

Standaardsetting

Bij een standaardsetting wordt een punt op de vaardigheidsschaal bepaald, zie figuur hieronder. Voor het bepalen van de standaarden zijn meerdere soorten procedures beschikbaar.

Procedures voor standaardsetting

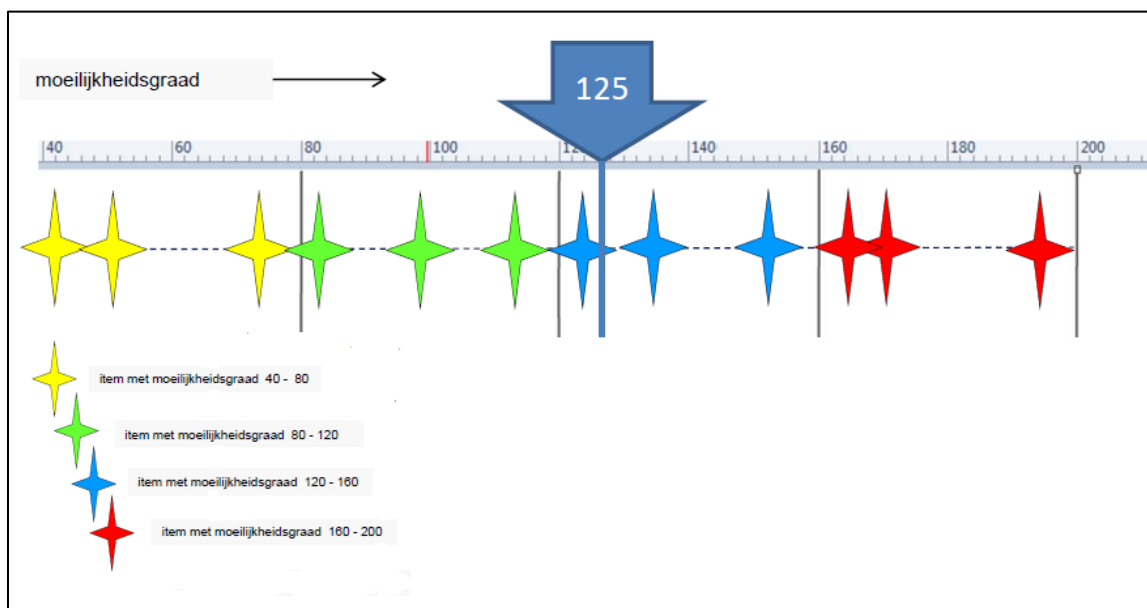
Een eerste onderscheid kan gemaakt worden tussen procedures waarbij:

- beoordelaars of experts gebruikt worden;
- standaarden uit voorgaand onderzoek toegepast worden op nieuwe examen-afnames (zie: **Ankering**).

Bij de standaardsetting-procedures met beoordelaars of experts zijn er twee subtypen te onderscheiden:

- procedures waarbij de experts kandidaten beoordelen;
- procedures waarbij de experts opgaven beoordelen.

Bij de COE's en de centrale rekentoets VO wordt alleen van procedures gebruik gemaakt waarbij opgaven beoordeeld worden. Dit kan weer op meerdere manieren. Hieronder worden twee gebruikte procedures met beoordelaars van opgaven besproken: **Angoff-procedure** en **Bookmark-procedure**. Merk op dat deze methoden niet beperkt hoeven te zijn tot de cesuur voor een voldoende, maar ook toegepast kunnen worden voor het cijfer 5 of 8.



Standaard (125) op een vaardigheidsschaal

Ankering

Een anker is een set van opgaven die uit voorgaand onderzoek stamt. Dit kunnen opgaven uit eenzelfde examen van een eerder tijdstip zijn, of opgaven uit andere examens. Wel is nodig dat de vaardigheid die in het voorgaande onderzoek gemeten werd, dezelfde vaardigheid is die met het

huidige examens gemeten wordt. Dus in een rekentoets VO kunnen ankeropgaven van vorig jaar zitten, of ankeropgaven uit de COE's Rekenen, of uit de [Referentiesets](#) rekenen.

Een voordeel van het opnemen van ankers in een examenperiode is dat standaarden die bij voorgaand onderzoek door experts al zijn vastgesteld, direct gehanteerd kunnen worden bij de nieuwe examens. Dit laatste noemen we standaardsetting via ankering.

Toetsen of examenvarianten die via ankers aan elkaar gekoppeld zijn, kunnen afgebeeld worden op dezelfde schaal. Meestal worden de itemparameters van de ankeropgaven gefixeerd, waarmee de vaardigheidsschaal van het voorgaande onderzoek ook gehanteerd wordt voor de afbeelding van de nieuwe opgaven en vaardigheden van nieuwe kandidaten.

Het kan voorkomen dat de itemparameters van het anker niet passen bij de populatie die deelneemt aan de nieuwe examenperiode. Bijvoorbeeld doordat het een andere populatie betreft (vo versus mbo), of omdat er gerichte training heeft plaatsgevonden tussen de oude en nieuwe periode. In dergelijke gevallen horen bij het anker twee sets parameters: één op de oude vaardigheidsschaal en één op de nieuwe. Als het anker uit minstens 40 opgaven bestaat, is een standaard op de oude schaal toch te vertalen naar een standaard op de nieuwe schaal. De standaard op de oude schaal wordt namelijk eerst vertaald naar een ruwe score op het anker, en deze wordt vervolgens weer vertaald in een positie op de nieuwe schaal.

Als er zowel een standaardsetting via ankering heeft plaatsgevonden, als een standaardsetting met experts, dan kunnen de verkregen cesuren van elkaar afwijken. Bij zo'n afwijking moet eerst gekeken worden of het verschil te wijten valt aan toeval. Indien er structureel afwijkingen geconstateerd worden, moet er wellicht een herziening van de oude standaard plaatsvinden.

Angoff-procedure

Bij een Angoff-procedure moeten de experts een grenskandidaat in gedachten nemen. Een grenskandidaat is een kandidaat die het vereiste niveau net beheerst. De experts geven per item aan of een grenskandidaat deze goed maakt of niet. Bij een gemodificeerde Angoff-procedure wordt per item door de experts de kans ingeschat dat een grenskandidaat de opgave goed maakt. Dit is hetzelfde als inschatten hoeveel van 100 grenskandidaten de opgave goed maken. Optellen van de kansen van alle opgaven van een bepaald examen of toets geeft de grensscore die gehaald moet worden op het betreffende examen of toets om te voldoen aan minimale eisen: de cesuur. Deze cesuur wordt vervolgens afgebeeld op de vaardigheidsschaal, waarmee de standaard bepaald is.

Bookmark-procedure

Als er ook afnamegegevens beschikbaar zijn, kan ook gewerkt worden met een Bookmark-procedure, waarbij gebruik gemaakt wordt van de empirische ordening van de items naar moeilijkheid ([Van der Schoot](#), 2001 en 2008). Deze procedure voor het vaststellen van een standaard voor een bepaald niveau maakt gebruik van een reeks opgaven die naar moeilijkheid zijn gerangschikt. Deskundigen geven aan welke opgave nog wel en welke niet meer beheerst zouden moeten worden door een grenskandidaat. In de reeks opgaven plaatsen zij een 'bookmark'. Experts geven dus aan hoeveel items van een set opgaven beheerst moeten worden op een bepaald niveau. Dit oordeel wordt weer omgezet in een punt op de vaardigheidsschaal.