College voor Toetsen en Examens

CHECKLIST VOOR HET BEOORDELEN VAN DE KWALITEIT VAN OBSERVATIECATEGORIEËN EN TOETSOPGAVEN

1 INLEIDING

Deze checklist fungeert als hulpmiddel voor het beoordelen van de onderwijskundige kwaliteit van observatiecategorieën en toetsopgaven (c.q. items) die onderdeel worden van een tot het stelsel toegelaten doorstroomtoets of erkend LVS-instrument. Meer specifiek kan met deze checklist worden getoetst of:

De observatiecategorieën van een observatie- en registratie-instrument – type (I) en (II) voldoen aan:

- de kwaliteitscriteria: relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit, zoals opgenomen in kwaliteitscriterium O.I.4 uit paragraaf 2.2.1 uit het Beoordelingskader voor observatie- en registratie-instrumenten groep 1 – 8 primair onderwijs en voor instrumenten binnen leerlingvolgsystemen groep 3 – 8 primair onderwijs;
- de constructievoorschriften uit deze onderhavige checklist.

De toetsopgaven van een LVS-instrument – type (III) en (IV) voldoen aan:

- de kwaliteitscriteria: relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit, zoals opgenomen in de kwaliteitscriteria O.IV.2 en O.IV.3 uit paragraaf 2.3.2 uit het Beoordelingskader voor observatie- en registratie-instrumenten groep 1 – 8 primair onderwijs en voor instrumenten binnen leerlingvolgsystemen groep 3 – 8 primair onderwijs;
- de constructievoorschriften uit deze onderhavige checklist.

De toetsopgaven van de wettelijk verplichte onderdelen en de niet-wettelijk verplichte onderdelen Nederlandse taal en rekenen van een doorstroomtoets voldoen aan:

- de kwaliteitscriteria V2 en V3 uit paragraaf 3.2 uit het Beoordelingskader Doorstroomtoetsen;
- de kwaliteitscriteria: relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit;
- de constructievoorschriften uit deze onderhavige checklist.

2 KWALITEITSCRITERIA

Deze checklist onderscheidt vijf kwaliteitscriteria.

Relevantie (R): de observatiecategorie of de toetsopgave meet wat deze beoogt te meten; de observatiecategorie of de toetsopgave heeft een directe relatie met het gedragskenmerk of het leerdoel en het te verwachten niveau.

Objectiviteit (O): inhoudsdeskundigen (vakdeskundigen) zijn het onder alle omstandigheden eens over een eenduidige interpretatie van de observatiecategorie of het juiste antwoord van de toetsopgave. Ofwel de observatiecategorie of de toetsopgave is zodanig geformuleerd dat alle inhoudsdeskundigen tot dezelfde interpretatie of hetzelfde modelantwoord komen.

Efficiëntie (E): de observatiecategorie of de toetsopgave is niet moeilijker geformuleerd dan nodig is en bevat alleen informatie die nodig is om de categorie te observeren of om de opgave te beantwoorden. Dit criterium geldt niet voor redundante informatie die nodig is om een vaardigheid te toetsen of bijvoorbeeld afbeeldingen die zijn toegevoegd om tekst te ondersteunen of te onderbreken.

Specificiteit (S): de observatiecategorie of de toetsopgave bevat geen aanwijzingen richting een gewenst ontwikkelingsniveau of richting het juiste antwoord. Alleen leerlingen die het gedragskenmerk laten zien of die het toetsonderdeel beheersen worden op het bijpassende ontwikkelingsniveau ingedeeld of kunnen de opgave juist beantwoorden.

Neutraliteit (N): de inhoud en de presentatie van de observatiecategorie of de toetsopgave is niet aanstootgevend voor de leerling.

Voor de observatiecategorieën of de toetsopgaven van de niet (wettelijk) verplichte onderdelen, zoals genoemd in artikel 8, tweede lid van de Wet op het Primair Onderwijs (WPO) en artikel 11, derde lid WEC, zijn sommige kwaliteitscriteria mogelijk niet van toepassing.

In deze checklist worden de vijf kwaliteitscriteria geoperationaliseerd in constructievoorschriften voor gesloten vragen en voor open vragen. Op deze manier kunnen de in de inleiding genoemde kwaliteitscriteria per type instrument objectief worden geëvalueerd. De constructievoorschriften zijn van toepassing op zowel papieren als digitale toetsen en op zowel klassieke als meer innovatieve itemtypes. Voor observatiecategorieën wordt per instrument en vorm gekeken welke criteria van toepassing zijn. De verschijningsvormen van instrumenten en itemtypes zijn meer divers dan hier wordt beschreven. Per type instrument of itemtype wordt nagegaan welke criteria wel of niet van toepassing zijn.

3 CONSTRUCTIEVOORSCHRIFTEN GESLOTEN VRAAG

Een gesloten vraag is een vraagtype waarbij de leerling moet kiezen uit een beperkt aantal antwoordmogelijkheden die vooraf gegeven zijn. Hierbij zijn verschillende innovatieve vraagvormen mogelijk. Sireci en Zenisky (2006) beschrijven veertien verschillende vraagvormen, die gebruikt worden in digitale toetsen. Voorbeelden zijn de één-uit-meervraag, meer-uit-meervraag, matchvraag, volgorde- of rangschikvraag, hotspotvraag, sleepvragen en invul- en aanvulvraag.

Een gesloten vraag kan bestaan uit drie onderdelen:

1. De situatiebeschrijving, zijnde de achtergrondinformatie die nodig is om de gesloten vraag te kunnen beantwoorden;
2. De stam, zijnde de vraag zelf;
3. De alternatieven, zijnde de sleutel (c.q. het juiste antwoord) en de afleider(s) (c.q. de overige onjuiste antwoordmogelijkheden). Bij een gesloten vraag zonder antwoordalternatieven zijn de criteria ten aanzien van alternatieven niet van toepassing.

3.1 CONSTRUCTIEVOORSCHRIFTEN PER ONDERDEEL VAN DE GESLOTEN VRAAG (INCLUSIEF BIJPASSEND KWALITEITSCRITERIUM)

3.1.1 Situatiebeschrijving

- De situatiebeschrijving bij de toetsopgave is gebaseerd op een realistische en veelvoorkomende praktijksituatie. (R)
- De situatiebeschrijving is nodig voor het kunnen beantwoorden van de opgave. (R/E)
- De situatiebeschrijving is een brede, rijke en eerlijke weergave van de maatschappij. (N)
- Een verwijzing naar context of bron in de situatiebeschrijving is duidelijk aangegeven. (O)
- Een afbeelding, video- of audiofragment in de situatiebeschrijving is van goede kwaliteit. (E)

3.1.2 Stam

- De toetsopgave sluit aan bij het beoogde toetsdoel. (R)
- De toetsopgave sluit aan bij het beheersingsniveau . (R)
- De toetsopgave bevat correct taalgebruik. (E)
- De toetsopgave bevat geen onnodige moeilijke woorden of afkortingen. (E)
- Bij de constructie van de toetsopgave is rekening gehouden met voldoende contrast in kleur en helderheid. (E)
- De toetsopgave bevat geen figuurlijk taalgebruik, tenzij dat het meetdoel is. (E)
- De toetsopgave bevat geen dubbele ontkenningen. (E)
- De toetsopgave bevat waar mogelijk actieve formuleringen. (E)
- De stam bevat zo min mogelijk ik-zinnen, vraagzinnen en gebiedende wijs zinnen. (E)
- Er wordt in de toetsopgave niet gevraagd naar een mening van de leerling, tenzij dat het meetdoel is. (O/N)
- De vraag in de toetsopgave is ondubbelzinnig en kan maar op één manier worden geïnterpreteerd. (O/E/S)

- De toetsopgave heeft een overzichtelijke lay-out en vormgeving. (E)
- De toetsopgave is niet aanstootgevend, vrij van bijvoorbeeld discriminerende en voor bepaalde bevolkingsgroepen kwetsende inhoud. (S/N)
- De toetsopgave levert geen vertekend beeld door leeftijd, sekse, cultuur of levensbeschouwing op. (N)

3.1.3 Alternatieven

- De toetsopgave bevat afleiders die aannemelijk zijn voor leerlingen. (S)
- De toetsopgave bevat antwoordalternatieven die elkaar uitsluiten. (S)
- De toetsopgave bevat antwoordalternatieven die in een logische volgorde zijn gerangschikt. (E/S)
- De toetsopgave bevat antwoordalternatieven die homogeen zijn qua lengte, woordgebruik en formulering. (S)
- De toetsopgave bevat antwoordalternatieven die overzichtelijk zijn weergegeven (qua lay-out en vormgeving). (S)
- De toetsopgave bevat antwoordalternatieven die bij voorkeur positief geformuleerd zijn. (E)
- De toetsopgave bevat antwoordalternatieven die helder en eenduidig zijn geformuleerd. (S)
- De toetsopgave bevat antwoordalternatieven die taalkundig correct geformuleerd zijn. (E)
- De sleutel van de toetsopgave mag niet tot discussie leiden tussen vakinhoudelijke deskundigen. (O)
- De alternatieven zijn niet aanstootgevend, vrij van bijvoorbeeld discriminerende en voor bepaalde bevolkingsgroepen kwetsende inhoud. (S/N)
- De alternatieven leveren geen vertekend beeld door leeftijd, sekse, cultuur of levensbeschouwing op. (N)

4 CONSTRUCTIEVOORSCHRIFTEN OPEN VRAAG

Een open vraag is een vraagtype waarbij een leerling het antwoord zelf moet formuleren. Open vragen kunnen, gekarakteriseerd naar de lengte van het antwoord, onderverdeeld worden in bepaalde vraagvormen, namelijk: kort-antwoordvraag, lang-antwoordvraag, opstelvraag.

Een open vraag kan bestaan uit drie onderdelen:

1. De situatiebeschrijving, zijnde de achtergrondinformatie die nodig is om de open vraag te kunnen beantwoorden;
2. De te beantwoorden vraag;
3. Het antwoordmodel, zijnde een eenduidige en volledige beschrijving van het modelantwoord.

4.1 CONSTRUCTIEVOORSCHRIFTEN PER ONDERDEEL VAN DE OPEN VRAAG (INCLUSIEF BIJPASSEND KWALITEITSCRITERIUM)

4.1.1 Situatiebeschrijving

- De situatiebeschrijving bij de toetsopgave is gebaseerd op een realistische en veelvoorkomende praktijksituatie. (R)
- De situatiebeschrijving is nodig voor het kunnen beantwoorden van de vraag. (R/E)
- De situatiebeschrijving is een brede, rijke en eerlijke weergave van de maatschappij. (N)
- Een verwijzing naar context of bron in de situatiebeschrijving is duidelijk aangegeven. (O)
- Een afbeelding, video- of audiofragment in de situatiebeschrijving is van goede kwaliteit. (E)
- De situatiebeschrijving heeft een overzichtelijke lay-out en vormgeving. (E)
- De situatiebeschrijving is niet aanstootgevend, vrij van bijvoorbeeld discriminerende en voor bepaalde bevolkingsgroepen kwetsende inhoud. (S/N)
- De situatiebeschrijving levert geen vertekend beeld door leeftijd, sekse, cultuur of levensbeschouwing op. (N)

4.1.2 Vraag

- De toetsopgave sluit aan bij de inhoud van het beoogde toetsdoel. (R)
- De toetsopgave sluit aan bij het beheersingsniveau. (R)
- De toetsopgave bevat correct taalgebruik. (E)
- De toetsopgave bevat geen onnodige moeilijke woorden of afkortingen. (E)
- Bij de constructie van de toetsopgave is rekening gehouden met voldoende contrast in kleur en helderheid. (E)
- De toetsopgave bevat geen figuurlijk taalgebruik, tenzij dat het meetdoel is. (E)
- De toetsopgave bevat geen dubbele ontkenningen. (E)
- De toetsopgave bevat waar mogelijk actieve formuleringen. (E)
- De toetsopgave bevat zo min mogelijk ik-zinnen, zo min mogelijk gebiedende wijs zinnen en niet meer dan één vraagzin. (E)
- Er wordt in de toetsopgave niet gevraagd naar een mening van de leerling, tenzij dat het meetdoel is. (O/N)
- De vraag in de toetsopgave is ondubbelzinnig en kan maar op één manier worden geïnterpreteerd. (O/E/S)
- De vraag in de toetsopgave bevat, indien nodig, concrete instructies voor de gewenste lengte en vorm van het antwoord. (O)

- De toetsopgave heeft een overzichtelijke lay-out en vormgeving. (E)
- De toetsopgave is niet aanstootgevend, vrij van bijvoorbeeld discriminerende en voor bepaalde bevolkingsgroepen kwetsende inhoud. (S/N)
- De toetsopgave levert geen vertekend beeld door leeftijd, sekse, cultuur of levensbeschouwing op. (N)

4.1.3 Antwoordmodel

- De toetsopgave bevat een volledig antwoordmodel. (O)
- De toetsopgave bevat een antwoordmodel met een duidelijk scoringsvoorschrift. (O)
- De toetsopgave bevat, indien van toepassing, een duidelijke instructie voor de beoordelaars. (O)

5 REFERENTIES

American Educational Research Association. (2014). *Standards for Educational and Psychological Testing.* Washington, DC: American Educational Research Association.

Cito CETE-CvTE (2015). *Het construeren van examenopgaven*. Cito: Arnhem

Downing, S.M. (2006). Selected- Response Item formats in Test Development. In S.M. Downing & T.M. Haladayna (eds.). *Handbook of test development* (pp.287-302). Lawrence Erlbaum Associates Publishers.

Molkenboer, H., (2018). *180 gevalideerde richtlijnen voor toetsen. Overzicht.* Enschede*:* Bureau voor Toetsen & Beoordelen.

Sanders, P., Brouwer, A., Eggen, T., & Veldkamp, B. (2018). *RCEC beoordelingssysteem voor de kwaliteit van studietoetsen en (praktijk)examens.* Apeldoorn: RCEC.

Sireci, S.G., & Zenisky, A.L. (2006). Innovative item Formats in Computer-Based testing: In Pursuit of Improved Construct Reprsentation. In S.M., Downing & T.M., Haladayna (eds.). *Handbook of test development* (pp.329-347). Lawrence Erlbaum Associates Publishers.

Teelen (2015). *Toetskwaliteit in de praktijk. Hoe maak ik goede toetsen met gesloten en open vragen?* Wilp: Teelen B.V.