

Aanvulling op COTAN Beoordelingssysteem m.b.t. unproctored gegevensverzameling

Het onderwerp ‘unproctored gegevensverzameling’ valt onder basisvraag 4.3 ‘Wat is de kwaliteit van de verstrekte normgroepen?’, specifiek onder basisvraag 4.3.b ‘Zijn de normgroepen representatief?’, aangezien deze basisvraag tevens gaat over de manier van gegevensverzameling. De volgende passage wordt toegevoegd aan het COTAN Beoordelingssysteem voor de Kwaliteit van tests (2010) op pagina 25 aan het eind van de tekst bij ‘Aanwijzingen bij vraag 4.3.b’ en vóór ‘Aanwijzingen bij vraag 4.4’.

Unproctored gegevensverzameling

Bij voorgaande is steeds uitgegaan van zogeheten ‘proctored’ gegevensverzameling (d.w.z. onder toezicht van de testleider). Met de komst van computers en internet en daarmee (meer) computergebaseerde tests is het aantrekkelijk geworden om ook ‘unproctored’ gegevens te verzamelen (d.w.z. zonder toezicht van de testleider). Unproctored gegevensverzameling heeft als grote voordeel dat meer mensen op verschillende locaties, zelfs uit verschillende landen, de test kunnen invullen. Een groot nadeel is dat de identiteit van de kandidaat nooit met zekerheid kan worden vastgesteld en dat daarmee de deur naar ‘sjoemelen’ open staat. Overigens lijkt de invloed van de nadelen per type test te verschillen. Onderzoek laat zien dat de invloed ervan op onder andere de factorstructuur, de ‘criterion-related’ validiteit en verschillen in gemiddelden verwaarloosbaar is voor persoonlijkheidsvragenlijsten, niet-cognitieve vragenlijsten en/of niet-vaardigheidstests (zie bijvoorbeeld Chuah, Drasgow, & Roberts, 2006). Bij capaciteiten- en vaardigheidstests is meer voorzichtigheid geboden, onder andere omdat daar niet te controleren is of er bepaalde hulpmiddelen zijn gebruikt tijdens de afname.

De gegevens die unproctored zijn verzameld voor persoonlijkheidsvragenlijsten, niet-cognitieve vragenlijsten en/of niet-vaardigheidstests ten behoeve van de normering, betrouwbaarheids- en validiteitsonderzoek worden op dezelfde wijze beoordeeld als gegevens die proctored zijn verzameld. Hierbij gelden vergelijkbare voorwaarden:

- de representativiteit van de normgroepen dient op dezelfde wijze te worden aangetoond.
- de situatie waarin de gegevens verzameld worden moet overeenkomen met het doel waarvoor de test later gebruikt gaat worden (bijvoorbeeld voor selectie of loopbaanbegeleiding). Als dit niet het geval is, dan wordt de beoordeling ‘onvoldoende’.
- de gegevens moeten verzameld zijn onder dezelfde afnamecondities (proctored of unproctored) als waarin de test daadwerkelijk wordt afgenomen. Indien de gegevens unproctored zijn verzameld en de test wordt proctored afgenomen dient aannemelijk te worden gemaakt dat het niet uitmaakt of de gegevens proctored of unproctored zijn verzameld of waarom unproctored de aangewezen manier is om dit te doen en/of eventueel te prefereren is boven proctored gegevensverzameling. Dit kan bijvoorbeeld door op basis van empirische gegevens de equivalentie van beide afnamecondities aan te tonen of middels een onderbouwing op basis van wetenschappelijke literatuur. Als niet aannemelijk wordt gemaakt dat het niet uitmaakt of de gegevens proctored of unproctored zijn verzameld, dient het oordeel voor deze beoordelingsvraag naar beneden te worden bijgesteld (d.w.z. ‘goed’ wordt ‘voldoende’ en ‘voldoende’ wordt ‘onvoldoende’).

Voor de beoordeling van capaciteiten- en vaardigheidstests geldt het volgende:

- indien de gegevens ten behoeve van de normering, betrouwbaarheids- en validiteitsonderzoek unproctored zijn verzameld en de test dient in de praktijk proctored te worden afgenomen, is het oordeel 'onvoldoende'.
- indien de gegevens ten behoeve van de normering, betrouwbaarheids- en validiteitsonderzoek unproctored zijn verzameld en de test dient volgens de auteur(s)/testuitgever in de praktijk ook unproctored te worden afgenomen, kan het oordeel maximaal 'voldoende' zijn, onder de voorwaarde dat tevens de equivalentie van beide afnamecondities op basis van empirische gegevens is aangetoond. Indien de equivalentie niet (voldoende) wordt aangetoond, geldt de beoordeling 'onvoldoende'.

Hierbij dient benadrukt te worden dat de COTAN adviseert om zowel de afname van als de verzameling van de normgegevens bij capaciteiten- en vaardigheidstests (zowel in *low stakes* als in *high stakes* situaties) proctored te doen.

- indien de gegevens ten behoeve van de normering, betrouwbaarheids- en validiteitsonderzoek proctored zijn verzameld en de test dient in de praktijk unproctored te worden afgenomen, kan het oordeel 'goed' zijn, onder de voorwaarde dat tevens een uitgebreide en duidelijke passage in de handleiding is opgenomen met betrekking tot goed en verantwoord testgebruik bij unproctored testafnames. Indien een dergelijke passage ontbreekt en/of niet (voldoende) duidelijk is, geldt het oordeel 'voldoende'.

Hierbij geldt aanvullend, net als bij persoonlijkheidsvragenlijsten, niet-cognitieve vragenlijsten en/of niet-vaardigheidstests, dat de representativiteit van de normgroepen op dezelfde wijze als bij proctored gegevensverzameling dient te worden aangetoond en dat de situatie waarin de gegevens verzameld worden overeen moet komen met het doel waarvoor de test later gebruikt gaat worden (bijvoorbeeld voor selectie of loopbaanbegeleiding). Als dit niet het geval is, dan wordt de beoordeling 'onvoldoende'.

Daarnaast geldt voor alle soorten tests dat een volledige beschrijving dient te zijn opgenomen van hoe fraude bij de testafname kan worden voorkomen en over hoe de test beveiligd is tijdens de gegevensverzameling om er zo zeker mogelijk van te zijn dat ook daadwerkelijk de deelnemer achter het scherm zit. Hierbij kan gedacht worden aan de 'controlled mode', zoals geformuleerd door de International Test Commission (2005), waarbij de test alleen tijdens een bepaalde periode beschikbaar wordt gemaakt voor de kandidaat zelf door middel van een gebruikersnaam en wachtwoord. Verder is het van belang dat beschreven wordt hoe gecontroleerd gaat worden op frauduleuze en/of dubieuze afnames en wat de criteria zijn om bepaalde gegevens niet op te nemen in de normgegevens. Opmerkingen over hoe dit geregeld is, zullen worden opgenomen in de Toelichting bij de beoordeling.

Chuah, S. C., Drasgow, F., Roberts, B. W. (2006). Personality assessment: Does the medium matter? No. *Journal of Research in Personality*, 40, 359-376.

International Test Commission (ITC) (2005). *ITC Guidelines on Computer-Based and Internet Delivered Testing*. Downloaded electronically on June 19th 2014 from https://www.intestcom.org/files/guideline_computer_based_testing.pdf