

## Hoe komt een toets tot stand?

Bij het maken van een toets die betrouwbaar meet wat u wilt weten, komt heel wat kijken. Toetsen maken is een vak apart. Wij, bij Cito, beoefenen dit vak al sinds 1968! Graag laten wij u in deze flyer in vogelvlucht zien hoe een toets tot stand komt.

### De acht stappen om een toets te maken op een rijtje

Hieronder ziet u in een oogopslag welke 8 stappen er doorlopen worden om tot een goede en betrouwbare toets te komen. Vervolgens lichten we de 8 stappen nader toe.

#### 1 Toetsdoel bepalen

- Wat willen we meten?
- Waarom willen we meten?

#### 2 Toetsopzet bepalen

Denken aan:

- omvang
- inhoud
- niveau
- opgavenvorm

#### 3 Opgaven maken

Leerkrachten en toetsdeskundigen:

- maken individueel conceptopgaven
- screenen en bespreken (elkaars) opgaven
- komen gezamenlijk tot uitgewerkte opgaven

#### 4 Proefafname

- Een grote, representatieve groep leerlingen maakt de opgaven.

#### 5 Opgaven selecteren

Opgaven evalueren op basis van proefafname:

- welke opgaven zijn te makkelijk/te moeilijk?
- meten de opgaven wat we willen meten?
- maken de opgaven goed onderscheid tussen vaardige en minder vaardige leerlingen?
- zijn er reacties vanuit het veld op specifieke opgaven?

#### 6 Toets samenstellen

Denken aan:

- gevarieerde inhoud
- gevarieerd niveau (moeilijkheidsgraad)
- gemiddeld niveau
- gevarieerde opgavenvormen

#### 7 Normeringsonderzoek

Een grote, representatieve groep leerlingen maakt de opgaven uit de toetsboekjes.

Aan de hand van de resultaten wordt:

- de definitieve toets samengesteld
- de normering vastgesteld

#### 8 Uitgave

De uitgave bestaat uit:

- toetsmateriaal
- registratieformulieren
- handleiding
- inhoudsverantwoording
- wetenschappelijke verantwoording

## 1 Toetsdoel bepalen

Een school 'toetst niet om te toetsen'. Toetsen hebben altijd een doel. Bij het bepalen van het toetsdoel kijken we naar twee vragen:

### ***Wat willen we meten?***

Als we Begrijpend lezen willen meten, wat verstaan we hier dan precies onder? Definiëren we rekenen als 'rekenen in een context' of bijvoorbeeld als 'automatiseren van de basisoperaties'? Bij het definiëren van de vaardigheid vormen de kerndoelen en veelgebruikte methodes een leidraad. Voor de *makers* van toetsen is het belangrijk om zich steeds te realiseren wat de toets beoogt te meten. Zo willen we met de toets Taal voor kleuters taalvaardigheid meten en niet de werkhouding. Voor *gebruikers* is het ook goed het doel voor ogen te houden: met de toets Taal voor kleuters kan een leerkracht uitspraken doen over taalvaardigheid, niet over werkhouding. Uiteraard kan de concentratie van een kind invloed hebben op het toetsresultaat. Toch kunt u op basis van het resultaat op de toets Taal voor kleuters geen uitspraak doen over concentratievermogen. Om het concentratievermogen te meten, zet u bijvoorbeeld een observatie-instrument in.

### ***Waarom willen we meten?***

Cito biedt u toetsen waarmee u de vorderingen van leerlingen op een systematische wijze volgt. Aan de hand van toetsresultaten vergelijkt u (groepen) leerlingen met zichzelf, met anderen en met te bereiken doelen. Daardoor krijgt u informatie over het resultaat van het geboden onderwijs op leerling-, groeps- en schoolniveau. Ook kunt u aandachtspunten formuleren en waar nodig maatregelen treffen om uw onderwijs te verbeteren.

Er zijn verschillende redenen om te meten:

- *Het niveau van de leerling vaststellen:* als u een toets bij een leerling heeft afgenomen, bepaalt u het niveau van de leerling door de vaardigheidsscore en het vaardigheidsniveau vast te stellen. Deze gebruikt u om het onderwijsaanbod op de leerling af te stemmen. Let op! De LVS-toetsen zijn géén beheersingstoetsen. Leerlingen hoeven dus niet alle opgaven goed te maken. De toetsen bevatten opgaven van uiteenlopende moeilijkheidsgraad.
- *De ontwikkeling van de leerling volgen:* als er meerdere toetsen op hetzelfde vakgebied zijn afgenomen, is het vervolgens mogelijk om de ontwikkeling van de leerling te volgen in de tijd. Door deze ontwikkeling te vergelijken met de ontwikkeling van een leerling met een gelijk niveau (I t/m V of A t/m E) kan bepaald worden of deze ontwikkeling voldoende is.
- *Diagnostische informatie verzamelen:* sommige toetsen, zoals AVI/DMT of Rekenen-Wiskunde, leveren naast een vaardigheidsscore ook informatie op over welke (deel)gebieden (mogelijk) extra aandacht nodig hebben en/of welke strategie een leerling toepast. Met behulp hiervan kan een leerkracht zijn onderwijs nog beter afstemmen op de leerling. De toetsen van het Cito Volgstelsel primair onderwijs geven een eerste indicatie van de richting waarin de leerkracht verder kan kijken. Diagnostische toetsen of analyses bieden verdere uitkomst.
- *Voorspellen waar een leerling uitkomt/wat bij een leerling past:* de Entreetoets en de Eindtoets Basisonderwijs zijn voorbeelden van toetsen met een voorspellende waarde. Een belangrijke functie van de Eindtoets Basisonderwijs is het geven van het advies welk brugklastype het beste past bij de leerling.

## 2 Toetsopzet bepalen

Een toets moet voldoen aan verschillende eisen. Een groot deel van de eisen komt voort uit wensen van scholen en kenmerken van de leerlingen. Een ander deel bestaat uit vastgestelde eisen met betrekking tot bijvoorbeeld betrouwbaarheid en validiteit. Het doel van de toets bepaalt onder andere het type opgave, de moeilijkheidsgraad van de opgaven en de variatie in moeilijkheidsgraad tussen de opgaven.

De volgende aandachtspunten zijn in elk geval van belang:

- *Omvang:* hoe lang mag de afname van een toets duren en uit hoeveel opgaven moet de toets bestaan? Een taak uit de Rekentoets voor groep 3 is bijvoorbeeld korter dan een taak voor groep 8. Deels omdat de leerlingen in groep 3 minder lang achter elkaar kunnen werken, deels

omdat de stof in groep 8 meer omvat dan in groep 3. Om betrouwbaar te meten, is een minimaal aantal opgaven nodig. Om rekening te houden met de concentratie van de leerling en tegelijkertijd toch voldoende informatie te kunnen verzamelen, kan de toets worden aangepast door deze in meerdere delen ('taken') op te knippen.

- *Inhoud*: welke (sub)onderdelen komen in de toets? In een schema leggen we vast hoeveel opgaven van elke soort in de toets komen. Bijvoorbeeld: welke spellingscategorieën komen in de toets? En hoeveel opgaven van elke categorie?
- *Niveau*: bij welke leerlingen wordt de toets afgenomen? Hierbij wordt ook bekeken of deze leerlingen uiteenlopende niveaus hebben óf dat de niveaus juist dicht bij elkaar liggen.
- *Opgavenvorm*: hoe zien de opgaven eruit? Hierbij worden bijvoorbeeld de volgende keuzes gemaakt: open vragen of meerkeuzevragen, papier of digitaal, met of zonder antwoordblad, met of zonder plaatjes, individuele of groepsgewijze afname, soort vraagstelling.

### 3 Opgaven maken

Leerkrachten, pabodocenten en andere ervaringsdeskundigen maken in een constructiegroep onder leiding van toetsdeskundigen van Cito opgaven:

- Van Cito krijgen constructiegroepleden een constructieopdracht. Daarin staat bij Spelling bijvoorbeeld aangegeven voor welke spellingcategorieën ze opgaven moeten maken en welk vraagtype het betreft (met/zonder plaatjes, dictee /meerkeuze).
- De constructiegroepleden maken individueel opgaven en leveren deze in.
- Tijdens een constructiegroepvergadering bespreken de constructiegroepleden elkaars opgaven. Hierbij wordt o.a. kritisch gekeken naar inhoud, taalgebruik, niveau en context van de opgaven.
- Een Cito-medewerker werkt de opgaven verder uit en zet bijvoorbeeld tekenopdrachten uit bij tekenaars om de benodigde tekeningen te maken.

### 4 Proefafname

Nadat alle opgaven meerdere malen zijn bekeken en verbeterd, volgt de proeftoetsing. Van elke opgave willen we weten hoe moeilijk deze opgave is, hoe goed de opgave onderscheid maakt tussen vaardige en minder vaardige leerlingen en of de opgave meet wat we willen meten. Een grote, representatieve groep leerlingen maakt daarom de opgaven.

### 5 Opgaven selecteren

De gegevens uit de proefafname vormen de basis voor de selectie van opgaven.

- *Moeilijkheidsgraad*: van elke opgave bepalen we de moeilijkheidsgraad door te kijken naar het percentage leerlingen dat de opgaven goed maakt. Te makkelijke of te moeilijke opgaven worden verwijderd of bewaard voor een andere toets.
- *Alternatief*: bij meerkeuze-opgaven bekijken we of het alternatief dat correct is duidelijk door meer leerlingen gekozen wordt dan de overige alternatieven. Springt het 'correcte' alternatief (het juiste of beste alternatief) er niet duidelijk uit, dan is de opgave 'ambigu' en komt deze niet in aanmerking voor opname in de uiteindelijke toets. Daarnaast moeten alle alternatieven even waarschijnlijk zijn. Wanneer één van de alternatieven in het proefonderzoek zelden of nooit is gekozen, komt deze opgave ook niet in de toets.
- *Vaardigheid*: ook kijken we welke leerlingen een opgave goed of juist fout maken. Maken vaardige leerlingen relatief eenvoudige opgaven bijvoorbeeld inderdaad goed? Hoe vaardiger de leerling, des te groter de kans moet zijn om de opgave goed te kunnen maken.
- *Reactie veld*: uiteraard nemen we ook reacties uit het veld mee. Welke opgaven zijn volgens leerkrachten minder goed of juist wel goed? Soms verwijderen we opgaven die volgens de statistieken heel goed functioneren, omdat leerkrachten ze als 'slechte opgaven' ervaren.

Al deze informatie gebruiken we om te komen tot een verzameling opgaven waaruit een toets kan worden samengesteld. Het is belangrijk om te realiseren dat het antwoord van een leerling op één enkele opgave weinig zegt over de werkelijke vaardigheid van een kind. Daarbij kunnen immers allerlei toevalligheden een grote rol spelen. Het gaat juist om de antwoorden op een grote reeks van deze opgaven. Daardoor worden toevalligheden uitgesloten en krijgen leerkrachten betrouwbare informatie.

## 6 Toets samenstellen

Uit de geselecteerde opgaven wordt een toets samengesteld. Hierbij wordt gekeken naar:

- *Gevarieerde inhoud*: de verschillende onderdelen van een vakgebied moeten volgens afgesproken verhoudingen in de toets naar voren komen.
- *Gevarieerd niveau (moeilijkheidsgraad)*: een toets bevat opgaven van verschillende moeilijkheid, van gemakkelijk tot relatief moeilijk voor de doelgroep. De toets bevat gemakkelijke, moeilijkere en opgaven van gemiddelde moeilijkheidsgraad. In de toets worden deze meestal afwisselend aangeboden om te voorkomen dat een leerling verveeld of gefrustreerd raakt.
- *Gemiddeld niveau*: voor de totale toets geldt dat een gemiddelde leerling een ruime meerderheid van de vragen goed moet kunnen maken.
- *Gevarieerde opgavenvormen*: afwisseling in opgavevorm, bijvoorbeeld lange vragen versus korte vragen en/of opgaven met versus opgaven zonder plaatje. De toets ziet er zo aantrekkelijk uit, waardoor de leerling zijn aandacht beter kan vasthouden gedurende de hele toets.

## 7 Normeringsonderzoek

- Een grote, representatieve groep leerlingen uit de groep waarvoor de toets bestemd is, maakt de opgaven uit de toetsboekjes. Hierbij worden alle aanwijzingen voor de afname uit de handleiding opgevolgd. Alleen dan is het mogelijk de toetsprestaties van leerlingen met elkaar te vergelijken. Dat betekent dat ook bij de uiteindelijke afname van de toets de normeringstabellen alleen gebruikt kunnen worden als de toets afgenomen wordt volgens de aanwijzingen in de handleiding.
- De leerlinggegevens worden teruggestuurd naar Cito. Cito berekent vervolgens de vaardigheidsniveaus en stelt de normgroepen (I t/m V en A t/m E) vast. Deze normgegevens komen dus tot stand op basis van de gegevens uit het normeringsonderzoek.

## 8 Uitgave

De uitgave bestaat uit:

- *Toetsmateriaal*: de toetsboekjes, digitale toetsen en eventuele antwoordbladen of voorleesbladen.
- *Registratieformulieren*: hierop kan een leerkracht bijhouden welke vorderingen een leerling maakt. Dit kan uiteraard ook met behulp van een computerprogramma, zoals het Computerprogramma LOVS.
- *Handleiding*: een duidelijke omschrijving over hoe de toets moet worden afgenomen. Daardoor is de manier van afname op verschillende scholen en op verschillende toetsmomenten zo gelijk mogelijk (gestandaardiseerde afname). Dit maakt een eerlijke vergelijking van resultaten mogelijk. Door aanwijzingen over de scoring bepalen leerkrachten de juiste score van de leerling en kunnen zij de resultaten gebruiken om de leerling te volgen.
- *Inhoudsverantwoording*: hierin staat beschreven hoe de toets tot stand is gekomen. Bij sommige toetsen staat ook informatie waarmee scholen betekenis kunnen geven aan een toetsscore. Wat kan een leerling bij een bepaalde vaardigheidsscore? En welke moeilijkheidsgraad hebben de opgaven in de toets?
- *Wetenschappelijke verantwoording*: hierin staan onder andere de normeringsonderzoeken beschreven. De wetenschappelijke verantwoording wordt samen met het overige toetsmateriaal ter beoordeling voorgelegd aan de Commissie Testaangelegenheden Nederland (COTAN).

## Wilt u meer weten?

Graag verwijzen wij u naar de onderstaande internetadressen voor meer informatie.

- [www.cito.nl](http://www.cito.nl)
- <http://toetswijzer.kennisnet.nl/html/toetsconstructie/home.htm>
- <http://toetswijzer.kennisnet.nl/html/toetsenopschool/default.shtm>