Diversity and Inclusion in Assessment

Alice Keane CPsychol

Senior Talent Assessment Consultant

In any assessment process, it's important to consider issues of fairness and equality of opportunity to give every candidate an equal chance to do their best, and for legal and ethical reasons. Including psychometric tests in your process is a great way to do this, as they help to reduce bias because they are standardised and objective unlike other methods such as interviews. In order to maximise their effectiveness, there are several best practice steps that organisations should incorporate over time.

1. At least once every 5 years, conduct a **job analysis** of the position for which you are administering the test. A job analysis will help you determine if the job has changed in any way that requires adjustments to your selection process.

2. Undertake, and periodically reassess, a local validation study to establish the criterion-related **validity** of the assessment. This will demonstrate the strength of the relationships between test scores and job performance in your candidates.

3. **Evaluate adverse impact** by comparing the selection rates for individuals from protected subgroups (e.g., gender or ethnicity) with selection rates of historically advantaged groups. Information needed to facilitate these analyses includes applicant demographics (e.g., voluntary information on gender, race/ethnicity, and age), assessment scores, and employment status (e.g., hired/not hired). Adverse impact should be monitored at each step of the selection process, and not just in relation to psychometric tests.

4. **Combine assessments** in order to minimise any potential bias. For example, using both verbal and numerical reasoning tests will balance out the gender biases inherent in these tests (men tend to perform better on numerical reasoning and women tend to perform better on verbal reasoning). Alternatively, using a personality or values assessment alongside your cognitive ability test in your decision making can help to reduce any potential bias.

5. Periodically re-examine **cut scores** in light of recent validity results, adverse impact, market data, and other factors (e.g., projected workload), and make adjustments as necessary.

6. When sufficient samples of employees and candidates have been obtained (e.g., > 25 per demographic group), analyse the information collected to see if the selection procedure predicts equally for the majority group and protected groups. The Cleary model is a commonly-used approach to evaluate the fairness of selection tools (Guion, 1998). This model utilises regression analysis to determine if a test demonstrates **differential validity** or prediction among subgroups of applicants.

7. No test taker should be disadvantaged in their access to your assessment, so **reasonable adjustments** should be offered to all candidates and provided to those who require them in order to level the playing field. These include but aren't limited to: provision of extra time to complete a test, rest breaks, large print or Braille versions of tests, screen readers and voice recognition software.

8. Provide a **practice test** and encourage your candidates to complete it. Practice tests allow candidates to feel more at ease when completing the test. Research has shown that score gains between practice tests and actual tests were greater for Black and Hispanic candidates, which means this can help to minimise adverse impact in a selection process.

Adverse impact, or a disproportionately negative effect on potential candidates because of their gender, ethnicity, religion, disability, age or sexual orientation, is to be expected in any selection process, to some degree. The important thing is to monitor it, to minimise it as much as possible and to ensure that the assessments you use are valid and justifiable in terms of the job requirements so that your process is defendable.

According to the Uniform Guidelines on Employee Selection Procedures (Equal Employment Opportunity Commission, 1978), adverse impact is indicated when the selection rate for one group is less than 80% (or 4 out of 5) of another. The degree of adverse impact seen in the use of any assessment, will vary between organisations and selection processes as it is specific to the sample of candidates rather than to the assessment itself.

Adverse impact is not a legal term that implies guilt or a psychometric term that implies unfairness or test bias; rather, it simply describes differences between groups on a testing process. Many employment tests, including cognitive ability tests, result in adverse impact. Adverse impact is not normally due to forms of bias inherent to the test (Sackett, et al., 2001).

An assessment with adverse impact can still be used for selection, but the organisation must demonstrate that the test is job-related, predicts performance, and is consistent with business necessity. A local validation study, in which test scores are correlated with indicators of on-the-job performance, provides evidence to support the use of the test in a particular job context. A local study that demonstrates that the test is equally predictive for protected subgroups, helps to further establish test fairness.

Does adverse impact mean the test is not appropriately predicting performance for all groups?

No. Racial group differences are commonly observed in scores on standardised knowledge, skill, ability, and achievement tests where African Americans tend to repeatedly score approximately one standard deviation lower than Whites while Hispanics score approximately two thirds of a standard deviation lower than Whites (Sackett, et al., 2001). However, an extensive body of literature has shown that these tests do not under-predict minority performance (e.g., American Educational Research Association, American Psychological Association, & National Council of Measurement in Education. 1999; Sackett & Wilk. 1994).

References

Campion, M. C., Campion, E. D., & Campion, M. A. (2019). Using practice employment tests to improve recruitment and personnel selection outcomes for organizations and job seekers. *Journal of Applied Psychology, 104* (9), 1089-1102.

Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43* (166), 38295–38309.

Guion, R. M. (1998). Assessment, measurement, and prediction for personnel decisions. Mahwah, NJ: Lawrence Erlbaum and Associates.

Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. American Psychologist, 49, 929-954. doi: 10.1037/0003-066X.49.11.929

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education. American Psychologist, 56, 302-318. doi: I0.IO37/AJO03-O66X.56.4.302