



Diversiteit en inclusie in beoordelingen

Alice Keane, erkend psychologe
Senior Talentbeoordelingsconsultant

A decorative graphic consisting of numerous small pink dots arranged in a circular pattern, with the dots becoming larger towards the center.

**MORE INSIGHT
MORE IMPACT™**

In elk beoordelingsproces is het van belang dat je rekening houdt met kwesties van eerlijkheid en gelijkheid van kansen om elke kandidaat een gelijke kans te geven om zijn/haar best te doen, en omwille van wettelijke en ethische redenen. Het opnemen van psychometrische tests in je proces is een fantastische manier om dit te doen. Ze helpen vooroordelen verminderen omdat ze gestandaardiseerd en objectief zijn in tegenstelling tot andere methodes zoals sollicitatiegesprekken. Om de effectiviteit ervan te maximaliseren zijn er verschillende stappen voor best practices die organisaties mettertijd zouden moeten overwegen.

1. Ten minste om de 5 jaar een **functieanalyse** uitvoeren van de functie waarvoor je de test afneemt. Een functieanalyse helpt je te bepalen of de functie zodanig veranderd is dat er aanpassingen nodig zijn aan je selectieproces.
2. Een lokaal valideringsonderzoek uitvoeren en periodiek herbeoordelen, om de **validiteit** van de beoordeling na te gaan in functie van de criteria. Dit zal aantonen hoe sterk het verband is tussen de testcores en de werkprestaties bij je kandidaten.
3. **Negatieve impact evalueren** door selectiescores van mensen uit beschermde subgroepen (bijvoorbeeld geslacht of etniciteit) te vergelijken met selectiescores van historisch bevoordeelde groepen. De informatie die nodig is voor deze analyses bestaat uit demografische gegevens van de kandidaat (bijvoorbeeld vrijwillige informatie over het geslacht, ras/eticiteit en leeftijd), beoordelingscores en werksituatie (in dienst/niet in dienst). Je moet negatieve impact in elke stap van het selectieproces in de gaten houden en niet alleen in verband met psychometrische tests.
4. **Beoordelingen combineren** om alle potentiële vooroordelen te minimaliseren. Bijvoorbeeld, door zowel verbale als numerieke redeneringstests te gebruiken, breng je vooroordelen met betrekking tot geslacht in deze tests in evenwicht (mannen presteren doorgaans beter op numeriek redeneren en vrouwen presteren meestal beter op verbaal redeneren). Door bij je besluitvorming een persoonlijkheids- of waardebeoordeling te gebruiken naast je cognitieve vaardigheidentest, kan je potentiële vooroordelen verminderen.
5. Bekijk periodiek de **de minimumscores om te slagen** opnieuw in het kader van recente validiteitsresultaten, negatieve impact, marktgegevens en andere factoren (bijvoorbeeld verwachte werklast) en breng zo nodig aanpassingen aan.
6. Wanneer je voldoende referentiemateriaal van medewerkers en kandidaten hebt verkregen (bijvoorbeeld > 25 per demografische groep), analyseer je de verzamelde informatie om te zien of de selectieprocedure in gelijke mate predictief is voor de meerderheidsgroep en de beschermde groepen. Het Cleary-model is een vaak gebruikte benadering om de eerlijkheid van selectietools te evalueren (Guion, 1998). Dit model gebruikt regressieanalyse om te bepalen of een test wijst op **differentiële validiteit** of voorspelling tussen kandidaatsubgroepen.
7. Geen enkele deelnemer mag worden benadeeld bij deelname aan jouw beoordeling. Daarom moeten er **redelijke aanpassingen** worden aangeboden aan alle kandidaten en voorzien voor degenen die deze nodig hebben om een gelijke beoordeling te kunnen bieden. Deze omvatten maar zijn niet beperkt tot: het voorzien van extra tijd voor een test, rustpauzes, grote print of Braille-versies van tests, schermlezers of software voor stemherkenning.
8. Deel een **oefentest** en moedig je kandidaten aan om deze test te maken. Oefentests zorgen ervoor dat kandidaten zich meer op hun gemak voelen tijdens het afleggen van de test. Onderzoek heeft aangetoond dat de scoretoename tussen oefentests en de echte tests groter was voor donkerkleurige kandidaten en hispanics wat betekent dat dit de negatieve impact in een selectieproces kan minimaliseren.

Negatieve impact, of een disproportioneel negatief effect op potentiële kandidaten omwille van hun geslacht, etniciteit, religie, handicap, leeftijd of seksuele geaardheid is, tot op zekere hoogte, te verwachten in elk selectieproces. Het belangrijkste is dat je het selectieproces overziet, de verschillen zoveel mogelijk beperkt en ervoor zorgt dat de beoordeling die je gebruikt valide en gerechtvaardigd is in het oogpunt van de functievereisten zodat je proces verdedigbaar is.

Volgens de uniforme richtlijnen voor selectieprocedures van werknemers (Commissie gelijke tewerkstellingskansen, 1978), is er een aanwijzing voor negatieve ene groep minder is dan 80% (of 4 van de 5) ten opzichte impact wanneer het selectiepercentage van de van de andere groep. De graad van negatieve impact die te zien is in het gebruik van een beoordeling varieert per organisatie en selectieproces omdat het specifiek is voor de groep kandidaten en niet voor de beoordeling zelf.

Negatieve impact is geen wettelijke term die schuld impliceert, of een psychometrische term die wijst op oneerlijkheid of testvooroordelen. Het is eerder een eenvoudige beschrijving van verschillen tussen groepen in een testproces. Veel wervingstests, met inbegrip van cognitieve vaardigheidstests, zorgen voor negatieve impact. Negatieve impact is normaal gesproken niet te wijten aan vormen van vooroordelen die inherent aan de test verbonden zijn (Sackett, et al., 2001).

Een beoordeling met negatieve impact kan nog steeds worden gebruikt voor selectie, maar de organisatie moet aantonen dat de test functiegerelateerd is, prestatievoorspellend en een zakelijke noodzaak is. Een lokale validatiestudie waarin testcores worden gecorreleerd met indicatoren van prestaties in de praktijk, levert bewijs ter ondersteuning van het gebruik van de test in een bepaalde werkcontext. Een lokale studie die aantoont dat de test in gelijk mate een voorspelling geeft voor beschermde subgroepen, zal de eerlijkheid van de test verder bevestigen.

Betekent negatieve impact dat de test de prestaties voor alle groepen niet op de juiste manier voorspelt?

Nee. Verschillen tussen raciale groepen worden vaak waargenomen in scores op gestandaardiseerde kennis, vaardigheden, capaciteiten, en prestatietests waar Afrikaanse Amerikanen meestal ongeveer één standaardafwijking lager scoren dan blanke mensen terwijl hispanics ongeveer twee derde standaardafwijking lager scoren dan blanke mensen (Sackett, et al., 2001). Een uitgebreide literatuurreeks heeft echter aangetoond dat deze tests de prestaties van minderheden niet te laag voorspellen (bijvoorbeeld de American Educational Research Association, American Psychological Association, & National Council of Measurement in Education. 1999; Sackett & Wilk. 1994).

Referenties

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: 10.1037/AJ003-066X.56.4.302