener's personal hearing aid or cochlear implant device.

4   Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

5   Leighton, J. P. (2017). *Using think-aloud interviews and cognitive labs in educational research: Understanding qualitative research*. New York, NY: Oxford University Press.

6   Johnstone, C. J., Bottsford-Miller, N. A., & Thompson, S. J. (2006). *Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and English language learners* (Technical Report 44). Minneapolis, MN: National Center on Educational Outcomes.

7   King, T. C., & Laitusis, C. C. (2008). *Sample cognitive interview protocol*. Princeton, NY: Educational Testing Service.

8   Lazarus, S. S., Thurlow, M. L., Rieke, R., Halpin, D., & Dillon, T. (2012). *Using cognitive labs to evaluate student experiences with the read aloud accommodation in math* (Technical Report 67). Minneapolis, MN: National Center on Educational Outcomes.

RESOURCES

Websites

- Lists of sign language tests: www.signlang-assessment.info/index.php/home-en.html
- Americans with Disabilities testing accommodations: www.ada.gov/regs2014/testing_accommodations.html

General Resources

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Dorans, N. J., & Cook, L. L. (Eds.). (2016). *Fairness in educational assessment and measurement*. New York, NY: Routledge.
- Hambleton, R. K., Bartram, D., & Oakland, T. (2011). Technical advances and guidelines for improving testing practices. In P. Martin, F. Cheung, M. Knowles, M. Kyrios, L. Littlefield, & J. Overmier (Eds.), *IAAP handbook of applied psychology* (pp. 338–361). Malden, MA: John Wiley & Sons.
- Lane, S., Raymond, M. R., & Haladyna, T. M. (Eds.). (2016). *Handbook of test development*. New York, NY: Routledge.

Design Thinking

- Benson, J., & Dresdow, S. (2014). Design thinking: A fresh approach for transformative assessment practice. *Journal of Management Education, 38*(3), 436–461.
- Tran, R., & Gopalakrishnan, S. (2013). *Design thinking for evaluation and learning*. Retrieved from https://ssir.org/articles/entry/design_thinking_for_evaluation_and_learning

Evidence-Centered Design

- Haertel, G., Haydel DeBarger, A., Cheng, B., Blackorby, J., Javitz, H., Ructtinger, L., & Hansen, E. G. (2010). *Using evidence-centered design and Universal Design for learning to design science assessment tasks for students with disabilities* (Assessment for Students With Disabilities Technical Report 1). Menlo Park, CA: SRI International.
- Mislevy, R. J. (2011). *Evidence centered design for simulation-based assessment* (CRESST Report 800). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Universal Design

- National Center for Universal Design for Learning: http://udlguidelines.cast.org
- National Center on Educational Outcomes: www.nceo.info/Assessments/universal_design/overview