

De diagnostische tussentijdse toets

Tussenstand in ontwikkeling





Samenvatting

In het kader van het opbrengstgericht werken op school heeft het ministerie van OCW het initiatief genomen tot de ontwikkeling van een diagnostische tussentijdse toets (DTT) voor de doorstroom relevante vakken Nederlands, Engels en wiskunde/rekenen van vmbo, havo en vwo (Ministerie van OCW, 2012). Cito heeft daartoe in 2012 in opdracht van OCW en College voor Examens een voorstudie uitgevoerd, waarvan dit document de bevindingen samenvat. Met dit initiatief gaf OCW invulling aan het Actieplan Beter Presteren (2011). De kaders voor de DTT worden gevormd door het concept wetsvoorstel 'Leerlingvolgsysteem en diagnostische tussentijdse toets VO' (Ministerie van OCW, 2012).

Zoals verwoord in het concept wetsvoorstel is de DTT een diagnostische tussentijdse toets die aan het einde van de onderbouw zal worden afgenomen om leerlingen en docenten te informeren over de prestaties van de leerling, zodat de docent gericht actie kan ondernemen voor verbetering en maatwerk. Hiermee is de diagnostische tussentijdse toets een ijkpunt in de doorlopende leerlijn van eindtoets primair onderwijs naar eindexamen voortgezet onderwijs. Het doel van de DTT is diagnostisch en de toetsing heeft een formatief in plaats van summatief karakter. Summatieve toetsen hebben als functie de leerprestaties van een leerling te beoordelen en vervolgens een beslissing te nemen met betrekking tot selectie, classificatie, plaatsing of certificering (Sanders, 2011). Toetsen hebben een formatieve functie als de resultaten gebruikt kunnen worden om richting en vorm te geven aan het onderwijsleerproces. Deze toetsen worden ook wel formatieve toetsen genoemd. De diagnose levert een sterkte-zwakke profiel van de leerling voor het betreffende vak. Behalve docenten, leerlingen en ouders kunnen ook schoolleiders en beleidsmakers de informatie uit de DTT gebruiken, bijvoorbeeld voor schoolbeleid of landelijk onderwijsbeleid. Het is echter nadrukkelijk de bedoeling dat de DTT gebruikt zal worden om opbrengstgericht te werken. Scholen beslissen zelf in hoeverre ze resultaten openbaar maken.

De DTT is een adaptieve toets die betrekking heeft op de doorstroom relevante vakken Nederlands, Engels en wiskunde/rekenen. De toets kent vijf verschillende niveaus, namelijk vmbo-bb, vmbo-kb en vmbo-gl/tl, havo en vwo en zal worden afgenomen in het tweede leerjaar van vmbo respectievelijk het derde leerjaar van havo/vwo. De DTT wordt digitaal afgenomen via het nieuwe ICT-platform FACET (CvE, 2012).

De ontwikkeling van de DTT is gebaseerd op het raamwerk voor *Evidence Centered Design* (Mislevy, Steinberg, Breyer, Almond, & Johnson, 2002). Dit heeft geleid tot het onderscheiden van

- een leerlingmodel dat de aard van de te toetsen vaardigheid beschrijft in termen van attributen en deelattributen,
- een taakmodel dat de sleutelkenmerken van taken bevat en
- een rapportagemodel dat aangeeft hoe de toetsuitkomsten zo beschreven en gevisualiseerd kunnen worden dat de docent de diagnoses adequaat kan gebruiken.

Voor het onderdeel schrijfvaardigheid van Nederlands en Engels en voor het vak wiskunde is in deze voorstudie een start gemaakt met de ontwikkeling van het leerlingmodel, het taakmodel en het rapportagemodel.

De ontwikkeling van de DTT Nederlands heeft zich geconcentreerd op het domein schrijfvaardigheid. Als basis voor de attributen is het model van Deane (2011) genomen. Deze attributen zijn nader gespecificeerd in deelattributen die voortkomen uit de concept-tussendoelen onderbouw (SLO, 2012) met betrekking tot schrijfvaardigheid. Voor de opzet van deze DTT wordt gedacht aan een getrapt model, waarin de leerling eerst een schrijfofdracht maakt waarmee het beginniveau wordt bepaald en vervolgens een adaptieve deeltoets doorwerkt.

Net als bij Nederlands heeft het werk aan de DTT Engels zich op het domein schrijfvaardigheid gericht. Bij de opzet van het leerlingmodel is eveneens gekozen voor het model van Deane voor de bepaling van de attributen. De deelattributen zijn afgeleid uit de concept-tussendoelen onderbouw. Ook voor schrijfvaardigheid Engels gaan de gedachten uit naar een getrapt model, bestaande uit een open schrijfofdracht gevolgd door adaptieve deeltoetsen, waarvan het beginniveau mogelijk mede bepaald wordt door de resultaten op de schrijfofdracht.

Voor de DTT wiskunde zijn de attributen van het leerlingmodel net als bij Nederlands en Engels gebaseerd op de vakinhoudelijke domeinen zoals geformuleerd in de concept-tussendoelen onderbouw (SLO, 2012). Daarnaast is als invalshoek voor het leerlingmodel gekozen voor de wiskundig-didactische aspecten structuur, ambiguïteit en samenhang, die de inhoudelijke domeinen overstijgen en moeilijkheden van leerlingen kunnen verklaren.

Voor de verdere ontwikkeling en implementatie van de DTT zal nog veel werk verzet moeten worden. Aandachtspunten voor Cito hierbij zijn opgave- en itemontwikkeling, psychometrische zaken als kalibratie en validatie, toetsassemblage en het vormgeven van adaptiviteit. Dit vraagt om een gefaseerd groeimodel dat een doorlooptijd van enkele jaren kent. Na deze voorstudie volgen onder andere een try-out fase, een fase van grootschalig itemontwikkeling, een pretest, een eerste afname en de uiteindelijke operationele fase. Het lijkt efficiënt om deze fasen niet voor alle vakken en domeinen gelijktijdig te doorlopen, maar via een gefaseerde aanpak. Op deze manier wordt extreme belasting voor scholen, toetsconstructeurs, Toetswijzercommissies en Vaststellingscommissies voorkomen. In de volgende fasen wordt een voorstudie verricht naar de overige vaardigheden (rekenen, en voor Nederlands en Engels lees-, spreek- en luistervaardigheid). Mede op basis van de voorstudies stelt de toetswijzercommissie een toetswijzer op die functioneert als basis voor de toetsconstructie. Via onder andere de nieuwsbrief VO van OCW wordt nadere informatie over het wetsvoorstel en de implementatie van de DTT, verspreid.

Inhoudsopgave

Samenvatting.....	1
Inhoudsopgave	3
1 Diagnostische tussentijdse toets: waarom, wat en hoe?.....	5
1.1 Waarom een diagnostische tussentijdse toets?	5
1.2 Wat is de DTT?	5
1.3 Hoe kan de DTT tot stand komen?	7
2 Naar een diagnostische tussentijdse toets voor Nederlands	10
2.1 Aanzet tot een taakmodel voor schrijfvaardigheid Nederlands	12
2.2 Naar een rapportagemodel schrijfvaardigheid Nederlands	13
3 Naar een diagnostische tussentijdse toets voor Engels	15
3.1 Een leerlingmodel voor schrijfvaardigheid Engels	15
3.2 Aanzet tot een taakmodel voor schrijfvaardigheid Engels	16
3.3 Naar een rapportagemodel schrijfvaardigheid Engels	17
4 Naar een diagnostische tussentijdse toets voor wiskunde	19
4.1 Een leerlingmodel voor de DTT wiskunde	19
4.2 Aanzet tot een taakmodel voor de DTT wiskunde	20
4.3 Naar een rapportagemodel voor de DTT wiskunde	21
5 Ontwikkeling en implementatie in praktijk	23
5.1 Ontwikkeling en uitwerking	23
5.2 Tijdpad naar implementatie	24
Literatuur.....	26

1 Diagnostische tussentijdse toets: waarom, wat en hoe?

In het kader van het opbrengstgericht werken op school heeft het ministerie van OCW het initiatief genomen voor de ontwikkeling van een diagnostische tussentijdse toets in het voortgezet onderwijs en voor de doorstroom relevante vakken Nederlands, Engels en wiskunde/rekenen in het bijzonder. Als eerste stap heeft Cito in 2012 in opdracht van OCW en College voor Examens een voorstudie uitgevoerd waarin het concept van een diagnostische tussentijdse toets nader wordt verkend en op haalbaarheid wordt beproefd.

Deze publicatie vat de bevindingen van de voorstudie samen. In dit eerste hoofdstuk wordt ingegaan op het waarom, het wat en het hoe van de diagnostische tussentijdse toets, afgekort tot DTT. Hoewel het gaat over toetsen voor verschillende vakken en op verschillende niveaus, spreken we in dit verslag over *de* diagnostische tussentijdse toets in enkelvoud.

1.1 Waarom een diagnostische tussentijdse toets?

De kwaliteit van het Nederlandse onderwijs vraagt om voortdurende aandacht. In dit kader heeft het vorige kabinet in het *Actieplan Beter Presteren* en het *Actieplan Leraar 2020—een krachtig beroep!* een aantal doelen geformuleerd ten aanzien van het onderwijs. Onderdeel daarvan is dat men een tussentijdse meting in de onderbouw van het voortgezet onderwijs zal invoeren die docenten en leerlingen helpt het basisoniveau in de zogeheten doorstroom relevante vakken te versterken. De invoering van de diagnostische tussentijdse toets (DTT) is bedoeld als stimulans tot de verdere realisatie van een cultuur van opbrengstgericht werken.

De diagnostische tussentijdse toets is een ijkpunt in de doorlopende leerlijn van eindtoets primair onderwijs naar eindexamen voortgezet onderwijs. De aard en functie van de DTT zijn echter anders dan die van de eindtoets PO en het eindexamen VO. Terwijl de eindtoets PO en het eindexamen VO afsluitende toetsen zijn, is de DTT primair een diagnostisch instrument. De toets dient docenten en leerlingen inzicht te geven in hun positie in het leerproces op weg naar het eindexamen. Het zwaartepunt ligt dus op de informatiewaarde over de individuele leerling. De informatie die de diagnostische toets oplevert moet de docent in staat stellen gerichte actie te ondernemen voor verbetering, maatwerk of bijvoorbeeld verrijking voor excellente leerlingen. Op deze wijze moet de diagnostische tussentijdse toets bijdragen aan het realiseren van hogere prestaties van leerlingen in het voortgezet onderwijs (Ministerie van OCW, 2012).

Als gevolg van het wetsvoorstel zal de uitvoering van de DTT berusten bij het College voor Examens (CvE). Stichting Cito zal onder regie van het College voor Examens (CvE) en in nauwe samenspraak met het onderwijsveld de diagnostische tussentijdse toets ontwikkelen. De basis hiervan vormen de door SLO geformuleerde concepttussendoelen onderbouw.

1.2 Wat is de DTT?

De DTT is een diagnostische tussentijdse toets die wordt afgenomen aan het einde van de onderbouw (leerjaar 2 vmbo, leerjaar 3 havo en vwo), dus vóór de leerlingen in het derde leerjaar vmbo en het vierde leerjaar havo en vwo gericht gaan toewerken naar de examens. De toets heeft betrekking op de doorstroom relevante vakken Nederlands, Engels en wiskunde/rekenen en kent vijf verschillende niveaus, namelijk vmbo-bb, vmbo-kb, vmbo-gl/tl, havo en vwo.

De DTT wordt gekenmerkt door de vier trefwoorden diagnostisch, tussentijds, adaptief en digitaal, die we nu achtereenvolgens kort bespreken.

Diagnostisch

Het doel van de DTT is primair diagnostisch: de toets is erop gericht om het leerproces van de leerlingen te ondersteunen door gerichte diagnostische informatie op te leveren waarmee dat leerproces verder ingericht kan worden. Deze diagnose betreft vooral een sterkte-zwakke profiel van de leerling. Het gaat dus om formatieve in plaats van summatieve toetsing, die diagnostische informatie oplevert. De Bruyn, Ruijsenaars, Pameijer en Van Aarle (2003) beschrijven diagnosticeren als een cyclisch proces van verhelderen, verkennen, verklaren en indiceren:

- Verhelderen: Waar zit de leerling ten opzichte van het hoofdoel of de hoofdoelen?
- Verkennen: Hoe presteert de leerling op subdoelen?
- Verklaaren: Hoe komt het dat de leerling sterk of zwak is binnen dit (sub)domein?
- Indiceren (Adviseren): Op welke doelen moet de interventie zich richten?

Deze diagnostische informatie is in de eerste plaats van belang voor docenten, leerlingen en ouders. De DTT geeft hen inzicht in de sterke en zwakke punten van de leerling aan het einde van de onderbouw VO. De rapportage zal dus individuele rapporten bevatten, waarin de leerprestaties voor elk doorstroom relevant vak worden weergegeven in relatie tot de standaard of norm van de referentiegroep. Deze informatie kan de docent gebruiken om de voortgang van de leerlingen in kaart te brengen en het onderwijsproces en het didactisch handelen zo nodig hierop aan te passen. Ook kan rapportage op groepsniveau plaatsvinden, waaruit blijkt hoe een groep presteert op de verschillende (deel-)domeinen binnen een vak. Dit maakt het identificeren van leerlingen en groepen leerlingen met specifieke zwakke of sterke punten eenvoudig. Voor de leerlingen zal het werken aan de toets een leereffect hebben; daarnaast zullen leerling en ouders kunnen profiteren van de diagnostische informatie, op basis waarvan bijvoorbeeld het studiegedrag kan worden aangepast.

Naast deze primaire groep van belanghebbenden is de DTT ook relevant voor vaksecties en schoolleiding, die uit de diagnose van de DTT kunnen afleiden hoe de sterke en zwakke punten van hun leerlingen zich verhouden tot die van andere scholen en tot het landelijke beeld. Scholen kunnen zo hun eigen bevindingen vergelijken met die van vergelijkbare groepen scholen en deze informatie gebruiken voor het opbrengstgericht werken. Het gaat hier overigens om niet-openbare, schooleigen informatie; scholen beslissen zelf in hoeverre ze hun resultaten openbaar maken. Tevens kunnen de resultaten van de DTT beleidsmakers inzicht geven in het functioneren van het onderwijsstelsel als geheel. Door de informatie over de leerprestaties verkregen uit DTT te aggregeren op stelselniveau, ontstaat zicht op het functioneren van het stelsel en op mogelijke consequenties voor het te voeren beleid.

Tussentijds

Het tweede trefwoord, tussentijds, geeft aan dat de DTT er niet is om leerlingen te certificeren of te classificeren, maar om een tussenbalans op te maken. Deze tussenbalans identificeert sterke en zwakke punten en kan aanleiding zijn tot het opzetten van remediërende activiteiten, of het bieden van aanvullende uitdagingen. De concepttussendoelen onderbouw, die door SLO zijn opgesteld, vormen hierbij het vakinhoudelijke vertrekpunt.

Adaptief

Het derde trefwoord, adaptief, geeft aan dat de toets zich aanpast aan het niveau van de leerling. Immers, als alle leerlingen op dezelfde manier worden getoetst met een diagnostische toets die alle niveaus en onderwerpen dekt, dan zou de toets te lang worden en tevens veel opgaven bevatten die voor een bepaalde leerling niet relevant zijn. Van een efficiënt diagnose-instrument is dan geen sprake. In de DTT zal de adaptiviteit daarom allereerst gestalte krijgen door de toets bij aanvang aan te passen aan de leerling. Daarnaast wordt gedurende de toetsafname nagegaan wat we al weten over de individuele leerling op grond van zijn/haar antwoorden. Zo wordt bepaald welke opgaven (of toetsdelen) we nog moeten voorleggen aan deze leerling om tot betrouwbare en valide informatie te komen over zijn sterke en zwakke punten. Op deze manier wordt meer verfijnde informatie verkregen in dezelfde afnametijd. De adaptiviteit heeft primair tot doel de potentiële zwakke en sterke punten van de leerling in kaart te brengen.

Digitaal

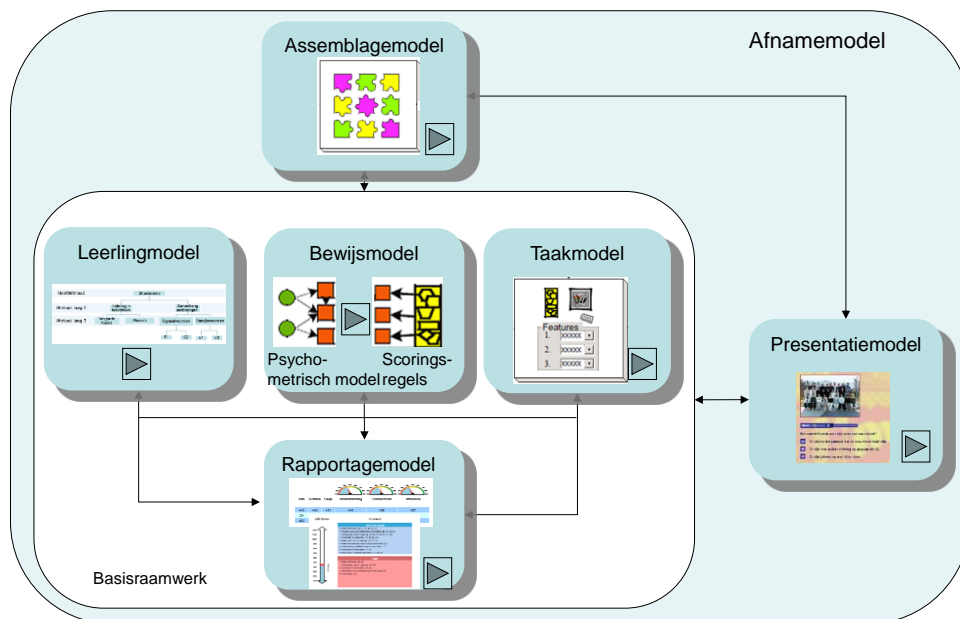
Het vierde en laatste trefwoord van de DTT, digitaal, geeft aan dat het gaat om een toets die digitaal wordt afgenomen en beoordeeld. De beoogde adaptiviteit gedurende de toetsing is alleen op deze manier te realiseren. Verdere voordelen van digitale afname zijn dat de belasting voor school en docent beperkt blijft en dat de verwerking van de gegevens op verschillende schaal efficiënt kan plaatsvinden. Leerlingen en docenten krijgen dus snel toegang tot de rapportage. Ook is het mogelijk om opgaven authentieker te presenteren. De DTT wordt afgenomen in het nieuwe ICT-platform FACET (CvE, 2012).

Vanwege het diagnostische en tussentijdse karakter van de meting is de DTT geen instrument om het niveau van een school of een docent te bepalen. Het is aan de scholen zelf om te beslissen in hoeverre

schoolgegevens uit de DTT openbaar gemaakt worden. Omdat de DTT bedoeld is als hulpmiddel bij opbrengstgericht werken, zal de inspectie in het toezicht zich vooral richten op de vraag of scholen dit instrument adequaat voor dat doel inzet (OCW, 2012). Een tweede risico is dat de docent overgaat tot het bekende “teaching to the test” door gericht variaties van items uit de toets met leerlingen te gaan oefenen of het lesprogramma erop in te richten. Graag benadrukken we dat de diagnose van de DTT geen doel op zichzelf is, maar slechts een hulpmiddel is voor betere prestaties op termijn.

1.3 Hoe kan de DTT tot stand komen?

Een landelijke diagnostische, digitale en adaptieve tussentijdse toets is een nieuw fenomeen in het Nederlandse onderwijs. De ontwikkeling van een dergelijke innovatieve toets vraagt om een solide onderliggend raamwerk. Voor dit doel is een aangepaste versie ontwikkeld van het raamwerk voor *Evidence Centered Design* (Mislevy, Steinberg, Breyer, Almond, & Johnson, 2002; Mislevy, Steinberg, Almond & Lukas, 2006). Figuur 1-1 geeft de belangrijkste componenten van dit raamwerk weer.



Figuur 1-1: Componenten van het DTT raamwerk (naar Mislevy et al., 2002).

De basis voor de DTT wordt gevormd door een *leerlingmodel*. Een leerlingmodel beschrijft de aard van de vaardigheid die getoetst wordt in de vorm van een samenhangende structuur van hoofd-, deel- en subattributen. De attributen verwijzen naar deelvaardigheden, strategieën en kenniselementen zoals regels, procedures, begrippen en principes die deel uitmaken van de totale vaardigheid of die ingezet moeten

worden bij taken binnen het domein of vakgebied. In

- 1 **Sociale vaardigheden:** schrijfsituatie en schrijfdoelen
- 2 **Conceptuele vaardigheden:** beslissen over concepten en
- 3 **Tekstuele vaardigheden:** schrijfproduct plannen en
- 4 **Verbale vaardigheden:** gedachten omzetten in woorden en
- 5 **Orthografische vaardigheden:** regels en conventies van taal beheersen

Figuur 1-2 staan bijvoorbeeld vijf attributen die deel uitmaken van een leerlingmodel voorschrijfvaardigheid Nederlands. Het is geïnspireerd op literatuur over schrijfvaardigheid (Deane et al., 2008; Deane, 2011) en wordt verder uitgewerkt in hoofdstuk 2. Deelattributen van het attribuut Tekstuele vaardigheden zouden bijvoorbeeld zijn het indelen van een tekst in alinea's en het verbinden van verschillende alinea's.

- 1 **Sociale vaardigheden:** schrijfsituatie en schrijfdoelen
- 2 **Conceptuele vaardigheden:** beslissen over concepten en
- 3 **Tekstuele vaardigheden:** schrijfproduct plannen en
- 4 **Verbale vaardigheden:** gedachten omzetten in woorden en
- 5 **Orthografische vaardigheden:** regels en conventies van taal beheersen

Figuur 1-2: Attributen van schrijfvaardigheid Nederlands naar Deane (2011)

In een leerlingmodel worden ook leerprogressies beschreven die leerlingen tijdens de verwerving van een vaardigheid doormaken. Daarbij wordt aandacht besteed aan de toenemende complexiteit van taaksituaties en taakcondities waarin de leerling leert te presteren. In een leerlingmodel voor een toets zijn alle attributen en onderliggende deelattributen opgenomen waarover uitspraken gedaan moeten worden. Attributen zijn scorebaar via items op de toets en op basis daarvan wordt diagnostische feedback gegeven die inzicht geeft in sterke en zwakke punten van leerlingen.

Een tweede onderdeel van het raamwerk is het *taakmodel*. Dit beschrijft de sleutelkenmerken van taken en taaksituaties waarmee de benodigde gegevens worden verzameld die aangeven waar de leerling staat met betrekking tot de attributen in kwestie. Daarnaast beschrijft het taakmodel de wijze waarop taken worden aangeboden en de van de leerling verwachte responsen. Het is daarmee de basis voor de ontwikkeling van opgaven.

Een derde onderdeel van het raamwerk is het *bewijsmodel*. Dit bevat regels voor toekenning van scores en een meetmodel waarmee waargenomen metingen worden verbonden aan de veronderstelde attributen in het leerlingmodel.

In het *rapportagemodel* wordt aangegeven op welke wijze de toetsuitkomsten beschreven en gevisualiseerd kunnen worden, zodat de docent de informatie uit de DTT kan gebruiken om richting en vorm te geven aan zijn/haar onderwijs. Het spreekt voor zich dat de doeltreffendheid van de diagnostische toets sterk afhangt van de kwaliteit van de rapportage, die in elk geval duidelijk moet maken waar de leerling staat ten opzichte van de leerdoelen.

Het *assemblagemodel* beschrijft hoe de opgaven in de toets worden geselecteerd, rekening houdend met voldoende dekking van de attributen en met de vaardigheid van de leerling (adaptieve opgavenselectie), met als doel maximale informatie te genereren in een bepaalde tijd.

Het *presentatiemodel* beschrijft hoe de opgaven worden voorgelegd aan de leerlingen, dus welke technische en materiële voorwaarden nodig zijn om de taakuitvoering op te roepen.

Tot slot is er een *afnamemodel* nodig. Het afnamemodel beschrijft hoe de voorgaande modellen samenwerken bij de toetsafname en in welke afnameomgeving de toets wordt afgenomen. Denk aan zaken als het gebruikte platform, dat in het geval van de DTT de FACET omgeving zal zijn, aan de administratieve vereisten en aan het moment van afname in het schooljaar.

Bij ontwikkeling van de DTT dienen de bovenstaande modellen in onderlinge samenhang te worden ontwikkeld. De meetpretenties, zoals beschreven in het leerlingmodel, dienen te worden waargemaakt in passende taken (opgaven) en daarbij passende bewijsmodellen. Gedurende de ontwikkeling worden de deelmodellen steeds explicieter uitgewerkt door continue afstemming en bijstelling. Deze dynamische vorm van toetsontwikkeling bevordert de toetskwaliteit en helpt ontwikkelaars om hun werk conceptueel en operationeel te structureren en optimaal af te stemmen. De verschillende onderdelen van het raamwerk vragen verschillende typen ontwerpers: toetsdeskundigen, onderwijskundigen, didactici, leraren, materiedeskundigen, psychometrici en ICT-deskundigen. Om deze expertise te combineren in een innovatieve en functionele diagnostische tussentijdse toets is de uitdaging waarvoor we staan.

2 Naar een diagnostische tussentijdse toets voor Nederlands

In dit hoofdstuk gaan we na hoe de DTT voor Nederlands op basis van de uitgangspunten van hoofdstuk 1 gestalte kan krijgen. Aan de orde komen achtereenvolgens een leerlingmodel, een aanzet tot een taakmodel en een voorbeeld van de mogelijke rapportage. Hierbij merken we op dat dit alles nog volop in ontwikkeling is en dus slechts een indicatief karakter heeft; een en ander wordt in hoog tempo nader geconcretiseerd. Tijdens de voorstudie zijn ook twee veldraadplegingen verricht, waarbij de ideeën zijn voorgelegd aan docenten. Hier zijn zinvolle opmerkingen gemaakt die mee worden genomen in het vervolg. Over het algemeen was men enthousiast over de uitwerkingen.

Het vakonderwijs Nederlands in het voortgezet onderwijs concentreert zich traditioneel op vier taalvaardigheden: de productieve vaardigheden Spreken, Gesprekken voeren en Schrijven en de receptieve vaardigheden Lezen en Luisteren.

Het toetsen van receptieve vaardigheden is door middel van lees- en luistertoetsen relatief eenvoudig uit te voeren. De toetsing van productieve vaardigheden is echter lastig. Het beoordelen van geproduceerde taal is bijvoorbeeld zeer arbeidsintensief voor de docent. Dit geldt in het bijzonder voor schrijfvaardigheid, tenzij onderdelen van het beoordelen geautomatiseerd kunnen worden. Hoewel geautomatiseerde afname en beoordeling nog verre van eenvoudig zijn, zijn daar wel mogelijkheden voor. Daarmee lijkt het verkennen van de mogelijkheden van een diagnostische tussentijdse toets voor schrijfvaardigheid als eerste vaardigheid bij Nederlands het meest interessant. In dit hoofdstuk beperken we ons dan ook tot schrijfvaardigheid. Het leerlingmodel voor schrijfvaardigheid zal via een vergelijkbare werkwijze naar verwachting dan ook worden verbreed naar leerlingmodellen voor lees- en luistervaardigheid. Daarnaast worden de mogelijkheden voor spreekvaardigheid onderzocht, waarbij ook naar andere toetsvormen dan een digitale adaptieve toets zal worden gekeken.

Een leerlingmodel voor schrijfvaardigheid Nederlands

Een leerlingmodel beschrijft de aard van de vaardigheid die getoetst wordt in de vorm van een samenhangende structuur van hoofd- en deelattributen. Bij de opzet van het voorlopige leerlingmodel schrijfvaardigheid Nederlands is gekozen voor de attributen van het model van Deane en zijn de deelattributen bepaald vanuit de concept-tussendoelen.

Attributen op basis van het model van Deane

SLO heeft concept-tussendoelen opgesteld voor schrijfvaardigheid Nederlands (SLO, 2012). Deze concept-tussendoelen vormen de basis van het leerlingmodel. Om het leerlingmodel de noodzakelijke cognitieve grondslag te geven, worden de attributen van het model beschreven in termen van het cognitieve model van Deane (Deane et al., 2008; Deane, 2011), dat in hoofdstuk 1 al kort aan de orde is geweest (zie

- 1 **Sociale vaardigheden:** schrijfsituatie en schrijfdoelen
- 2 **Conceptuele vaardigheden:** beslissen over concepten en
- 3 **Tekstuele vaardigheden:** schrijfproduct plannen en
- 4 **Verbale vaardigheden:** gedachten omzetten in woorden en
- 5 **Orthografische vaardigheden:** regels en conventies van taal beheersen

Figuur 1-2). Sinds 2007 worden volgens dit model in de Verenigde Staten formatieve en summatieve toetsen voor schrijven ontwikkeld, die informatie opleveren over de kwaliteit van zowel het schrijfproduct als het schrijfproces van de leerling. Hierbij wordt gestreefd naar een toetssysteem dat documenteert wat de leerlingen hebben gepresteerd ('assessment of learning'), dat ondersteunend is voor het plannen en bijstellen van didactisch handelen ('assessment for learning') en dat door leerlingen wordt gezien als een waardevolle bezigheid op zich ('assessment as learning'). Het model omvat de vijf attributen sociale en strategische vaardigheden, conceptuele vaardigheden, tekstuele vaardigheden, verbale vaardigheden en orthografische vaardigheden.

Bij het beoordelen van schrijfvaardigheid volgens dit model kunnen observaties gedaan worden op alle vijf attributen, die tevens mogelijke knelpunten in het individuele schrijfproces blootleggen. Bij het schrijven van eenvoudige argumentatieve teksten kunnen bijvoorbeeld moeilijkheden op verschillende lagen worden vastgesteld: een leerling heeft bijvoorbeeld geen duidelijk beeld van de groep tot wie hij zich richt, heeft moeite met het verzamelen en ordenen van argumenten (pro, contra), slaagt er niet in de bedachte argumenten ordelijk en samenhangend op papier te zetten, vindt niet de passende woorden, bouwt niet de benodigde zinnen, of maakt veel spelfouten dan wel fouten in de zinsbouw.

Deelattributen op basis van concept-tussendoelen

Voor de deelattributen zijn de concept-tussendoelen van het SLO geanalyseerd en ingepast in de hierboven beschreven vijf lagenstructuur. Alle genoemde concept-tussendoelen zijn in het geschetste leerlingmodel geplaatst. De op deze manier verkregen deelattributen staan veelal niet op zichzelf maar interacteren met elkaar.

Het leerlingmodel met attributen en deelattributen

In totaal kent het leerlingmodel schrijfvaardigheid Nederlands zo vijf attributen en 23 deelattributen. Tabel 2-1 geeft een overzicht van dit model.

Het kan verleidelijk zijn om dit model op te vatten als een nauwkeurige chronologische beschrijving van de stappen die een schrijver successievelijk neemt, of zelfs om het model prescriptief te hanteren.

Taalverwerving verloopt echter schoksgewijs (Clark & Clark, 1977; Pinker, 1996). Onderzoek naar schrijfprocessen laat tevens zien dat er geen vaste volgorde bestaat waarin de schrijver zich bezig houdt met de deelcomponenten (Braaksma, Rijlaarsdam, van den Bergh, & van Hout-Wolters, 2007); evenmin is er reden om te veronderstellen dat er een ideale volgorde zou bestaan in de deeltaken die een schrijver gedurende het schrijfproces uitvoert (Schilperoord, 1996).

Een mogelijke meerwaarde van dit leerlingmodel is dat we lees- en schrijfvaardigheid in samenhang met elkaar kunnen toetsen en diagnostische uitspraken kunnen doen op het niveau van de verschillende lagen. Waarschijnlijk kunnen ook luistervaardigheid en spreekvaardigheid in ditzelfde model worden ondergebracht.

Tabel 2-1: Een leerlingmodel voor schrijfvaardigheid Nederlands (naar Deane, 2011)

<p>1. Sociale en strategische vaardigheden</p> <p>Correct inschatten in welke schrijfsituatie je je begeeft, wat de schrijfdoelen zijn, welk tekstgenre dit betreft en welke verplichtingen je aan je lezer hebt.</p>	<p>1.1 voorkennis bij de lezers inschatten</p> <p>1.2 verhouding schrijver - lezers goed kunnen inschatten</p> <p>1.3 kiezen voor een formele of informele schrijfstijl</p> <p>1.4 toonzetting kiezen</p>
<p>2. Conceptuele vaardigheden</p> <p>Begrijpen en beslissen welke ideeën, concepten en gedachtegang moeten worden overgebracht aan de lezersdoelgroep</p>	<p>2.1 schrijfdoel bepalen</p> <p>2.2 bepalen wat voor soort tekst gewenst is</p> <p>2.3 deugdelijke argumenten geven</p> <p>2.4 optimale volgorde van tekstuele elementen bepalen</p>
<p>3. Tekstuele vaardigheden</p> <p>Een schrijfproduct plannen en realiseren waarin beoogde concepten en gedachtegang gestructureerd worden overgebracht en waarin bepaalde beoogde doelen aan de kant van de lezer centraal staan</p>	<p>3.1 een goede tekstopbouw realiseren</p> <p>3.2 logische volgorde van inhoudelijke elementen bepalen</p> <p>3.3 genre-conventies hanteren en handhaven</p> <p>3.4 een goede alinea-indeling geven</p> <p>3.5 verbanden tussen tekstdelen en alinea's aanduiden</p> <p>3.6 lay-out gebruiken die de tekstopbouw ondersteunt</p>
<p>4. Verbale vaardigheden</p> <p>De gedachten omzetten in concrete woorden en zinnen die de bedoelde betekenis overbrengen.</p>	<p>4.1 correcte zinsbouw hanteren</p> <p>4.2 grammaticale fouten vermijden</p> <p>4.3 een vlotte schrijfstijl hanteren</p> <p>4.4 een goede woordkeuze laten zien</p> <p>4.5 variatie in woordgebruik aanbrengen</p> <p>4.6 verwijswwoorden goed en duidelijk gebruiken</p> <p>4.7 verbanden tussen / binnen zinnen expliciet maken</p>
<p>5. Orthografische vaardigheden</p> <p>De zinnen in geschreven vorm op papier zetten conform de regels en conventies van de Nederlandse taal.</p>	<p>5.1 correcte spelling hanteren</p> <p>5.2 leestekens correct hanteren</p>

2.1 Aanzet tot een taakmodel voor schrijfvaardigheid Nederlands

In Tabel 2-1 is een leerlingmodel gepresenteerd dat een overzicht geeft van de te meten attributen en deelattributen. Dit vormt het uitgangspunt voor het taakmodel, dat de sleutelkenmerken beschrijft van taken en taaksituaties van de DTT. Op grond van dit leerlingmodel zijn prototypische voorbeeldopgaven ontwikkeld, die in moeilijkheidsgraad voor de verschillende niveaus kunnen worden aangepast.

Voor de opbouw van de diagnostische toets Schrijfvaardigheid kan de volgende getrapte toetsopzet worden gehanteerd, die leidt tot toetsen die in principe voor elke leerling passend zijn.

1. Een (geleide) schrijfopdracht, afgestemd op het niveau van de leerling, die zo mogelijk door middel van een geautomatiseerd tekstanalysesysteem wordt beoordeeld. Deze opdracht kan aanwijzingen geven over het niveau van onderliggende vaardigheden als spelling en grammatica, woordenschat en zinsbouwvaardigheid. Voor niveau vmbo-b kan de opdracht bijvoorbeeld de tekst van een persoonlijke brief of e-mail betreffen, of een verslag van een persoonlijke gebeurtenis. Denk voor havo/vwo bijvoorbeeld aan een formele brief die bedoeld is om de lezers te overtuigen.
2. Een aantal deeltoetsen waarin op aan de leerling aangepaste wijze verschillende deelattributen aan de orde komen die niet of in onvoldoende mate in beeld zijn gebracht met behulp van de schrijfopdracht. Hierin wordt met gesloten vragen en meerkeuzevragen meer informatie verzameld over deze hoofd- en deelattributen. Indien mogelijk bepaalt het globale resultaat van de schrijfopdracht het instapniveau voor de adaptieve deeltoets. Het eerste deel levert dus algemene informatie over het hoofdattribuut en over een gedeelte van de deelattributen. De deeltoetsen leidt tot diagnostische informatie over de deelattributen.

Voor een aantal deelattributen zijn interactieve voorbeeldopgaven gemaakt. In deze opgaven kunnen de leerlingen hun antwoord ingeven met de muis, waarna ze direct het effect ervan zien op de tekst die ze bewerken. Bij de constructie van deze voorbeeldopgaven is nadrukkelijk de hele taalvaardigheidsrange vanaf vmbo-b tot en met vwo in de beschouwing betrokken. Dit vraagt mogelijk om specifieke uitgangsteksten voor de verschillende niveaus.

Figuur 2-1 bevat een prototypische voorbeeldopgave, die is gekoppeld aan het attribuut Tekstuele vaardigheden en aan deelattribuut 3.2 Logische volgorde van inhoudelijke elementen bepalen.

Je hebt via internet een iPod gekocht. Toen de iPod bezorgd werd, bleek hij niet te werken. Je hebt de winkel gebeld en kreeg te horen dat je het garantiebewijs per post op moest sturen.

Maak het bericht hieronder af door de onderdelen in de juiste volgorde in de brief te zetten.

Opnieuw

Aan: Hoppa Electronica

Zomerdijk 10
1253 IP Batenburg

Batenburg, 25 mei 2012,

Geachte heer, mevrouw

Van: Harold Headbanger

In plaats daarvan stuur ik u een kopie. Ik vertrouw erop dat hiermee de iPod onder de volledige garantie valt.

Met vriendelijke groet,
Harold

Gisteren hebben we elkaar aan de telefoon gesproken. Ik heb beloofd het garantiebewijs van de iPod op te sturen.

Na overleg met mijn vader lijkt het me verstandiger om niet het echte garantiebewijs op te sturen; als deze brief kwijt zou raken, heb ik dan namelijk geen enkel bewijs meer.

Figuur 2-1: Voorbeeldopgave DTT Nederlands Schrijfvaardigheid

2.2 Naar een rapportagemodel schrijfvaardigheid Nederlands

Vanwege het diagnostische doel van de DTT is de rapportage van de resultaten een essentieel aspect. Het rapportagemodel geeft aan hoe de toetsuitkomsten beschreven en gevisualiseerd kunnen worden, zodat de docent de informatie uit de DTT kan gebruiken om richting en vorm te geven aan het onderwijs.

Het rapportagemodel Schrijfvaardigheid Nederlands kan voorzien in rapportages op verschillende niveaus: het individuele niveau (voor leerling en docent), het klassenniveau (voor de docent) het schoolniveau (voor docent en schoolleider) en het landelijke niveau (voor schoolleider en beleidsmaker). Deze rapportages kennen een verschillende mate van detaillering.

Elk van deze rapportages geeft allereerst informatie over de prestaties van de leerling(en) op de inhoudelijke domeinen. Op basis van de resultaten zal per domein worden gekeken of een leerling onder, op of boven het niveau presteert dat verwacht mag worden van een leerling van het betreffende schooltype. Daarnaast geven de rapportages ook meer gedetailleerde diagnostische informatie in termen van de deelattributen die corresponderen met de concept-tussendoelen. De definitieve vorm van de rapportage wordt in overleg met het veld, onder regie van CvE, nader ingevuld.

Figuur 2-2 geeft een indruk van een mogelijke rapportage op leerlingniveau voor havo/vwo. Het bovenste deel bevat de globale niveauaanduiding; in dit geval presteert de leerling op havoniveau. Linksonder in de figuur staan deze prestaties uitgesplitst per attribuut; de zwarte punt geeft de relatieve prestatie van de leerling weer ten opzichte van de erboven genoemde schooltypen. Rechts onder staat de diagnostische informatie met betrekking tot de deelattributen.



Figuur 2-2: Voorbeeldrapportage op leerlingniveau van toetsuitkomsten Schrijfvaardigheid Nederlands voor havo/vwo

3 Naar een diagnostische tussentijdse toets voor Engels

In dit hoofdstuk wordt besproken hoe de DTT voor Engels er op basis van de uitgangspunten van hoofdstuk 1 uit zou kunnen zien. Aan de orde komen achtereenvolgens een leerlingmodel, een aanzet tot een taakmodel en een voorbeeld van de mogelijke rapportage, die vanzelfsprekend nader uitgewerkt dienen te worden en vooralsnog een voorlopig karakter hebben. Tijdens de voorstudie zijn ook twee veldraadplegingen verricht, waarbij de ideeën zijn voorgelegd aan docenten. Hier zijn zinvolle opmerkingen gemaakt die mee worden genomen in het vervolg. Over het algemeen was men enthousiast over de uitwerkingen.

Net als bij het leren van Nederlands zijn bij het leren van Engels verschillende vaardigheden van belang. Dit zijn Spreken, Gesprekken voeren, Schrijven, Lezen en Luisteren. De eerste twee zijn productief: de leerling produceert taal. De laatste twee zijn receptief, omdat de leerling taal ontvangt. Zowel voor de productieve als de receptieve vaardigheden zal er een adaptieve toets ontwikkeld worden. In de voorstudie is er echter voor gekozen één domein grondig te analyseren. Om vergelijkbare redenen als bij de DTT Nederlands (zie hoofdstuk 2) is de keuze gevallen op schrijfvaardigheid. Schrijfvaardigheid is immers wat onderbelicht in het voortgezet onderwijs en kan volgens het veld een impuls gebruiken. De vaardigheid schrijven ondervindt ook een opleving onder invloed van *social media* als Facebook, MSN en andere chat-omgevingen.

De voorstudie heeft geresulteerd in een uitgewerkt leerlingmodel voor schrijfvaardigheid Engels, dat kan worden verbreed naar leerlingmodellen voor lees-, luister- en spreekvaardigheid. Bij de ontwikkeling van het leerlingmodel voor Engels is veelvuldig overleg geweest met de sectie Nederlands binnen Cito; dit om de taalmodellen in de basis op elkaar te laten aansluiten.

3.1 Een leerlingmodel voor schrijfvaardigheid Engels

Een leerlingmodel beschrijft de aard van de vaardigheid die getoetst wordt in de vorm van een samenhangende structuur van hoofd- en deelattributen. Bij de opzet van het voorlopige leerlingmodel schrijfvaardigheid Engels zijn de attributen geïnspireerd op het model van Deane en zijn de deelattributen bepaald vanuit de concept-tussendoelen en onderzoek aangaande het Europees Referentiekader (ERK)¹.

Attributen op basis van het model van Deane

Het model van Deane vormt een basis voor een algemeen leerlingmodel Taal en Schrijven (Deane et al., 2008; Deane, 2011). Dit model is echter niet zonder meer toepasbaar op Engels, omdat het ontwikkeld is voor eerste taalverwerving. Hierdoor omvat het een aantal aspecten van schrijven die bij tweede taalverwerving Engels op onderbouwniveau niet aan de orde komen. Daarentegen kent tweede taalverwerving andere aspecten, zoals compenserende strategieën en woordenboekgebruik, die bij eerste taalverwerving geen rol spelen. Daarom heeft het leerlingmodel de volgende vijf attributen: Afstemming op publiek en doel, Coherentie, Woordenschat en woordgebruik, Spelling, interpunctie en grammatica en Ondersteunende vaardigheden.

Deelattributen op basis van concept-tussendoelen en het Europees Referentiekader

Ten behoeve van het identificeren van deelattributen zijn de door SLO opgestelde concept-tussendoelen, gebaseerd op het ERK, geanalyseerd op toetsbaarheid. Bij de analyse is tevens gekeken naar het onderzoek over de koppeling van het curriculum Moderne Vreemde Talen aan het Europees Referentiekader (Van Hest, Beltman & Kleintjes, 2001). Om een zo volledig mogelijk beeld te krijgen van

¹ Het ERK heeft bij Engels een directe rol gespeeld bij het formuleren van het leerlingmodel. Dit in tegenstelling tot de referentieniveaus bij Nederlands: de referentieniveaus speelden alleen een indirecte rol, via de concept- tussendoelen. Bij Engels zijn de concept-tussendoelen en, in mindere mate, het cognitieve model van Deane ook de basis voor het leerlingmodel. Maar omdat het model van Deane ontwikkeld is voor eerste taalverwerving, zijn bij Engels tevens het onderzoek over de koppeling van het curriculum Moderne Vreemde Talen aan het ERK (Van Hest, Beltman & Kleintjes, 2001) en de algemene "can do statements" van het ERK bepalend geweest voor het leerlingmodel Engels.

gebruikelijke attributen bij schrijfvaardigheid is tevens gekeken naar de algemene “can do statements” van het ERK en ook naar de attributen van Meijerink (2008). Dit heeft geleid tot een tiental deelattributen.

Het tekstkenmerk *Onderwerp* uit de concept-tussendoelen wordt opgenomen als taakkenmerk. Datzelfde geldt voor het element *Inhoud*, dat in het geval van een diagnostische schrijftoets niet wordt beoordeeld omdat er een context wordt aangereikt. Dit om ook leerlingen met minder ideeënrijkdom in staat te stellen een toetsbaar schrijfproduct op te leveren.

Het leerlingmodel met attributen en deelattributen

Gecombineerd leiden de attributen en deelattributen tot het leerlingmodel schrijfvaardigheid Engels dat is samengevat in Tabel 3-1. Omdat taalverwerving schoksgewijs verloopt, dient deze tabel niet als een sequentie of hiërarchie te worden opgevat. Van een lineaire opbouw in vaardigheden is geen sprake. Voor de toetsing van alle hoofd-, deel- en daaronder liggende attributen zal een toetsmatrijs worden uitgewerkt.

Tabel 3-1: Een leerlingmodel Schrijfvaardigheid Engels

Hoofdattribuut	Deelattribuut	
1 Afstemming op publiek en doel	1.1 Kan conventies behorend bij een tekstsoort gebruiken. 1.2 Kan toonzetting en register afstemmen op publiek en schrijfdiel.	5 Ondersteunende vaardigheden 5.1 kan het woordenboek gebruiken 5.2 kan compenserende strategieën toepassen
2 Coherentie	2.1 Kan bij de taak passende structuurwoorden gebruiken: gebruik van voegwoorden en verwijzwoorden. 2.2 Kan tekststructuur en verbanden aanbrengen.	
3 Woordenschat en woordgebruik	3.1 Kan bij de taak passende woorden en woordcombinaties gebruiken. 3.2 Kan het woordgebruik functioneel variëren (inclusief als compenserende strategie).	
4 Spelling, interpunctie en grammatica	4.1 Kan passende spelling en interpunctie hanteren. 4.2 Kan woordvolgorde en zinsconstructie functioneel hanteren.	

3.2 Aanzet tot een taakmodel voor schrijfvaardigheid Engels

Tabel 3-1 geeft een overzicht van de te meten attributen en deelattributen die samen het leerlingmodel vormen. Dit model is het uitgangspunt voor het taakmodel, dat de sleutelkenmerken beschrijft van taken en taaksituaties van de DTT. Op grond hiervan zijn prototypische voorbeeldopgaven ontwikkeld. Bij de ontwikkeling hiervan is rekening gehouden met de verschillende niveaus. Bij elke prototypische opgave is het beoogde niveau (vmbo-b, vmbo-k, vmbo-gt, havo en vwo) aangegeven; de moeilijkheidsgraad kan voor de verschillende niveaus worden aangepast. Verder is het van belang dat bij de ontwikkeling van items en schrijfoopdrachten aspecten als taakkenmerken, stimuli, taakstructurering en responsekenmerken worden uitgewerkt en beschreven.

Een diagnostische toets voor een productieve vaardigheid zoals schrijfvaardigheid vereist natuurlijk enige mate van productie. Daarom is gekozen voor een getrapte toetsopzet, die leidt tot toetsen die in principe voor elke leerling verschillend zijn. De DTT schrijfvaardigheid Engels kan dan bestaan uit twee onderdelen:

1. Een (geleide) schrijfoopdracht, afgestemd op het niveau van de leerling, die door middel van een digitaal tekstanalysesysteem automatisch wordt beoordeeld. Deze opdracht geeft aanwijzingen over het niveau van onderliggende vaardigheden als spelling en grammatica, woordenschat, veelvoorkomende fouten. Voor niveau vmbo-b kan dit een kaartje, formulier of chat zijn, voor havo/vwo bijvoorbeeld een (semi-zakelijke) brief. Opdracht 1 levert informatie op over de diverse attributen.

- Op basis van de resultaten van opdracht 1 volgt een adaptieve toets die de hoofd- en deelattributen verder in kaart brengt. Het instapniveau daarvan is op basis van opdracht 1 bepaald. Opdracht 2 levert aanvullende diagnostische informatie op de deelattributen.

Voor zowel het eerste deel, de schrijfoopdracht, als het tweede deel zijn prototypische voorbeeldopgaven ontwikkeld. Figuur 3-1 bevat een dergelijke opgave waarin de leerling uit een pull down menu een bijwoord of bijvoeglijk naamwoord kiest en dit vervolgens via drag and drop naar de juiste plaats in de zin sleept. Het item toetst attribuut 3, Woordenschat en woordgebruik, en deelattribuut 3.1: Kan bij de taak passende woorden en woordcombinaties gebruiken. Deze opgave heeft niveau vmbo-kb/gt, maar is eenvoudig aan te passen voor andere niveaus.

Choose the right name.

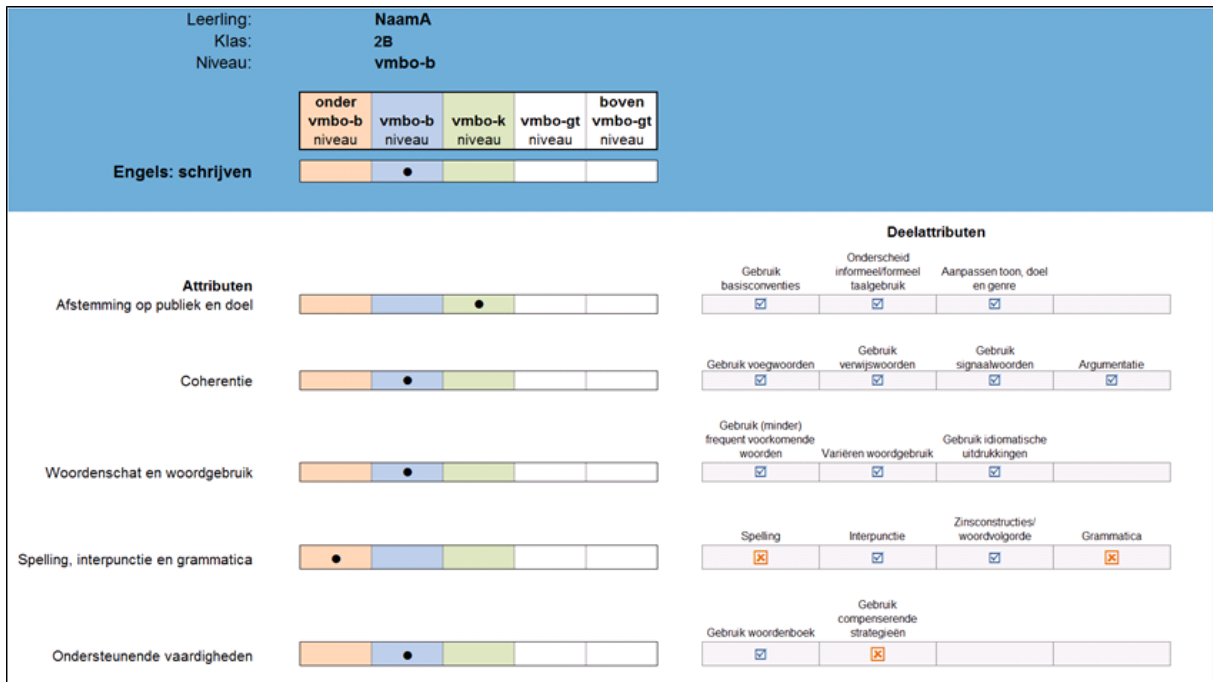
De leerling kiest uit een keuzemenu de juiste benaming.

Figuur 3-1: Voorbeeldopgave DTT Engels Schrijfvaardigheid

3.3 Naar een rapportagemodel schrijfvaardigheid Engels

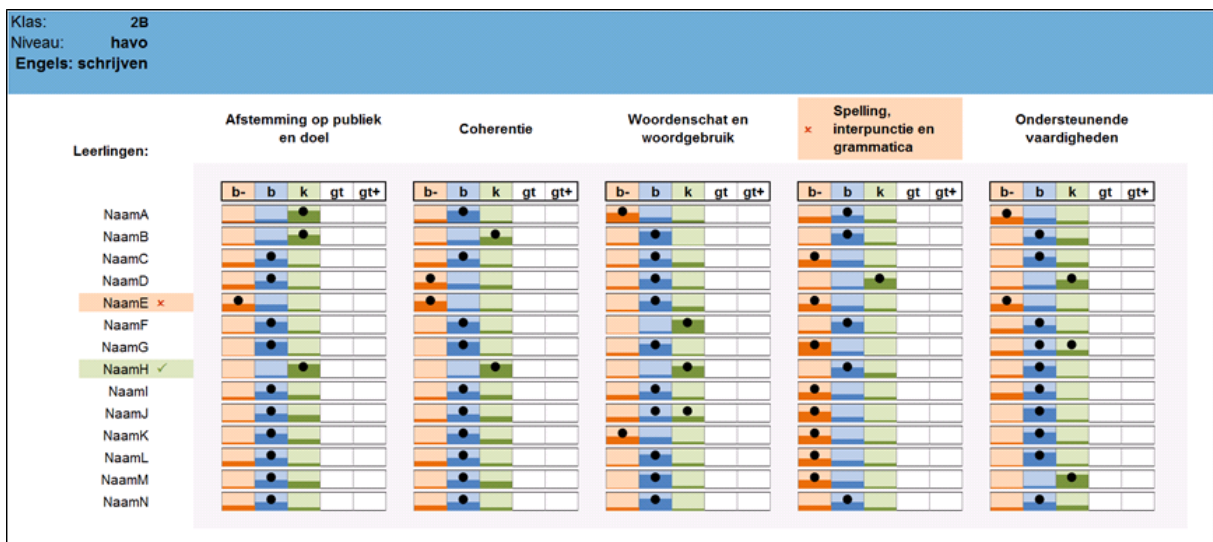
Net als bij Nederlands is bij de DTT Engels de rapportage essentieel voor het realiseren van het diagnostische doel. Ook het rapportagemodel Schrijfvaardigheid Engels kan voorzien in rapportages op verschillende niveaus: het individuele niveau (voor leerling en docent), het groepsniveau (voor de docent) het schoolniveau (voor docent en schoolleider) en het landelijke niveau (voor schoolleider en beleidsmaker). Deze rapportages kennen een verschillende mate van detaillering.

Elk van deze rapportages geeft allereerst informatie over de prestaties van de leerling(en) op de inhoudelijke domeinen. Op basis van de resultaten zal per domein worden gekeken of een leerling onder, op of boven het niveau presteert dat verwacht mag worden van een leerling van het betreffende schooltype. Daarnaast geven de rapportages ook meer gedetailleerde diagnostische informatie in termen van de deelattributen die corresponderen met de concept-tussendoelen. De definitieve vorm van de rapportage wordt in overleg met het veld, onder regie van CvE, nader ingevuld.



Figuur 3-2: Voorbeeldrapportage op leerlingniveau van toetsuitkomsten schrijfvaardigheid Engels voor vmbo

Voor Schrijfvaardigheid Engels zal gelijksoortige rapportage beschikbaar komen als voor Nederlands (zie Figuur 2-2). Figuur 3-2 en Figuur 3-3 geven een indruk van een mogelijke rapportage op leerlingniveau en op groepsniveau. Deze vormen van rapportage informeren de docent en de schoolleider over het niveau van klas ten opzichte van vergelijkbare klassen.



Figuur 3-3: Voorbeeldrapportage op groepsniveau van toetsuitkomsten Schrijfvaardigheid Engels voor vmbo

4 Naar een diagnostische tussentijdse toets voor wiskunde

In dit hoofdstuk wordt beschreven hoe de DTT wiskunde er op basis van de uitgangspunten van hoofdstuk 1 uit zou kunnen zien. Aan de orde komen achtereenvolgens een leerlingmodel, een aanzet tot een taakmodel en een voorbeeld van de mogelijke rapportage, die nog nader uitgewerkt worden en vooralsnog een voorlopig karakter hebben. Tijdens de voorstudie zijn ook twee veldraadplegingen verricht, waarbij de ideeën zijn voorgelegd aan docenten. Hier zijn zinvolle opmerkingen gemaakt die mee worden genomen in het vervolg. Over het algemeen was men enthousiast over de uitwerkingen.

In de ministeriële stukken is sprake van “wiskunde, alsmede rekenen” maar op welke manier het “alsmede” gestalte krijgt en hoe zich dit verhoudt tot de referentietoetsen rekenen zal nog moeten worden bepaald. In dit hoofdstuk is daarom gekozen om de aandacht vooralsnog op wiskunde te richten.

4.1 Een leerlingmodel voor de DTT wiskunde

Een leerlingmodel beschrijft de aard van de vaardigheid die getoetst wordt in de vorm van een samenhangende structuur van attributen. Net als bij Nederlands en Engels is een aantal attributen in het leerlingmodel voor de DTT wiskunde gebaseerd op de vakinhoudelijke domeinen zoals geformuleerd in de concept-tussendoelen onderbouw (SLO, 2012). Anders dan bij Nederlands en Engels is er bij de DTT wiskunde niet voor gekozen om vakinhoudelijke attributen uit te werken in deelattributen, maar om in plaats daarvan wiskundig-didactische aspecten in het leerlingmodel op te nemen die bijdragen aan de diagnose van moeilijkheden van leerlingen.

Attributen op basis van vakinhoud

SLO heeft vakinhoudelijke concept-tussendoelen voor vmbo en havo/vwo geformuleerd, onder andere op basis van het werk van de vernieuwingscommissie wiskunde cTWO (SLO, 2012). Deze concept-tussendoelen worden beschreven in domeinen. Voor vmbo zijn de (sub)domeinen beschreven voor verschillende niveaus en ook voor havo en vwo is een uitsplitsing gemaakt.

Het vmbo kent de domeinen Getallen, Verhoudingen, Meten en meetkunde en Verbanden en formules. Voor havo en vwo zijn de domeinen Getallen en variabelen, Verhoudingen, Meten en meetkunde en Verbanden en formules en Informatieverwerking en onzekerheid. Daarnaast kennen zowel vmbo als havo/vwo het domein Inzicht en handelen. Omdat dit domein overkoepelende wiskundige vaardigheden (SLO, 2012) betreft die niet apart getoetst worden, wordt er in de DTT niet concreet over gerapporteerd.

Deze domeinen van de concept-tussendoelen vormen de attributen van het leerlingmodel. De diagnostische tussentijdse toets zal immers docent en leerling informatie moeten geven over de mate waarin leerlingen deze concept-tussendoelen beheersen.

Attributen op basis van wiskundig-didactische aspecten

Wiskunde leren is een complex proces voor veel leerlingen. Op het eerste gezicht onbegrijpelijke fouten blijken vaak een diepere oorzaak te hebben, die soms domeinoverstijgend is. Om de diagnostische waarde van de DTT te vergroten, is een drietal wiskundig-didactische aspecten geïdentificeerd die licht kunnen werpen op deze complexiteit en op onderliggende moeilijkheden van leerlingen. Daarmee biedt het resultaat van de DTT aanknopingspunten voor diagnose en remediëring. Uit vakdidactische literatuur (zie bijvoorbeeld Drijvers, Van Streun, & Zwaneveld, 2012) zijn de volgende aspecten gekozen als wiskundig-didactische attributen: wiskundige structuur, ambiguïteit en samenhang.

Centraal in de wiskunde staat het zien en ontdekken van *wiskundige structuur*. Het kan bijvoorbeeld gaan om de structuur van formules, de structuur van een figuur of de structuur in een verzameling meetgegevens. Het zien en ontdekken van structuur speelt binnen alle wiskundedomeinen een rol. In het vmbo wordt leerlingen bijvoorbeeld gevraagd om $20 - 2 \times 8 - 3$ uit te rekenen. Daarvoor is het nodig dat ze de structuur zien van 20 min 2×8 en dan nog eens min 3 en niet van links naar rechts $18 \times 8 - 3$ lezen. Ook bij formules is het van belang om de structuur ervan te herkennen. Een vergelijkbaar voorbeeld betreft het herkennen van de structuur van een lineair verband, waarbij een leerling het verband tussen het vaste deel en het variabele deel in de bijbehorende formule doorgrondt.

Een tweede aspect dat het leren van wiskunde moeilijk maakt voor veel leerlingen is het ambigue karakter van symbolen en begrippen, die op meerdere manieren kunnen worden opgevat. Ook deze *ambigüiteit* komt voor in alle domeinen en op alle niveaus van de wiskunde. Neem bijvoorbeeld breuken zoals $\frac{3}{4}$ die zowel een deling (een proces) als het resultaat van die deling (een object) voorstellen, of expressies als $3x + 2$ die zowel een proces van vermenigvuldigen en optellen voorstellen als het resultaat daarvan (een object). Een ander voorbeeld van een ambigu symbool is het =-teken. In het basisonderwijs heeft het =-teken een operationele betekenis: het is het signaal dat de leerling wat moet gaan uitrekenen. In de regel $a + b = b + a$ gaat het echter om een relationele betekenis: er hoeft niet te worden uitgevoerd, maar er wordt aangegeven dat je bij optellen de volgorde mag verwisselen. Een laatste voorbeeld is de blikwisseling die leerlingen van havo-vwo bij het werken met statistische data nodig hebben: enerzijds moet elk datapunt als een meting op zich worden beschouwd, terwijl anderzijds bijvoorbeeld het gemiddelde een kenmerk van de dataset als geheel is.

Wiskunde is een samenhangend geheel en het overzien van deze samenhang is voor leerlingen niet vanzelfsprekend. Met *samenhang*, het derde vakdidactische attribuut, doelen we zowel op samenhang tussen domeinen als binnen domeinen. Binnen Verbanden en formules betekent samenhang bijvoorbeeld dat leerlingen de relaties zien tussen de verschillende representaties van verbanden, zoals grafiek, formule, tabel en context. Voorbeelden van samenhang tussen domeinen zijn het zien van relaties tussen breuken, verhoudingen, schalen en percentages in de domeinen Getallen en Verhoudingen; rekenregels voor machten uit het domein Getallen en variabelen en exponentiële verbanden uit het domein Verbanden en formules; en de relaties tussen gelijkvormigheid uit het domein Meten en meetkunde en het domein Verhoudingen.

Vakinhoudelijke en vakdidactische attributen gecombineerd

De combinatie van de twee typen attributen levert een leerlingmodel op waarin vijf vakinhoudelijke attributen worden onderscheiden, die elk corresponderen met een domein uit de concept-tussendoelen. Als wiskundig-didactische attributen, die deze domeinen overstijgen, zijn drie aspecten onderscheiden: het zien van de wiskundige structuur, het omgaan met ambigüiteit en het zien van samenhang. Tabel 4-1 laat een kruising zien van deze twee typen attributen en visualiseert daarmee een leerlingmodel voor de DTT wiskunde.

Tabel 4-1: Attributen in een leerlingmodel voor de DTT wiskunde

Vakinhoudelijke attributen: domeinen	Wiskundig-didactische aspecten:		
	Wiskundige structuur	Ambigüiteit	Samenhang
Getallen (en variabelen, alleen voor havo/vwo)			
Verhoudingen			
Meten en meetkunde			
Verbanden en formules			
Informatieverwerking en onzekerheid (alleen voor havo/vwo)			

4.2 Aanzet tot een taakmodel voor de DTT wiskunde

Een taakmodel beschrijft de sleutelkenmerken van taken en taaksituaties van de DTT en vormt daarmee de basis voor de ontwikkeling van opgaven. De volgende bouwstenen vormen een aanzet voor het te ontwikkelen taakmodel.

Ten eerste is het net als bij andere vakken voor de DTT wiskunde van belang dat de opgaven in de toets de in de concept-tussendoelen beschreven domeinen (attributen) zorgvuldig afdekken. Zoals eerder opgemerkt zal domein, Inzicht en handelen, niet apart getoetst worden omdat de betreffende vaardigheden besloten zitten in de overige domeinen.

Ten tweede zullen de opgaven alle niveaus, variërend van vmbo basisberoepsgerichte leerweg tot en met vwo, moeten afdekken. Dit betekent dat er een verscheidenheid aan opgaven geconstrueerd zal moeten worden waarin ook aandacht is voor items die op de grens van deze niveaus liggen. Omdat de DTT gericht is op klas 2 van vmbo en klas 3 van havo/vwo, is een niveauverschil tussen vmbo en havo/vwo vanzelfsprekend.

Ten derde zullen de items zo veel als mogelijk gekoppeld dienen te worden aan de attributen structuur, ambiguïteit en samenhang. Op basis van een analyse van de gegevens uit een pretest zal de relatie tussen deze potentiële knelpunten en items nader worden uitgewerkt. De adaptiviteit van de toets zal er onder meer uit bestaan dat zowel domeinspecifieke als overstijgende moeilijkheden van leerlingen nader worden gediagnostiseerd.

Ten vierde, zullen contexten in items van de diagnostische tussentijdse toets wiskunde een vergelijkbare rol spelen als gebruikelijk is in de schoolboeken. Voor het vmbo betekent dit dat er veelal zal worden gestart met probleemsituaties waarin leerlingen vanuit de context kunnen redeneren en informele strategieën kunnen gebruiken. Voor havo/vwo zal eerder worden overgegaan naar meer wiskundige situaties.

Bij het beoordelen van wiskunde, tenslotte, is niet alleen het antwoord van belang, maar ook hoe leerlingen tot het antwoord gekomen zijn (de uitwerking). De digitale afname biedt de mogelijkheid dat er gebruik wordt gemaakt van software die in staat is om de uitwerkingen van de leerlingen op te slaan en te analyseren. Het is de verwachting dat een analyse van dergelijke uitwerkingen kan helpen bij het diagnosticeren van fouten, gekoppeld aan de drie potentiële knelpunten (wiskundige) structuur, ambiguïteit en samenhang.

Figuur 4-1 bevat een prototypische voorbeeldopgave voor havo/vwo, die is gekoppeld aan het domein Meten en meetkunde en aan het attribuut Ambiguïteit, omdat dat in dit item naar verwachting het grootste knelpunt is.

	er zijn altijd evenwijdige zijden	de diagonalen snijden elkaar altijd loodrecht	alle zijden zijn altijd even lang	alle hoeken zijn altijd even groot
parallelogram	ja	nee	nee	nee
vlieger	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
ruit	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
gelijkzijdige driehoek	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
vierkant	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Opnieuw

Niveau: vmbo-t/havo/vwo

Potentieel vakdidactisch aspect: ambiguïteit.

Toelichting: In deze opgave gaat het erom dat leerlingen eigenschappen van hoeken en lijnen kunnen koppelen aan een meetkundig object.

Figuur 4-1: Voorbeeldopgave DTT wiskunde domein Meten en meetkunde voor vmbo-t/havo/vwo

4.3 Naar een rapportagemodel voor de DTT wiskunde

Vanwege het diagnostische doel van de DTT is de rapportage van de resultaten een wezenlijk aspect. Het rapportagemodel geeft aan hoe de toetsuitkomsten beschreven en gevisualiseerd kunnen worden, zodat de docent de informatie uit de DTT kan gebruiken om richting en vorm te geven aan zijn/haar onderwijs.

5 Ontwikkeling en implementatie in praktijk

In dit hoofdstuk wordt nader ingegaan op de ontwikkeling en implementatie van de DTT vanuit het perspectief van Cito. De ontwikkeling en implementatie van de DTT zal onder verantwoordelijkheid van het CvE verder vormkrijgen en is een groeiproces. Enerzijds is de DTT een innovatief product, waardoor de ontwikkeling gepaard zal moeten gaan met onderzoeksactiviteiten die gemaakte keuzes en uitwerkingen evalueren en waar nodig bijsturen. Anderzijds betreft het een grootschalig product waarbij voor veel domeinen een adaptieve toets wordt gemaakt. Voor adaptieve toetsen zijn omvangrijke opgavenbanken nodig, evenals grote aantallen observaties voor de onderbouwing van de feedback en de diagnose. Een en ander maakt dat een doorlooptijd van enkele jaren nodig zal zijn, waarin opgavenbanken stapsgewijs worden gevuld en landelijke implementatie gefaseerd zal verlopen.

5.1 Ontwikkeling en uitwerking

Vrijwel alle onderwerpen die in dit document zijn aangestipt vragen om verdere ontwikkeling, uitwerking en onderbouwing voor de realisatie van de DTT. Zonder in detail te treden of te streven naar volledigheid noemen we een aantal wezenlijke aandachtspunten vanuit Cito.

Opgave- en itemontwikkeling

Een eerste aandachtspunt betreft de opgave- en itemontwikkeling. Nieuwe itemtypen dienen te worden gespecificeerd evenals de beoordelingsmodellen die zich lenen voor automatische scoring. Vanzelfsprekend wordt gedacht aan items waarop goed/fout wordt gescoord, maar er zijn ook andere vormen mogelijk, zoals de voorbeelden in de eerdere hoofdstukken laten zien. Een en ander dient te worden afgestemd met de mogelijkheden van de digitale toetsomgeving FACET (CvE, 2012). Als itemspecificaties beschikbaar zijn, kan de ontwikkeling van toetsitems een aanvang nemen.

Psychometrische aspecten

Een tweede aandachtspunt betreft de psychometrische aspecten van de DTT. De toetsen moeten vooraf worden gekalibreerd om diagnostische uitspraken, zoals “de leerling beheerst het attribuut boven het vereiste niveau”, mogelijk te maken. Voor elke leerling kunnen diagnostische hypothesen (zie Figuur 5-1) geëvalueerd worden met behulp van Bayesiaanse statistische methoden. Bovendien dienen de betrouwbaarheid en validiteit van de diagnoses en onderliggende leerlingmodellen te worden onderzocht.

- | |
|--|
| <p>H+: de leerling beheerst het attribuut boven het vereiste niveau (bijvoorbeeld <i>boven</i> vmbo-gt niveau)</p> <p>H=: de leerling beheerst het attribuut op het vereiste niveau (bijvoorbeeld <i>op</i> vmbo-gt niveau)</p> <p>H-: de leerling beheerst het attribuut onder het vereiste niveau (bijvoorbeeld <i>onder</i> vmbo-gt niveau)</p> |
|--|

Figuur 5-1: De relevante diagnostische hypothesen

Assemblage en adaptiviteit

Als derde aandachtspunt noemen we de assemblage, en met name de adaptiviteit daarbinnen. Bij de samenstelling van de toets moet worden nagegaan hoeveel items nodig zijn om de leerlingmodellen te evalueren en tot welke afnametijd dit leidt. Tevens is de vraag op welke manier de adaptiviteit binnen de toets gestalte krijgt. De getrapte toetsvorm die in de hoofdstukken 2 en 3 aan de orde is gekomen, biedt bijvoorbeeld mogelijkheden voor adaptiviteit, omdat (deel)attributen die de leerling in het eerste deel van de toets beheerst niet nader gediagnostiseerd hoeven te worden. Andere mogelijkheden voor adaptiviteit doen zich voor als bepaalde (deel)attributen hiërarchisch geordend kunnen worden. Stel bijvoorbeeld dat het bij Schrijfvaardigheid zo is dat leerlingen die niet coherent kunnen schrijven, ook niet in staat zijn af te stemmen op publiek en doel. Dit zou betekenen dat het attribuut Afstemming op publiek en doel niet meer geëvalueerd hoeft te worden als het attribuut Coherentie niet beheerst wordt, wat zou leiden tot een verdere reductie van het aantal af te nemen items voor de leerlingen die Coherentie niet beheersen.

Rapportagemodel

Als vierde en laatste aandachtspunt noemen we de verdere uitwerking van het rapportagemodel. In de literatuur wordt opvallend weinig aandacht besteed aan het op een juiste manier interpreteren van rapportages door gebruikers. Het *EvidenceCentered Design*model (Mislevy et al., 2002, zie Figuur 1-1) bevat geen rapportagemodel. De ondersteuning van een correcte interpretatie door middel van heldere rapportages is echter wel een essentieel aspect van de DTT, omdat de toetsresultaten belangrijke informatie moet opleveren. Voor de diagnostische toets betekent dit concreet dat de rapportage het beoogde formatieve toetsgebruik moet ondersteunen. Voor de rapportage is een aantal principes en criteria geïdentificeerd, zoals het geven van visuele informatie, het ordenen van informatie, zinvolle groepering van data, en doelgericht kleurgebruik. Uitgaande van deze principes en criteria zal het rapportagemodel worden gespecificeerd. De definitieve vorm van de rapportage wordt in overleg met het veld, onder regie van CvE, nader ingevuld.

5.2 Tijdpad naar implementatie

Bij de ontwikkeling en implementatie van de DTT is sprake van een groeiproces waarin de volgende fasen worden onderscheiden.

1. Voorstudie
Het doel van de voorstudie is om het concept DTT nader uit te werken in een exemplarisch toetsontwerp en zo te beproeven op zijn haalbaarheid. Dit document is een samenvatting van de resultaten van de voorstudies tot nu toe. Voor de andere vaardigheden bij de talen en voor rekenen zal deze stap nog plaatsvinden.
2. Try-out
Het doel van de try-out is de verdere uitwerking van het toetsontwerp en de kwaliteitsborging van de toets middels onderzoek. Hiertoe is een kleinschalige afname van het ontwikkelde toetsmateriaal in een digitale omgeving noodzakelijk, omdat sommige items uitsluitend digitaal afgenomen kunnen worden. Ook maakt dit een realistische inschatting mogelijk van de antwoortijd voor elk itemtype via registratie van responstijden.
3. Grootschalige itemontwikkeling
De adaptiviteit en de dekking van alle attributen en deelattributen vragen om omvangrijke itemdatabanken. Op basis van de bevindingen van de eerdere fasen kan grootschalige itemontwikkeling worden gerealiseerd.
4. Pretest
Het doel van de pretest is het grootschalig pretesten van de gehele toetsinhoud. Op grond van de resultaten hiervan kan de psychometrische kalibratie plaats vinden en kunnen betrouwbaarheid en validiteit onderzocht worden.
5. Eerste afname
De eerste afname betreft de feitelijke invoering van de DTT in de beoogde opzet en infrastructuur. De eerste afname zal geëvalueerd worden.
6. Operationele fase
De operationele fase bestaat uit een jaarlijkse cyclus van eventuele bijstelling van de toetsprocedure, itemverversing, toetsafname, analyse en rapportage. Afhankelijk van de gekozen methode zal de kalibratie jaarlijks uitgevoerd moeten worden voor nieuw itemmateriaal.

Op basis van deze fasering is het mogelijk een tentatief tijdpad op te stellen. De invoering van de DTT voor de vaardigheden uit deze voorstudie wordt in het concept wetsvoorstel voorzien in het schooljaar 2015-2016. De doorlooptijd per domein bedraagt namelijk ongeveer 4,5 jaar vanaf de start van de voorstudie tot en met de eerste afname. Om een extreme belasting voor scholen, toetsconstructeurs, Toetswijzercommissies en Vaststellingscommissies te vermijden is een stapsgewijze invoering van vaardigheden per vak het meest realistisch en efficiënt. De komende tijd zal duidelijk worden hoe dit proces zal gaan verlopen voor de overige vaardigheden. De door het CvE ingestelde toetswijzercommissies en het verdere verloop van het wetgevingstraject zullen hier duidelijkheid in scheppen.

Literatuur

- Braaksmā, M., Rijlaarsdam, G., van den Bergh, H., & van Hout-Wolters, B. (2007). Observerend leren en de effecten op de organisatie van schrijfprocessen. *LevendeTalenTijdschrift*, 8(4), 3-15.
- Clark, H. & Clark, E. (1977). *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- College voor Examens (2012). *Invoeringsplan CES, versie 06*. Intern rapport. Utrecht: CvE.
- De Bruyn, E. E., J., Ruijssenaars, A. J. J. M., Pameijer, N. K., & Van Aarle, E. J. M. (2003). *De diagnostische cyclus. Een praktijkleer*. Leuven, België: Acco.
- Deane, P. (2011). *Writing Assessment and Cognition*. Research Report ETS RR-11-14. Princeton: Educational Testing Service.
- Deane, P., Odendahl, O., Quinlan, Th., Fowles, M., Welsh, C., Bivens-Tatum, J. (2008). *Cognitive Models of Writing: Writing Proficiency as a Complex Integrated Skill*. Research Report ETS RR-08-55. Princeton: Educational Testing Service.
- Drijvers, P., Van Streun, A., & Zwaneveld, B. (Red.) (2012). *Handboek wiskundendidactiek*. Utrecht: Epsilon Uitgaven.
- Meijerink, H. (2008). *Over de drempels met taal en rekenen : hoofdrapport van de Expertgroep doorlopende leerlijnen taal en rekenen*. Enschede: Expertgroep Doorlopende Leerlijnen Taal en Rekenen.
- Ministerie van OCW (2012). *Concept wetsvoorstel Leerlingvolgsysteem en diagnostische tussentijdse toets VO - versie voor internetconsultatie, maart 2012*. Opgeroepen in maart 2012 van <http://www.internetconsultatie.nl/tussentijdsetoetsvo>.
- Mislevy, R. J., Steinberg, L. S., Breyer, F. J., Almond, R. G., & Johnson, L. (2002). Making sense of data from complex assessments. *Applied Measurement in Education*, 15, 363-378.
- Mislevy, R. J., Steinberg, L. S., Almond, R. G., & Lukas, J. F. (2006). Concepts, terminology and basic models of evidence-centered design. In D. M. Williamson, R. J. Mislevy, & I. I. Bejar (Eds.), *Automated Scoring of Complex Tasks in Computer-Based Testing* (pp. 15-47). Mahwah NJ: Lawrence Erlbaum Associates.
- Pinker, S. (1996). *Language Learnability and Language Development*. Cambridge MA: MIT University Press.
- Sanders, P. (2011). Het doel van toetsen. In P. Sanders (Ed.), *Toetsen op school* (pp. 9-20). Arnhem: Cito.
- Schilperoord, J. (1996). *It's about time. Temporal aspects of cognitive processes in text production*. Dissertation. Amsterdam/Philadelphia: Rodopi.
- SLO (2012). *Concept-advies tussendoelen kernvakken onderbouw vo*. Enschede: SLO.
- Van Hest, E., Beltman E.H., & Kleintjes, F. (2001). *Koppeling mvt-examens aan het Europese Referentiekader*. Arnhem: Citogroep, unit Voortgezet Onderwijs.

Cito maakt wereldwijd werk van goed en eerlijk toetsen en beoordelen. Met de meet- en volgmethoden van Cito krijgen mensen een objectief beeld van kennis, vaardigheden en competenties.

Hierdoor zijn verantwoorde keuzes op het gebied van persoonlijke en professionele ontwikkeling mogelijk. Onze expertise zetten we niet alleen in voor ons eigen werk maar ook om advies, ondersteuning en onderzoek te bieden aan anderen.

Cito

Amsterdamseweg 13
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11
F (026) 352 13 56
www.cito.nl

Klantenservice

T (026) 352 11 11
F (026) 352 11 35
klantenservice@cito.nl

Fotografie: Ron Steemers