

Toetsen op School Voortgezet onderwijs

Hans Kuhlemeier en Alma van Til
Piet Sanders (redactie)



zeker weten

Toetsen op School Voortgezet onderwijs

Hans Kuhlemeier en Alma van Til

Piet Sanders (redactie)

Toetsen op School voortgezet onderwijs is een uitgave van Cito

- Redactie: Piet Sanders
- Auteurs: Hans Kuhlemeier en Alma van Til
- Opmaak: Service unit, MMS
- Foto omslag: Ron Steemers

© Stichting Cito Instituut voor Toetsontwikkeling Arnhem (2016)

Alle rechten voorbehouden. Niets uit dit werk mag zonder voorafgaande schriftelijke toestemming van Stichting Cito Instituut voor Toetsontwikkeling worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotokopie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook. Stichting Cito Instituut voor Toetsontwikkeling Arnhem heeft getracht alle rechthebbenden te achterhalen. Indien iemand meent als rechthebbende in aanmerking te komen, kan hij of zij zich tot Cito wenden.

Inhoud

	Inleiding	8
1	Het doel van toetsen	10
1.1	Toetsen voor het beoordelen van leerlingen	10
1.1.1	Toetsen om leerlingen te classificeren	10
1.1.2	Toetsen om leerlingen te certificeren	12
1.2	Toetsen voor het beoordelen van het onderwijsleerproces	19
1.2.1	Toetsen van het leren of toetsen voor het leren	19
1.2.2	Welke toetsvorm voor toetsen van het leren en toetsen voor het leren?	19
1.2.3	Waarom toetsen voor het leren?	20
1.3	Toetsen voor het beoordelen van groepen leerlingen en scholen	23
1.4	Toetsen voor het beoordelen van de kwaliteit van het Nederlandse onderwijs	24
2	De inhoud van toetsen en examens	28
2.1	De kerndoelen voor de onderbouw	28
2.2	De examenprogramma's en syllabi voor de bovenbouw	31
2.3	Referentieniveaus voor het voortgezet onderwijs	35
3	Het centraal examen en het schoolexamen	46
3.1	Het examensysteem in het voortgezet onderwijs	46
3.2	Het centraal examen	47
3.2.1	Soorten centrale examens	48
3.2.2	De voorbereiding op het centraal examen	50
3.2.3	De afname van het centraal examen	51
3.2.4	De correctie van het CSPE	52
3.2.5	De correctie van het centraal schriftelijk examen	55
3.3	Het schoolexamen	62
3.3.1	De inhoud en toetsvormen van het schoolexamen	62
3.3.2	Examenreglement, PTA en andere documenten	64
3.3.3	Kwaliteitseisen voor het schoolexamen	64
3.3.4	De kwaliteit van de schoolexamens	68
3.3.5	Het verschil tussen de cijfers op het CE en SE	70
3.3.6	Eindcijfers en einduitslag	70
3.4	De betrouwbaarheid van examens	71
3.5	Het gebruik van examens voor andere doeleinden	73
4	Toetsen met het ERK in de klas	80
4.1	De inhoud van ERK-toetsen	80
4.1.1	Vaardigheden apart of geïntegreerd toetsen	81
4.2	Toetsen met behulp van een toetsplan en een toetsmatrijs	82
4.2.1	De inhoudscategorieën van een toetsmatrijs voor het ERK	82
4.2.2	De gedragscategorieën van een toetsmatrijs voor het ERK	83
4.2.3	Een concreet voorbeeld van een toetsmatrijs	84

4.3	Het maken van opdrachten bij productieve vaardigheden	87
4.3.1	Opdrachten wel of niet in de doeltaal?	87
4.3.2	Het nadeel van volledig open opdrachten	89
4.3.3	Het uitlokken van een leerlingprestatie op een bepaald ERK-niveau	90
4.3.4	Is mijn B1-opgave jouw B1-opgave?	91
4.4	Het beoordelen van productieve vaardigheden	92
4.4.1	Welke beoordelingscriteria moeten worden toegepast?	92
4.4.2	Hoe moet gescoord worden?	93
4.5	Het maken van opdrachten voor receptieve vaardigheden	96
4.5.1	Het constructieproces	96
4.5.2	De vraagstelling	97
4.6	Het normeren van ERK-toetsen	98
4.6.1	ERK-toetsen en cijfers	99
5	Het meten van non-cognitieve onderwijsdoelen	102
5.1	Waarom aandacht besteden aan non-cognitieve doelen?	102
5.2	Sociale competentie, burgerschap en morele ontwikkeling	105
5.2.1	Sociale competentie	105
5.2.2	Burgerschap en sociale integratie	106
5.2.3	Morele ontwikkeling	109
5.3	Probleemgedrag als onderwijsdoelstelling?	112
5.4	Methoden voor het meten van non-cognitieve doelen	114
5.5	Instrumenten voor het meten van non-cognitieve doelen	118
5.6	Het zelf ontwikkelen van non-cognitieve instrumenten	124
	Geraadpleegde literatuur	130
	Index	138

Inleiding

Inleiding

In de publicatie *Toetsen op School* (2011) bespreken we in tien hoofdstukken een groot aantal aspecten van toetsen en examineren. De inhoud van *Toetsen op School* behoort volgens ons tot de basiskennis van iedere (beginnende) docent. *Toetsen op School* wordt vervolgd met aanvullende publicaties die zich richten op het toetsen en/of examineren in de verschillende onderwijssectoren. Het gaat hier om de publicaties *Toetsen op School primair onderwijs* (2011), *Toetsen op School voortgezet onderwijs* (2012) en de nog te verschijnen publicaties *Toetsen op School middelbaar beroepsonderwijs* en *Toetsen op School hoger onderwijs*.

In *Toetsen op School voortgezet onderwijs* gaan we in vijf hoofdstukken in op het toetsen en examineren in het voortgezet onderwijs. In hoofdstuk 1 bespreken we het doel van toetsen die in het voortgezet onderwijs worden afgenomen. Hierbij maken we een onderscheid tussen toetsen voor het beoordelen van leerlingen, toetsen voor het beoordelen van het onderwijsleerproces, toetsen voor het beoordelen van groepen leerlingen en scholen, en toetsen voor het beoordelen van de kwaliteit van het Nederlandse onderwijs. De inhoud van toetsen en examens is het onderwerp van het tweede hoofdstuk. Aan de orde komen de kerndoelen van de onderbouw, de examenprogramma's en syllabi voor de bovenbouw en de referentieniveaus voor het voortgezet onderwijs. In het derde hoofdstuk besteden we aandacht aan de centrale examens en schoolexamens. Dit zijn waarschijnlijk de meest belangrijke examens die in Nederland worden afgenomen. Uitgebreid aan de orde komen onder andere de afname van examens, de correctie van examens en de kwaliteit van examens. Het onderwerp van hoofdstuk 4 is hoe, uitgaande van het Europees Referentiekader (ERK), moderne vreemde talen in de klas getoetst kunnen worden. Dergelijke handreikingen zouden ook voor het toetsen van beta-, gamma- en kunstvakken goede diensten kunnen bewijzen. Hoofdstuk 5 is gewijd aan een onderwerp waar gewoonlijk in het onderwijs minder aandacht voor is, het meten van niet-cognitieve leerdoelen. Het hoofdstuk geeft de lezer een handzaam overzicht van het meten van niet-cognitieve leerdoelen.

Toetsen op School voortgezet onderwijs op internet

Deze publicatie en een ToetsSpecial met aanvullende informatie bij de publicaties van *Toetsen op School* zijn online beschikbaar. Ga voor meer informatie naar www.toetsenopschool.nl.

Piet Sanders

1 Het doel van toetsen

1 Het doel van toetsen

Alma van Til

In het voortgezet onderwijs (vo) wordt regelmatig getoetst. Maar niet elke toets wordt om dezelfde reden afgenomen. Zoals in hoofdstuk 1 van het algemene deel van Toetsen op School, zie www.toetsenopschool.nl, is aangegeven, kunnen we verschillende doelen voor toetsen onderscheiden. In dit hoofdstuk zijn de toetsen per toetsdoel gerangschikt. In paragraaf 1.1 bespreken we toetsen voor het beoordelen van leerlingen. Dit zijn toetsen die het leerproces afsluiten en waarmee kan worden bepaald wat de leeropbrengsten zijn. In paragraaf 1.2 behandelen we toetsen voor het beoordelen van het onderwijsleerproces. Deze toetsen worden tijdens of soms zelfs voorafgaand aan het leerproces afgenomen. Doel van de toetsen is het geven van informatie met behulp waarvan de docent de lessen kan aanpassen aan de behoeften van leerlingen. Deze aanpassingen kunnen betrekking hebben op één enkele leerling, maar soms ook op een hele groep. Toetsen voor het beoordelen van scholen worden beschreven in paragraaf 1.3 en toetsen voor het beoordelen van de kwaliteit van het onderwijs in paragraaf 1.4. De meeste toetsen worden overigens voor meerdere doelen gebruikt. Een voorbeeld daarvan is Toets 0 t/m 3 van het Cito Volgsysteem VO. Daarom zullen de verschillende toepassingen van de toets in meerdere paragrafen aan de orde komen.

1.1 Toetsen voor het beoordelen van leerlingen

De meeste toetsen die in het vo worden gebruikt, zijn bedoeld om leerlingen te beoordelen. Deze beoordeling kan om verschillende redenen plaatsvinden. Een eerste reden is determinatie: het toewijzen van leerlingen aan verschillende schooltypen tijdens de brugperiode. Een toets die voor dit doel veel wordt gebruikt is Toets 0 t/m 3 van het Cito Volgsysteem voortgezet onderwijs. Deze toets zullen we eerst bespreken. Vervolgens behandelen we een aantal certificerende toetsen. Bij deze toetsen worden leerlingen ook beoordeeld, maar dan met het doel om na te gaan of ze in aanmerking komen voor een getuigschrift met civiel effect.

1.1.1 Toetsen om leerlingen te classificeren

Vrijwel alle scholen in het vo kennen een brugperiode. Aan het einde van deze periode wordt bepaald welk schooltype het beste bij de leerling past. Toetsen die de school en de leerling kunnen helpen bij deze beslissing zijn Toets 0 t/m 3 van het Cito Volgsysteem voortgezet onderwijs, zie www.cito.nl > Onderwijs > Voortgezet onderwijs > Cito Volgsysteem. Dit is een instrument voor het bepalen van het niveau van leerlingen op een aantal kernvaardigheden en het volgen van dit niveau gedurende meerdere leerjaren. Door de prestaties op meerdere opeenvolgende meetmomenten te vergelijken, krijgen we inzicht in de groei, stilstand of achteruitgang van de vaardigheid van leerlingen. De rapportages van Toets 0 t/m 3 leveren informatie op waarmee de prestaties van leerlingen niet alleen vergeleken kunnen worden met leerlingen van hetzelfde schooltype, maar ook met die van leerlingen die een schooltype hoger of lager geplaatst zijn. Zo is het mogelijk om informatie te geven over welke leerweg of schooltype het beste past bij de leerlingen gezien de behaalde leerprestaties. Natuurlijk kan een school besluiten om hiervan af te wijken.

De opzet van Toets 0 t/m 3

Toets 0 t/m 3 is gericht op leerlingen in de eerste drie leerjaren van het voortgezet onderwijs. Om aan te sluiten bij de vaardigheid van de leerlingen en dus recht te doen aan niveauverschillen zijn er voor elk afnamemoment drie toetsversies beschikbaar. Eén toetsversie

voor vmbo-leerlingen op het niveau basisberoepsgerichte leerweg en leerwegondersteunend onderwijs, één voor de overige vmbo-niveaus (kaderberoepsgericht en gemengde en theoretische leerweg) en één voor havo- en vwo-niveau. Voor de eerste drie leerjaren van het vo zijn toetsen gemaakt voor vier afnamemomenten:

- Toets 0: eerste klas september/oktober
- Toets 1: eerste klas april/juli
- Toets 2: tweede klas februari/mei
- Toets 3: derde klas april/juli

We meten de volgende vaardigheden met Toets 0 t/m 3:

- Nederlands leesvaardigheid
- Nederlands woordenschat
- Taalverzorging
- Engels leesvaardigheid
- Engels woordenschat
- Rekenen/wiskunde
- Wiskunde (h/v versie Toets 2 en Toets 3)

Er is voor deze vaardigheden gekozen omdat Nederlands, Engels en wiskunde verplichte vakken zijn die worden gegeven vanaf de brugklas tot en met het eindexamen. Daarnaast zijn Nederlands leesvaardigheid, Nederlands woordenschat en rekenen/wiskunde ondersteunend voor allerlei andere schoolvakken. Om bijvoorbeeld goed te kunnen presteren bij geschiedenis is een goed begrip van Nederlandse teksten (en woorden) essentieel. Het onderdeel taalverzorging is relatief nieuw. Het toevoegen van opgaven met betrekking tot taalverzorging past bij de toegenomen aandacht voor deze vaardigheid in het vo en maakt het mogelijk om Toets 0 t/m 3 te relateren aan de referentieniveaus van de commissie Meijerink. Voor enkele demo-opgaven van Toets 0 t/m 3 verwijzen we naar [PDF 1.1](#) op de ToetsSpecial Toetsen op School, <http://www.toetswijzer.nl/berichten/toetsspecial-toetsen-op-school-voortgezet-onderwijs>

De adviesfunctie van Toets 0 t/m 3

Het hart van de rapportages van Toets 0 t/m 3 wordt gevormd door een overzicht waarin de prestaties van leerlingen worden vergeleken met die van leerlingen van hetzelfde schooltype, maar ook met die van leerlingen die een schooltype hoger of lager geplaatst zijn. Dit is goed te zien in de niveaurapportage van leerling Nico Versteegh in figuur 1.1. Nico is een leerling die in een dakpanklas zit op niveau vmbo GT/havo. Hij heeft aan het begin van de brugklas Toets 0 gemaakt. Zijn mentor heeft Nico ingeschat als een vmbo GT-leerling en dit ingevuld op het antwoordblad, nog voordat Nico aan de toets begon. Deze niveau-inschatting van de mentor is weergegeven in de achtste regel bij 'niveau' en heeft dus niets te maken met de prestaties van Nico op de toets.

Om de prestaties van Nico te kunnen vergelijken met leerlingen van verschillende schooltypes, moeten we kijken naar de percentielscores. Volgens de rapportage heeft Nico op niveau vmbo BB een percentielscore van 77 voor taalverzorging (zie de vierde kolom). Dit geeft aan dat 77 procent van de BB-leerlingen even goed of slechter scoort dan Nico. In vergelijking met de andere BB-leerlingen doet Nico het dus goed. Vergelijken we Nico echter met havoleerlingen, dan zien we dat hij voor taalverzorging tot de 10 procent slechtst scorende leerlingen behoort (zie de percentielscores bij havo in de zevende kolom). Op havoniveau heeft Nico dus niet zo goed gepresteerd. Om te weten welk schooltype goed past bij Nico, kunnen we het beste kijken naar het totaal van deze toets (onderste regel). Daarop is te zien dat een kleine meerderheid van de vmbo GT-leerlingen even goed of slechter presteert dan Nico. Vergelijken we de prestaties van Nico met vmbo BB- of vmbo KB-leerlingen, dan zien we dat er slechts een heel kleine minderheid is die op dat niveau beter heeft gepresteerd dan hij. Voor havo geldt juist het

Rekentoets VO

In 2010 is de Wet referentieniveaus Nederlandse taal en rekenen van kracht geworden. Doel van deze wet is tot een betere aansluiting te komen tussen het taal- en rekenonderwijs in de verschillende onderwijssectoren. Dit komt de doorlopende leerlijn van leerlingen ten goede. De wet is gebaseerd op het werk van de Expertgroep Doorlopende Leerlijnen: 'Over de drempels met taal en rekenen' (2008) en 'Een nadere beschouwing' (2009). Wat de referentieniveaus precies inhouden wordt in hoofdstuk 2 besproken.

Als gevolg van de hiervoor genoemde wet krijgen leerlingen in het voortgezet onderwijs in het schooljaar 2013-2014 voor het eerst te maken met een centrale rekentoets als onderdeel van het eindexamen. Deze toets wordt geïjkt aan het Referentiekader rekenen. De centrale Rekentoets VO is bestemd voor alle leerlingen. Dus ook voor de leerlingen die geen wiskunde in hun pakket hebben. Een rekentoets als onderdeel van het eindexamen betekent:

- borging van een basisniveau rekenvaardigheid voor alle leerlingen;
- beschikbaarheid van gegevens over de rekenvaardigheid van leerlingen bij doorstroom naar vervolgonderwijs;
- eenduidige en objectieve meting van de rekenvaardigheid, zodat het rekenniveau van scholen over de jaren heen gevolgd kan worden.

De Rekentoets VO, zie www.cito.nl > Voortgezet onderwijs > Rekentoets VO, is een digitale toets. Hiervoor is gekozen omdat de opgaven zich goed lenen voor toetsen met behulp van de computer en vanwege een aantal logistieke redenen. In de toets stellen we vragen over vier sub-domeinen: Getallen, Verhoudingen, Meten en meetkunde en Verbanden. Bij een aantal opgaven is het de bedoeling dat de leerling een rekenmachine gebruikt. Deze rekenmachine verschijnt automatisch in beeld bij de betreffende opgave, zodat de leerling niet zijn eigen rekenmachine hoeft te gebruiken. In de aanloop naar de definitieve afname kunnen scholen zich inschrijven voor proeftoetsen die een beeld geven van het soort opgaven dat we zullen gebruiken.

De normering van de Rekentoets VO vindt plaats onder de regie van het College voor Toetsen en Examens. Dit gebeurt voor het eerst bij de pilotrekentoets van 2012. De evaluatie van prestaties kan pas definitief plaatsvinden als er een volledig representatieve afname van de toets is geweest. In de praktijk is hier uitsluitend sprake van als leerlingen even gemotiveerd zijn als bij de echte toets. In een proefafname is de motivatie van leerlingen vaak wat lager, omdat slechte prestaties voor hen weinig gevolgen hebben. Daarom heeft het de voorkeur om voor het normeren van de toetsen 'echte' afnames te gebruiken. Bij de introductie van nieuwe toetsen en toetsvormen geldt dat zowel leerlingen als docenten eerst even moeten wennen. Pas na één of twee jaar is het onderwijs helemaal ingespeeld op de nieuwe toetsen, zodat leerlingen dan optimaal kunnen presteren. Hierdoor ontstaat een klassiek kip en ei probleem: men kan pas goed een standaard bepalen als de toetsen twee of drie jaar operationeel zijn, maar om operationeel te zijn moet de toets een goede standaard hebben. Om dit probleem op te lossen, hanteren we in de eerste jaren dat de toets volledig operationeel is een semi-definitieve standaard (Béguin, 2010).

Maar ook voordat deze semi-definitieve standaard is vastgelegd, is er veel onderzoek nodig. De eerste aanzet tot een standaard is gemaakt bij het ijken van de diagnostische toetsen taal en rekenen (zie paragraaf 1.4). Dit zijn toetsen waarmee scholen een helder beeld kunnen krijgen van de taal- en rekenvaardigheid van hun leerlingen. De toetsen zijn gekoppeld aan de referentieniveaus zoals geformuleerd door de commissie Meijerink en ze zijn in drie opeenvolgende schooljaren afgenomen, voor het laatst in 2010-2011. De ijking had als doel om te bepalen hoeveel opgaven leerlingen goed moeten hebben om vast te stellen dat zij een bepaald referentieniveau beheersen. De uitgangspunten van de ijking waren conform het advies van de Expertgroep Doorlopende Leerlijnen en luiden als volgt:

- het ijkpunt voor 1F wordt vastgelegd op de toetsscore waarbij 75 procent van de leerlingen van groep 8 uit het basisonderwijs 2009 dit niveau ook daadwerkelijk beheerst;
- het ijkpunt voor 2F wordt vastgelegd op de toetsscore waarbij 50 procent van de leerlingen van klas 4 vmbo KB 2009 dit niveau ook daadwerkelijk beheerst;
- het ijkpunt voor 3F wordt vastgelegd op de toetsscore waarbij 75 procent van de leerlingen van klas 5 havo 2009 dit niveau ook daadwerkelijk beheerst.

Met het ijkpunt van de diagnostische toetsen taal en rekenen is dus gebruik gemaakt van informatie over prestatieniveaus van leerlingen in 2009. Dit leverde een allereerste indicatieve standaard op. Tijdens de introductie van de diagnostische toetsen lag de inhoud van de toetsen vaardigheden echter nog niet precies vast. Daarnaast ontwikkelt het niveau van de populatie zich nog. Daarom zal vanaf begin 2012 moeten blijken wat de ontwikkelingen zullen zijn ten aanzien van deze allereerste indicatieve standaard en of normering op meer inhoudelijke gronden wenselijk is.

Kijk- en luistertoetsen voor Nederlands en de moderne vreemde talen

De Kijk- en luistertoetsen voor Duits, Engels en Frans zijn schoolexamens, zie www.cito.nl > Voortgezet onderwijs > Schoolexamens. Ze zijn beschikbaar voor vmbo, havo en vwo.

Het doel van deze toetsen is bepalen in hoeverre leerlingen de buitenlandse taal begrijpen. Dit geldt in situaties waarbij:

- de sprekers van de doeltaal deze spontaan spreken;
- het taalgebruik niet zodanig regionaal, sociaal of individueel bepaald is dat het voor de gemiddelde spreker van de doeltaal problemen oplevert;
- het onderwerp geen specialistische kennis vereist;
- het taalmateriaal voor het BB, KB, GL, TL, havo- en vwo-programma een stijgende moeilijkheidsgraad vertoont wat betreft lexicale en syntactische elementen enerzijds en mate van abstractie en informatiedichtheid anderzijds.

Afhankelijk van het niveau maakt Cito in de toetsen gebruik van nieuws, reclameboodschappen, discussies, documentaires, lifestyle- en actualiteitenprogramma's, productdemonstraties, forumgesprekken, reportages, interviews en lezingen. Deze taalgebruiksvormen vertonen specifieke kenmerken van het kijken en luisteren. Voor het trainen van kijk- en luistervaardigheid zijn ook oefen cd-roms beschikbaar. Voor een demo van deze oefen cd-roms ga naar www.cito.nl > Voortgezet onderwijs > Schoolexamens > Kijk- en luistervaardigheid.

Ook voor het vak Nederlands zijn er toetsen Kijk- en luistervaardigheid. Deze toetsen zijn bedoeld voor het vmbo. Met de toetsen voor vmbo BB en KB toetsen we leerlingen op zoekend luisteren, fictie en het uitvoeren van instructies. Voor vmbo GL en TL zijn er de onderdelen tekstbegrip, fictie en sprekersbedoeling.

Wie de Kijk- en luistertoetsen moderne vreemde talen bestelt, kan net zoals bij Toets 0 tot en met 3 intekenen op scoringsservice. Dit houdt in dat de leerlingen de toets maken op antwoordbladen die Cito vervolgens nakijkt. De scholen krijgen een groepsoverzicht retour waarop ze kunnen aflezen welk cijfer de leerlingen voor de toets hebben gehaald. Per leerling geven we ook een ERK-indicatie. ERK staat voor Europees Referentiekader en bestaat uit een raamwerk met beschrijvingen van beheersingsniveaus voor de moderne vreemde talen. Meer informatie hierover staat in hoofdstuk 2 en 4. Doel van het raamwerk is om taalniveaus van leerlingen met elkaar te kunnen vergelijken, zowel nationaal als internationaal. Nu zijn de Kijk- en luistertoetsen geen ERK-toetsen. Ze zijn in de eerste plaats afgestemd op de examenprogramma's. Daarom kan het voorkomen dat een leerling die een voldoende haalt voor een Kijk- en luistertoets eenzelfde ERK-indicatie krijgt als een leerling die een onvoldoende haalt voor deze toets. Naast een cijfer en een ERK-indicatie geeft de scoringsservice bij de Kijk- en luistertoetsen ook informatie over hoe op leerling- en schoolniveau wordt gepresteerd ten opzichte van het landelijk gemiddelde.

We testen de Kijk- en luistertoetsen een jaar van tevoren op de doelgroep. Op basis van de resultaten van die proefafname wordt een toets samengesteld. Opgaven die in de proefafname niet goed functioneren, vallen af. Ondanks deze procedure zijn de toetsen van jaar tot jaar niet precies even moeilijk. Ook bevatten de toetsen elk jaar niet evenveel opgaven waardoor we jaarlijks een nieuwe adviesnormering moeten bepalen die de scores op de toets aan een cijfer koppelt. Omdat het een schoolexamen betreft, zijn scholen niet verplicht om zich hieraan te houden. De data die we retour krijgen bij de scoringservice bij de Kijk- en luistertoetsen dienen als basis voor de adviesnormering. We publiceren de adviesnormering op internet. Deze normering is voor iedereen beschikbaar, ook voor de scholen die geen scoringservice hebben besteld.

Bij de normeringsprocedure speelt het landelijk gemiddelde dat leerlingen van verschillende schooltypes op de toets hebben gehaald een belangrijke rol. Er wordt voor gezorgd dat het gemiddelde cijfer dat leerlingen behalen van jaar tot jaar min of meer constant blijft zodat docenten en leerlingen niet voor verrassingen komen te staan.

Net als bij Toets 0 t/m 3 uit het Cito Volgsysteem wordt bij de Kijk- en luistertoetsen relatief genormeerd. Het uitgangspunt daarbij is dat de vaardigheid van de populatie van jaar tot jaar hetzelfde blijft. Onderzoek wijst echter uit dat het erop lijkt dat er bij Frans en Duits een niveaudaling heeft plaatsgevonden (van Til, 2007). Dat deze niveaudaling niet tot hogere percentages onvoldoendes geleid heeft, komt doordat er relatief genormeerd wordt.

De vraag is natuurlijk of we niet beter kunnen kiezen voor een absolute norm, waarbij we van jaar tot jaar dezelfde standaard hanteren. Het antwoord op deze vraag ligt besloten in de positionering van de Kijk- en luistertoets. Het is geen verplichte toets, maar een schoolexamen dat onder verantwoording van de scholen zelf wordt afgenomen. Een adviesnormering die gebaseerd is op relatief normeren, past vooralsnog goed bij de wens van scholen voor een realistische normering waarbij we niet te veel eisen van leerlingen, maar ook niet te weinig. Een adviesnormering die gebaseerd zou zijn op een absolute norm past minder goed bij deze wens omdat dit bij de daling van het niveau van luistervaardigheid zou kunnen leiden tot hoge percentages onvoldoendes. Waarschijnlijk zou het gevolg daarvan zijn dat docenten hun eigen norm gaan bepalen waardoor de adviesnormering zijn doel voorbij zou schieten. Dit rechtvaardigt de keuze voor een relatieve normering. Het laat echter onverlet dat een stijging of daling van het niveau op deze manier ongemerkt zou kunnen plaatsvinden. Daarom is het van belang om met enige regelmaat te onderzoeken hoe het is gesteld met de luistervaardigheid van leerlingen in de moderne vreemde talen.

TaalstERK-toetsen

Met de TaalstERK-toetsen, zie www.cito.nl > Voortgezet onderwijs > Taaltoetsen bij het ERK, wordt de lees- en luistervaardigheid van leerlingen getoetst volgens de richtlijnen van het Europees Referentiekader (ERK). Er zijn TaalstERK-toetsen voor lees- en luistervaardigheid voor de talen Duits, Engels en Frans op niveau A1 t/m B2. De TaalstERK-toetsen kunnen worden ingezet vanaf de onderbouw tot en met de examenklas. Leerlingen die slagen voor een toets, krijgen een bij het betreffende niveau behorend certificaat. Dat certificaat is erkend door de Raad van Europa. Daarmee kunnen de leerlingen ook in het buitenland laten zien wat ze waard zijn. Ook in het mbo en hbo is het handig om over een TaalstERK-certificaat te beschikken, want daar is het ERK al helemaal ingeburgerd. In de loop der jaren kunnen leerlingen meerdere certificaten verzamelen. Dat laatste is natuurlijk afhankelijk van de taal en of de leerling het vak laat vallen of niet. Maar ook leerlingen die niet verder gaan met het vak Duits of Frans, hebben er baat bij om over één of meer certificaten te beschikken.

De TaalstERK-toetsen verschillen inhoudelijk van de examens leesvaardigheid en de Kijk- en luistertoetsen. Dat komt omdat er bij het zoeken naar geschikte teksten rekening is gehouden met de verschillende domeinen en descriptoren die beschreven worden in het ERK. Met domein bedoelen we de situatie waarin de tekst voor kan komen, bijvoorbeeld professioneel of publiek domein. Descriptoren beschrijven het soort taalhandeling dat wordt gevraagd, bijvoorbeeld oriënterend lezen of luisteren naar aankondigingen en instructies. Omdat de TaalstERK-toetsen inhoudelijk zijn gebaseerd op het ERK zijn toetsen ontstaan met een heel gevarieerd tekstaanbod. Voor een demo van de toetsen ga naar www.cito.nl > Voortgezet onderwijs > Taaltoetsen bij het ERK > TaalstERK Lezen en Luisteren.

Bij de constructie van de TaalstERK-opgaven zijn voor elk niveau opgaven gemaakt. Deze opgaven zijn in verschillende klassen en op verschillende schooltypen uitgetoetst. Zo zijn er bijvoorbeeld afnames georganiseerd voor de opgaven Frans A2 in de klassen 4 vmbo, 3 en 4 havo én 3 en 4 vwo. Dat komt omdat we van tevoren niet precies konden inschatten hoe moeilijk de opgaven uit het domein A2 zouden uitvallen. Op basis van de resultaten van de proefafnames zijn de definitieve toetsen samengesteld.

Om de TaalstERK-toetsen te kunnen normeren, is vervolgens een standaardonderzoek uitgevoerd. Dat was nodig omdat we niet aan de hand van leerlingprestaties konden vaststellen hoe vaardig leerlingen precies moesten zijn om in aanmerking te komen voor een certificaat op niveau A2. In het standaardonderzoek hebben Nederlandse experts een deel van de opgaven afzonderlijk bekeken en aangegeven hoe deze zich verhouden tot een bepaald ERK-niveau. De resultaten van deze gegevens zijn gekoppeld aan de hele toets waardoor we uiteindelijk konden vaststellen hoeveel punten een leerling voor de toets moest halen om een ERK-niveau te behalen.

De normeringsprocedure die bij de TaalstERK-toetsen is gevolgd, leidt tot een absolute normering. In tegenstelling tot een relatieve normering, wordt bij een absolute normering een standaard gekozen. Daarmee wordt gegarandeerd dat van jaar tot jaar precies dezelfde eisen worden gesteld. Dat is prettig voor werkgevers of vervolgoopleidingen, want zo weten ze precies wat ze kunnen verwachten van iemand die een TaalstERK-certificaat heeft behaald. Het certificaat heeft als het ware een vaste waarde en kan niet devalueren. Andere aspecten die samenhangen met een absolute normering komen naar voren uit de tabellen hierna. Daarin geven we voor luistervaardigheid aan welke TaalstERK-toets geschikt is voor welk leerjaar. De tabellen zijn gebaseerd op de resultaten van proefafnames ten behoeve van de toetsen.

Afname-advies TaalstERK Luisteren

In de tabellen staat in welke klassen de toetsen afgenomen kunnen worden. Tussen haakjes betekent dat de toets alleen bestemd is voor de betere leerlingen.

TaalstERK Luisteren Duits

	vmbo BB	vmbo KB	vmbo GL/TL	havo	vwo
Na enkele lessen in klas 2	A1	A1	A1	A1	A1
Klas 2 (halverwege)					A2
Klas 2 (eind)			(A2)	A2	
Klas 3 (eind)					(B1)
Klas 4 (halverwege)			A2		B1
Klas 4 (eind)				(B1)	
Klas 5 (halverwege)				B1	
Klas 6 (eind)					(B2)

TaalstERK Luisteren Engels

	vmbo BB	vmbo KB	vmbo GL/TL	havo	vwo
Na enkele lessen in klas 1	A1	A1	A1	A1	A1
Klas 1 (halverwege)					A2
Klas 1 (eind)			(A2)	A2	
Klas 2 (eind)			A2	(B1)	(B1)
Klas 3 (eind)		A2			B1
Klas 4 (eind)			(B1)	B1	
Klas 5 (eind)					(B2)
Klas 6 (eind)					B2

TaalstERK Luisteren Frans

	vmbo GL/TL	havo	vwo
Klas 1 (eind)	(A1)	A1	A1
Klas 2 (eind)	A1		
Klas 3 (eind)			(A2)
Klas 4 (eind)		(A2)	A2
Klas 5 (eind)		A2	(B1)
Klas 6 (eind)			(B2)

Als we de tabellen voor luistervaardigheid bekijken, dan valt op dat leerlingen voor Duits en Engels eerder een toets van een bepaald ERK-niveau halen dan voor Frans. Dat komt doordat deze laatste taal minder verwant is met het Nederlands dan Duits en Engels en dus minder makkelijk in het gehoor ligt. Bovendien worden de leerlingen buiten school maar zeer beperkt blootgesteld aan het Frans. De mate waarin ze met de taal kunnen oefenen, is dus heel beperkt. Ten slotte krijgen leerlingen Engels al aangeboden op de basisschool, terwijl het leertraject voor Frans over het algemeen pas start in de brugklas van het vo. Tot voor kort was het onmogelijk om de prestaties voor de verschillende talen onderling met elkaar te vergelijken. Door gebruik te maken van toetsen die leerlingen allemaal langs dezelfde meetlat leggen (in dit geval het ERK) kan dat nu wel.

Wat bij bestudering van de tabellen ook opvalt, is dat bepaalde ERK-niveaus niet altijd worden bereikt aan het einde van een schooljaar, maar soms ook halverwege. Dit geldt bijvoorbeeld voor Duits luistervaardigheid in het vwo. De toets B1 is geschikt voor een afnamemoment halverwege klas 4, wat wil zeggen dat de meeste leerlingen pas op dat moment niveau B1 bereiken. Daarnaast constateren we dat leerlingen er soms meerdere jaren over doen om een bepaald niveau te halen. De meeste vwo-leerlingen bereiken bijvoorbeeld voor luistervaardigheid Frans het niveau A1 aan het einde van klas 1. De toets A2 is echter pas geschikt voor leerlingen eind klas 4. Eind klas 3 zijn er al enkele leerlingen die het niveau zouden kunnen halen, maar voor de meesten geldt dit niet. Dat wil zeggen dat vwo-leerlingen er voor Frans over het algemeen wel drie jaar over doen om van niveau A1 op niveau A2 te komen.

Vooraf het feit dat er niet elk jaar een volgend ERK-niveau kan worden gehaald, kan lastig zijn om het ERK te gebruiken in de klas. Voor leerlingen en docenten is het namelijk prettig om elk jaar te kunnen zien hoeveel vooruitgang er is geboekt. Als het drie jaar duurt voor bijvoorbeeld niveau A2 wordt bereikt, is dat natuurlijk niet erg motiverend. Daarnaast is het jammer dat de ERK-niveaus en het Nederlandse onderwijssysteem niet mooi met elkaar in de pas lopen. Want het is veel eenvoudiger als de prestatie-eisen voor het eindexamen gelijkgeschakeld zouden kunnen worden met een bepaald ERK-niveau. Gelukkig is het ERK nog niet uitontwikkeld. Internationaal onderzoek zou mogelijk tot een herziening van de niveaus in Nederland kunnen leiden. Daarom gelden de hiervoor gepresenteerde tabellen als voorlopig.

1.2 Toetsen voor het beoordelen van het onderwijsleerproces

Deze paragraaf gaat over toetsen die als doel hebben om leerlingen zo optimaal mogelijk te laten leren oftewel toetsen voor het leren. Eerst zetten we de verschillen tussen toetsen van het leren en toetsen voor het leren op een rijtje. Daarna behandelen we toetsvormen die van toepassing zijn bij toetsen voor het leren en laten we enkele vragen de revue passeren waarop toetsen voor het leren een antwoord kunnen geven.

1.2.1 Toetsen van het leren of toetsen voor het leren

Toetsen van het leren zijn in principe afsluitende toetsen. Daarom worden ze meestal afgenomen aan het einde van een serie lessen, aan het einde van het schooljaar of in de eindexamenklas. Toetsen voor het leren moeten de leerbehoeften van leerlingen in kaart brengen en worden daarom niet na afloop van het leerproces afgenomen, maar tijdens dat proces.

Bij toetsen voor het leren is het ook mogelijk om een toetsmoment in te lassen voorafgaand aan het lesgeven. Dit is vooral handig bij vakken met een cyclische lesopbouw, waarbij bepaalde onderwerpen in de loop van de jaren steeds terugkomen en verder worden uitgebouwd. Of aan het begin van het leerjaar als je als docent te maken hebt met klassen waaraan je vorig jaar niet hebt lesgegeven. Het ligt misschien niet erg voor de hand om een serie lessen te beginnen met een toets, maar het kan veel tijd en frustratie schelen. Als blijkt dat leerlingen bepaalde delen van een onderwerp al goed beheersen, kan lesstof worden overgeslagen. Als blijkt dat veel kennis is weggezakt, is het goed om één of meerdere opfrislessen in te lassen en zo te voorkomen dat er over de hoofden heen wordt gepraat.

Waar bij toetsen van het leren altijd een cijfer wordt gegeven, is dat bij toetsen voor het leren meestal niet het geval. Ze dienen immers niet om leerlingen te beoordelen, maar geven de docenten en leerlingen informatie met behulp waarvan ze het leerproces goed op de behoeften kunnen afstemmen.

1.2.2 Welke toetsvorm voor toetsen van het leren en toetsen voor het leren?

Toetsen van het leren en toetsen voor het leren kunnen beide de vorm aannemen van een mondelinge, schriftelijke of digitale toets. Bij toetsen voor het leren is het daarnaast ook mogelijk om te evalueren door middel van individuele evaluatiegespreken, observaties of portfolio's. Ook het bespreken van huiswerk kan een docent informatie geven over het onderwijsleerproces. Hieronder zetten we een aantal evaluatievormen op een rij die in het voortgezet onderwijs veel worden gebruikt om informatie te verzamelen ten behoeve van het leerproces (Stiggins, 2001).

Vragen stellen tijdens de les

Bij het stellen van vragen tijdens de les combineert de docent als het ware evaluatie en instructie. Juist als een leerling iets niet weet, of een fout antwoord geeft, is het van belang om door te vragen, zodat je er als docent achter komt waarom de leerling het niet heeft begrepen. Daar kun je dan meteen dieper op ingaan. Om dit te kunnen doen is het van belang om niet alleen vragen te stellen naar feiten, maar om ook waarom-vragen te stellen (Phye, 1997). Soms kan een docent een leerling zelf iets laten ontdekken door meerdere waarom-vragen achter elkaar te stellen. Daarnaast is het goed om van tevoren te bedenken welke leerlingen een beurt gaan krijgen. Als je een idee wilt krijgen van de mate waarin een klas een bepaalde hoeveelheid stof beheerst, is het handig om veel leerlingen een enkele vraag te stellen. Is het echter de bedoeling om een beter beeld te krijgen van het functioneren van één enkele leerling, dan moet men deze ene leerling juist veel vragen stellen. Wees er ten slotte op bedacht dat elke docent die zijn leerlingen een beetje kent, niet onbevooroordeeld is tijdens het stellen van vragen. Dit

speelt bij mondelinge communicatie een grotere rol dan bij schriftelijke, omdat de houding van de docent tijdens het gesprek de prestaties van de leerling kan beïnvloeden. Voor de leerlingen maakt het natuurlijk een groot verschil of ze worden ondervraagd door een zeer sceptische leraar of door een die duidelijk maakt dat hij alle vertrouwen heeft in de goede afloop van de zaak. Natuurlijk is het onmogelijk elke leerling steeds opnieuw volkomen blanco tegemoet te treden. Het is wel belangrijk dat docenten zich open opstellen en steeds bereid zijn om het oordeel dat ze over een leerling hebben bij te stellen.

Huiswerkcontrole

Huiswerkcontrole geeft informatie over hoe goed leerlingen de stof beheersen. De informatie is vrij beperkt, omdat het meestal slechts enkele oefeningen betreft. Wil je als docent door middel van huiswerkcontrole echt een beeld krijgen van hoe leerlingen de stof beheersen? Dan zul je dergelijke controles dus heel regelmatig moeten doen. Beoordeel het gemaakte huiswerk niet met een cijfer. Als leerlingen thuis hun oefeningen maken, zitten ze in een fase waarin ze zich de stof nog eigen moeten maken. In deze fase is het geoorloofd om fouten te maken. Het geven van cijfers werkt dan juist contraproductief. Het zal met name de zwakkere leerlingen niet stimuleren om zelf aan de gang te gaan met de stof, maar eerder om alles steeds over te schrijven van andere leerlingen (Phye, 1997).

Gesprekken met anderen

Om een goed beeld te krijgen van een bepaalde leerling, voeren docenten nogal eens gesprekken met anderen. Meestal zijn dat de ouders en collega's. Vaak helpt dit een docent om het plaatje compleet te krijgen. Is deze leerling alleen zo ongeconcentreerd bij geschiedenis, of zit hij ook bij andere vakken steeds uit het raam te kijken? Bij dit soort gesprekken is het altijd goed om meerdere meningen te vragen. Wel moet de docent er rekening mee houden dat het praten over leerlingen er al snel toe leidt dat men in stereotypen gaat denken (Stiggins, 2001).

Toetsen voor het leren maken leerlingen duidelijk wat ze al wel beheersen en wat niet. Zo krijgen ze een idee van het leerproces en hun eigen positie daarin. Dat wil zeggen: wat ze al hebben geleerd, maar ook welke inspanningen ze nog moeten leveren om het einddoel te halen. Uit onderzoek komt naar voren dat onderwijs waarin toetsen voor het leren worden gebruikt effectiever is dan onderwijs waarin dat niet wordt toegepast. Hierbij moeten we echter wel opmerken dat toetsen voor het leren alleen tot succes leiden als er ook daadwerkelijk iets gedaan wordt met de uitkomst van de toetsen. Dus het afnemen van een bepaald soort toets is niet genoeg. Na afloop moet de docent bereid zijn om op basis van de resultaten zijn of haar onderwijs aan te passen. En natuurlijk moeten leerlingen op basis van de scores zinnige feedback krijgen, zodat ze weten wat hen te doen staat (Popham, 2004).

1.2.3 Waarom toetsen voor het leren?

Toetsen voor het leren kunnen leerlingen en docenten antwoord geven op één of meer van de volgende vragen (Nitko & Brookhart, 2007):

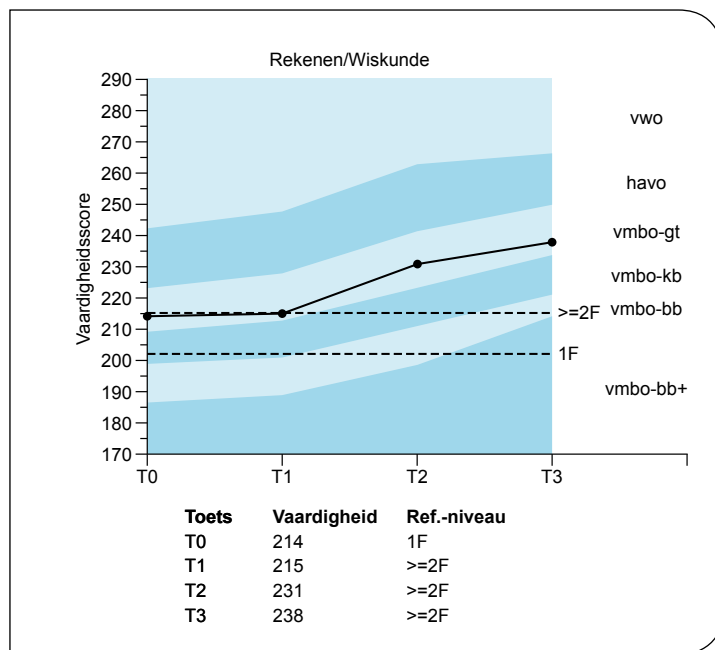
- 1 Worden er voldoende kennis en vaardigheden opgedaan?
- 2 Wat zijn de sterke en zwakke kanten van leerlingen?
- 3 Wat is de oorzaak van fouten die leerlingen maken?

Op deze vragen gaan we hierna verder in.

Worden er voldoende kennis en vaardigheden opgedaan?

Bij de eerste vraag gaat het erom waar leerlingen zich bevinden in het leerproces. Liggen ze op schema om straks bij de afsluitende toets een voldoende te halen? Als blijkt dat het merendeel van de leerlingen onvoldoende vorderingen heeft gemaakt, dan kan de docent beslissen om nog wat meer aandacht aan het onderwerp te besteden dat aan de orde was. Indien de toets aantoont dat slechts enkele leerlingen achterblijven, dan kan voor dit kleine groepje

remediëring worden toegepast. Om te bepalen of er voldoende kennis en vaardigheden worden opgedaan, is het prettig om leerlingen te kunnen volgen in de tijd. Dit is één van de functies van Toets 0 t/m 3 van het Cito Volgsysteem voortgezet onderwijs die nog niet ter sprake is gekomen. Hoe het volgen van leerlingen precies in z'n werk gaat, wordt uitgelegd aan de hand van figuur 1.2.



Figuur 1.2: Voortgangsrapportage leerling

In figuur 1.2 is voor rekenen/wiskunde een vaardigheidsscore van een leerling in beeld gebracht. De vaardigheidsscore is een getal dat uitdrukt in welke mate een leerling een bepaald onderdeel beheerst. Voor Toets 0 t/m 3 liggen alle vaardigheidsscores per leergebied op dezelfde schaal. Op deze manier kunnen de scores van verschillende toetsen met elkaar vergeleken worden. Dit maakt het mogelijk om te zien hoeveel de leerling vooruit is gegaan en of deze leerling zich naar verwachting ontwikkelt.

Op de horizontale as zijn Toets 0 t/m 3 afgebeeld, op de verticale as de vaardigheidsscores. De zwarte lijn geeft de voortgang van de leerling aan. Men kan er ook aan aflezen of de prestaties van deze leerling in overeenstemming zijn met de verwachte ontwikkeling. Wat die verwachte ontwikkeling is, is af te lezen aan de bandbreedtes (de licht en donker gekleurde vlakken). Deze geven aan hoe leerlingen van verschillende schooltypes zich ontwikkelen. De leerling in bovenstaande figuur ontwikkelt zich min of meer zoals verwacht. De lijn loopt immers ongeveer parallel aan het lichtblauw gekleurde vlak.

Wat we in de figuur ook kunnen aflezen, is dat over het algemeen de vooruitgang van leerlingen tussen het eerste en tweede afnamemoment relatief klein is. Het lijkt wel alsof de leerlingen tussen het begin en het einde van het eerste leerjaar wat minder voortgang boeken. Dit verschijnsel zien we terug bij de meeste getoetste vaardigheden binnen Toets 0 t/m 3. Tot nu toe is er geen sluitende verklaring voor gevonden. Het feit dat er minder tijd zit tussen Toets 0 en Toets 1 dan bij de andere toetsen zou een oorzaak kunnen zijn. Een andere mogelijke verklaring heeft wellicht te maken met het gegeven dat leerlingen in het eerste jaar van het voortgezet onderwijs erg moeten wennen aan hun nieuwe leeromgeving.

Ten slotte zijn er in figuur 1.2 ook twee stippellijnen te zien. Deze geven een indicatie van de referentieniveaus van de Expertgroep Doorlopende Leerlijnen voor rekenen. Voor de leerling die is afgebeeld in de figuur geldt dat hij of zij aan het einde van de brugklas, dit is het afnamemoment voor Toets 1, precies op niveau 2 F presteert. Het gaat hierbij echter om een schatting. Dat komt omdat Toets 0 t/m 3 afgestemd is op leerjaren en niet op één van de referentieniveaus. Daarom geldt dat slechts bij benadering kan worden vastgesteld hoe de leerling presteert ten opzichte van die referentieniveaus.

Wat zijn de sterke en zwakke kanten van leerlingen?

Voor leerlingen is het prettig om te weten wat hun sterke en zwakke kanten zijn. Ook voor docenten is dit soort informatie interessant, want zo kunnen ze goed aansluiten op de behoeften van de individuele leerling én op die van de klas. Hoe dat laatste in zijn werk gaat, laten we hierna zien aan de hand van een uitsnede van een niveaurapportage groep van de vernieuwde Toets 0 waarbij aangrenzende niveaus en percentielscores zijn weggelaten. Naast het determineren en volgen van leerlingen hebben Toets 0 tot en met 3 ook als functie om sterke en zwakke kanten van leerlingen in kaart te brengen. De gepresenteerde rapportage richt zich op het onderdeel rekenen/wiskunde dat is gemaakt door een brugklas aan het begin van het nieuwe schooljaar.

Alle scores in deze rapportage zijn groepsgemiddelden en stellen de docent in staat om de prestaties van zijn of haar brugklas te vergelijken met die van andere brugklassen in het land. In deze rapportage zijn de resultaten per vak opgesplitst in deelgebieden. Over het vak als geheel haalt de klas een percentielscore van 51 voor havo. Dat betekent dat 51 procent van de brugklassen in het land even goed scoort als deze klas of slechter. Bekijken we de scores van de klas per deelgebied, dan zien we dat de klas het heel goed doet op het onderdeel Meten en meetkunde, maar vrij slecht op het onderdeel Getallen. Als de docent van deze klas de prestaties van zijn leerlingen wil verbeteren, kan hij dus overwegen om in de lessen meer aandacht te besteden aan het onderdeel Getallen.

Onderdeel	Aantal opgaven	Aantal goed	Vaardigheidsscore	Percentiel	havo					
					0	20	40	60	80	100
Rekenen/wiskunde	68	48	230	51	-----	-----	-x-	-----	-----	-----
• Meten en meetkunde	17	15	251	90	-----	-----	-----	-----	-----	--x--
• Verbanden	17	11	234	66	-----	-----	-----	-x-	-----	-----
• Getallen	17	10	214	26	-----	-x-	-----	-----	-----	-----
• Verhoudingen	17	12	224	44	-----	-----	-x-	-----	-----	-----

Wat is de oorzaak van de fouten die leerlingen maken?

Stel voor dat je als docent hebt vastgesteld dat leerlingen moeite hebben met een bepaald onderwerp. Dan kun je in je lessen natuurlijk extra aandacht besteden aan dat onderwerp en alle aspecten nog eens de revue laten passeren. Je kunt echter ook proberen vast te stellen waarom leerlingen moeite hebben met de stof. Welke denkfout maken de leerlingen? Ontbreekt het hen aan bepaalde basiskennis? Passen ze een bepaalde regel systematisch verkeerd toe? Toetsen die antwoord geven op dit soort vragen heten diagnostische toetsen.

Het is overigens niet altijd nodig om een toets af te nemen om erachter te komen waarom leerlingen bepaalde fouten maken. Vaak kan een docent al snel een idee krijgen van waar de schoen wringt door in de loop van de les enkele vragen te stellen. En als de leerlingen een oefening maken, is het eenvoudig om ze hardop te laten denken. Het voordeel van deze laatste twee methodes is dat ze gemakkelijk zijn toe te passen.

1.3 Toetsen voor het beoordelen van groepen leerlingen en scholen

Toetsen van het leren kunnen gebruikt worden op het niveau van de leerling (zie paragraaf 1.1), maar ook op schoolniveau. De vraag is dan hoe goed de school presteert ten opzichte van het landelijk gemiddelde. Veel partijen zijn in deze informatie geïnteresseerd. In de eerste plaats leerlingen en hun ouders. Bij het zoeken van een school voor aankomende brugklassers zijn de schoolprestaties één van de criteria waarop gelet wordt. In de tweede plaats zijn de scholen zelf geïnteresseerd in de resultaten. Middelbare scholen zijn vaak met elkaar verwickeld in een strijd om de leerlingen en goede schoolprestaties leiden meestal tot een stijging van het leerlingenaantal. Ten slotte vraagt de Inspectie van het Onderwijs, zie www.onderwijsinspectie.nl, ook bepaalde toetsresultaten op, met name de SE en CE-cijfers. Deze gegevens worden gebruikt om opbrengsten in het voortgezet onderwijs te beoordelen.

Iedereen die geïnteresseerd is in de prestaties van vo-scholen kan daarvoor terecht bij de media (o.a. Trouw en Elsevier bieden prestatieoverzichten van scholen), maar ook op de website www.schoolvo.nl. Deze site is een initiatief van het project Vensters voor Verantwoording, zie www.venstersvoorverantwoording.nl. Dit project heeft als doel om gegevens van scholen over bijvoorbeeld slagingspercentages en doorstroom en uitstroom beschikbaar te stellen. Op deze manier kunnen scholen zich verantwoorden naar ouders, gemeenten en de Onderwijsinspectie. Daarnaast maakt Vensters voor Verantwoording het voor scholen mogelijk om onderling zaken te vergelijken en kwaliteitsstandaarden te ontwikkelen. De site biedt niet alleen informatie over de resultaten die een school boekt, maar ook over onderwijsbeleid, kwaliteit en bedrijfsvoering. Naast de site www.schoolvo.nl is er ook een site in ontwikkeling met de naam www.schoolkompas.nl. Deze helpt leerlingen van de basisschool en hun ouders bij het kiezen van een middelbare school.

Toetsen die gebruikt worden voor het beoordelen van groepen leerlingen en scholen zijn vaak dezelfde als de toetsen die gebruikt worden voor het beoordelen van individuele leerlingen. Deze toetsen zijn ter sprake gekomen in paragraaf 1.1.

COOL5-18

Het doel van COOL5-18, zie www.cool5-18.nl, is het verzamelen van gegevens om de ontwikkeling van leerlingen en de factoren die daarop van invloed zijn te beschrijven en te verklaren. Dit onderzoek volgt leerlingen van 5 tot 18 jaar in hun schoolloopbaan door het primair en voortgezet onderwijs en het mbo. Drie aspecten van de ontwikkeling van leerlingen staan in het onderzoek centraal:

- de cognitieve ontwikkeling: kennis en vaardigheden in het Nederlands, Engels en rekenen/wiskunde;
- de ontwikkeling van sociale competenties, waaronder burgerschapscompetenties;
- de sociaal-emotionele ontwikkeling.

Om de leerlingen op deze drie aspecten te volgen, worden op diverse momenten toetsen en vragenlijsten afgenomen. Daarnaast wordt de gehele schoolloopbaan van de leerlingen in kaart gebracht. De benodigde gegevens hiervoor worden verkregen door koppeling van de cohortgegevens aan de onderwijsnummerbestanden. Het Centraal Bureau voor de Statistiek

(CBS) beheert deze bestanden. Cool5-18 kan gebruikt worden als evaluatie-instrument voor schoolbesturen en gemeenten, maar fungeert daarnaast ook als ‘thermometer’ voor ontwikkelingen in het onderwijs en als informatiebron voor allerlei beleidsvragen. Voor de wetenschap biedt COOL5-18 een uniek databestand voor analyses op tal van onderwerpen. Ten slotte fungeert COOL5-18 als databron voor allerlei wetenswaardigheden over het onderwijs in Nederland.

Wat COOL5-18 zo bijzonder maakt, is dat het leerlingen volgt in de tijd. Schoolbesturen beschikken over een aantal instrumenten om te bepalen of hun leerlingen goed presteren. Eén daarvan is bijvoorbeeld het centrale eindexamen. Een directie kan de gemiddelde prestaties van haar leerlingen naast het landelijk gemiddelde leggen om een indruk te krijgen van de kwaliteit van het onderwijs. Maar deze vergelijking zegt niet alles. Binnen een school is het niveau van de eindexamenleerlingen namelijk niet constant. Dat komt omdat de capaciteiten van de leerlingen die eindexamen doen van jaar tot jaar kunnen verschillen. Stel een directie heeft bedacht dat leerlingen systematisch moeten leren plannen. Men bedenkt een plan aanpak waarna alle docenten in hun lessen aandacht geven aan systematisch plannen. Natuurlijk wil zo’n directie dan na verloop van tijd weten of alle inspanningen zoden aan de dijk zetten. Ze kan dan beslissen om de schoolgemiddeldes op de centrale examens te vergelijken met landelijke prestaties, maar de kans is groot dat dit een te beperkt beeld geeft. Omdat leerlingen in de tijd worden gevolgd, geeft COOL5-18 een directie meer houvast. Als blijkt dat leerlingen de eerste jaren onder het landelijk gemiddelde presteren, maar dat na de invoering van het systematisch leren plannen het beeld omslaat, dan is dat een teken dat het gewijzigde beleid succesvol is. Voor een voorbeeld van een schoolrapportage met uitleg ga naar www.cool5-18.nl > Voortgezet onderwijs > Downloads > Schoolrapportage.

1.4 Toetsen voor het beoordelen van de kwaliteit van het Nederlandse onderwijs

Het resultaat van een toets die de kwaliteit van het onderwijs beoordeelt, is voer voor beleidsmakers. Zij kunnen daaruit afleiden of het niveau van de leerlingen voldoende is en of ze de juiste dingen leren. Tegenvallende resultaten kunnen aanleiding zijn om veranderingen door te voeren.

Een voorbeeld van toetsen die gebruikt kunnen worden voor het beoordelen van de kwaliteit van het onderwijs zijn de diagnostische toetsen taal en rekenen. Dit zijn toetsen waarmee prestaties ten aanzien van taal en rekenen kunnen worden beschreven in termen van het Referentiekader van de Expertgroep Doorlopende Leerlijnen. Gedurende drie jaar, vanaf het schooljaar 2008-2009 konden vo-scholen zich kosteloos inschrijven voor deze toetsen. De scholen konden kiezen uit drie toetsversies: één op niveau 1/2F, één op niveau 2F en één op niveau 3F. Elke versie bestond weer uit vier verschillende onderdelen (lezen, luisteren, taalverzorging en rekenen). In 2010 hebben in totaal 81.692 leerlingen minstens één onderdeel van de toets gemaakt.

Om in kaart te brengen hoe leerlingen scoorden ten aanzien van de referentieniveaus is het noodzakelijk dat er bij de gebruikte meetinstrumenten grensscores zijn vastgesteld die aangeven hoe vaardig een leerling moet zijn om een bepaald referentieniveau te halen. De beschrijving van de referentieniveaus geeft aan welke inhoudselementen leerlingen moeten kennen en kunnen om een bepaald niveau te halen. Om vast te stellen of een leerling het referentieniveau voldoende beheerst, zijn opgaven ontwikkeld bij de referentieniveaus. Van deze opgaven weten we dan voor welk niveau ze gemaakt zijn en bij welk niveau ze dus passen. Binnen de ruimte die de referentieniveaus laten, bestaat dan nog steeds de mogelijkheid om moeilijkere en makkelijkere opgaven te maken. Vaak is pas na afname van die

opgaven echt bekend hoe moeilijk ze zijn. Het is goed mogelijk om bijvoorbeeld twee toetsen te maken die beide uit opgaven bestaan die het 2F niveau meten, maar waarvan de ene veel moeilijker opgaven bevat dan de andere. Als er dus een toets met opgaven bij een referentieniveau is ontwikkeld, moeten we nog onderzoeken of het hier om een makkelijke of een moeilijke toets gaat. We moeten vaststellen hoeveel opgaven een leerling goed moet hebben om het referentieniveau te halen. Als de opgaven wat gemakkelijker zijn dan moet een leerling logischerwijs meer opgaven goed hebben dan wanneer de opgaven wat moeilijker zijn (Béguin, 2010).

De relatie tussen een toets en de referentieniveaus wordt dus zowel inhoudelijk als in de gestelde prestatie-eis vastgesteld. Dit gebeurt op twee momenten:

- Tijdens de constructie van de toets, door die opgaven te selecteren die betrekking hebben op de juiste inhoudsdomen en waarvan het niveau past bij de te meten referentieniveaus.
- Bij het bepalen van de grensscore die het beste onderscheid maakt tussen kandidaten die een referentieniveau wel of niet hebben bereikt. Om deze grensscore te bepalen kan onder andere gebruik worden gemaakt van ankeritems. Dit zijn items die in meerdere toetsen voorkomen, zodat de score op de ene toets kan worden vergeleken met de score op de andere toets.

In het kader van de Wet op de referentieniveaus Nederlandse taal en rekenen worden voor het vo en mbo nieuwe toetsen en examens ontwikkeld (zie paragraaf 1.1.2). De normering van deze verplichte toetsen vindt plaats onder de regie van het College voor Toetsen en Examens, zie www.cvte.nl. Dit gebeurde voor het vo en mbo voor het eerst in 2011, bij pilots ten behoeve van de nieuwe toetsen en examens. In de aanloop hiernaartoe is bij de metingen van 2009 een eerste indicatieve standaard vastgelegd. Met behulp van deze standaard is bepaald hoeveel opgaven leerlingen goed moeten hebben om te bepalen dat zij een bepaald referentieniveau beheersen. In paragraaf 1.1.2 zijn de uitgangspunten van deze indicatieve standaard beschreven.

De resultaten op de verschillende toetsen die zijn gekoppeld aan de referentieniveaus worden jaarlijks gerapporteerd, zie www.cito.nl > Onderzoek en wetenschap > Onderzoek > Wetenschappelijk onderzoek bij producten > Publicaties > Taal en rekenen: meting en verantwoording. In dit rapport is te lezen hoe de leerlingen van een bepaald schooljaar hebben gepresteerd ten opzichte van referentieniveaus. Daarnaast beschrijven we of de leerlingen ten opzichte van een jaar daarvoor zijn vooruit gegaan. In paragraaf 1.1.2 hebben we laten zien dat de rapportages van Toets 0 t/m 3 niet bedoeld zijn om uitspraken te doen over de prestaties van leerlingen over de jaren heen. De diagnostische toetsen taal en rekenen zijn opgezet om dit wel te kunnen. Een voorwaarde daarvoor is het werken met een vaste standaard. Dit betekent dat de toetsen zijn geijkt (zie paragraaf 1.1.2) en dat we vervolgens bepalen hoe leerlingen in achtereenvolgende jaren presteren ten opzichte van die ijking (in dit geval 2F en 3F).

Voor het ministerie van OCW bevatten de verslagen met betrekking tot de prestaties voor taal en rekenen interessante informatie. Stel voor dat de rekenprestaties van de eindexamenleerlingen vmbo achteruit zouden gaan in plaats van vooruit. Nadere analyse zou dan uit kunnen wijzen dat dit vooral ligt aan de leerlingen in één van de vier sectoren. Met behulp van deze informatie kan het ministerie gericht maatregelen nemen. In dit voorbeeld moeten deze maatregelen speciaal toegespitst zijn op leerlingen uit de sector die achterblijft en hoeft er voor de overige sectoren niets geregeld te worden.

2 De inhoud van toetsen en examens

2 De inhoud van toetsen en examens

Hans Kuhlemeier

Onderwijs is gericht op het bereiken van doelen. De wenselijke doelen en inhouden van het voortgezet onderwijs zijn vastgelegd in de kerndoelen voor de onderbouw en de eindtermen en examenprogramma's voor de bovenbouw. De scholen concretiseren deze landelijke doelen en inhouden in schoolwerkplannen en programma's van toetsing en afsluiting. Auteurs van methoden zorgen voor de vertaling van de kerndoelen en eindtermen in concrete leerdoelen, leerstof en toetsen. Docenten, op hun beurt, bepalen welke leerstof ze daadwerkelijk in de klas behandelen en toetsen. De landelijke leerplankaders zijn minder uitgebreid, gedetailleerd en voorschrijvend dan voorheen. Daardoor hebben scholen meer vrijheid om zelf leerstof te kiezen en te toetsen. In dit hoofdstuk bespreken we de belangrijkste doelen en inhouden waar docenten zich in hun onderwijs en toetsing op moeten richten. In de verschillende paragrafen bespreken we achtereenvolgens de kerndoelen voor de onderbouw, de eindtermen en examenprogramma's voor de bovenbouw van vmbo en havo/vwo en de referentieniveaus voor Nederlandse taal en rekenen voor het gehele voortgezet onderwijs.

2.1 De kerndoelen voor de onderbouw

In 2006 heeft de overheid nieuwe kerndoelen voor de onderbouw van het voortgezet onderwijs gepubliceerd, zie <http://ko.slo.nl/kerndoelen>. De kerndoelen zijn geordend in zeven domeinen: Nederlands, Engels, wiskunde, mens en natuur, mens en maatschappij, kunst en cultuur, bewegen en sport. De kerndoelen heten kerndoelen omdat ze aangeven wat de kern van het vakgebied is. Kerndoelen worden regelmatig bijgesteld. Zo heeft de communicatieve functie van taal in de kerndoelen Nederlands een centrale plaats gekregen. Bij wijze van voorbeeld presenteren we hierna de tien kerndoelen voor het domein Nederlands.

Kerdoelen Nederlands

1. De leerling leert zich mondeling en schriftelijk begrijpelijk uit te drukken.
2. De leerling leert zich te houden aan conventies (spelling, grammaticaal correcte zinnen, woordgebruik) en leert het belang van die conventies te zien.
3. De leerling leert strategieën te gebruiken voor het uitbreiden van zijn woordenschat.
4. De leerling leert strategieën te gebruiken bij het verwerven van informatie uit gesproken en geschreven teksten.
5. De leerling leert in schriftelijke en digitale bronnen informatie te zoeken, te ordenen en te beoordelen op waarde voor hemzelf en anderen.
6. De leerling leert deel te nemen aan overleg, planning en discussie in een groep.
7. De leerling leert een mondelinge presentatie te geven.
8. De leerling leert verhalen, gedichten en informatieve teksten te lezen die aan zijn belangstelling tegemoet komen en zijn belevingswereld uitbreiden.
9. De leerling leert taalactiviteiten (spreken, luisteren, schrijven en lezen) planmatig voor te bereiden en uit te voeren.
10. De leerling leert te reflecteren op de manier waarop hij zijn taalactiviteiten uitvoert en leert, op grond daarvan en van reacties van anderen, conclusies te trekken voor het uitvoeren van nieuwe taalactiviteiten.

Waren er eerder 109 kerndoelen voor de basisvorming, tegenwoordig zijn het er nog maar 58. Tegelijkertijd zijn de nieuwe kerndoelen veel globaler geformuleerd dan de vorige versies. De publicatie van minder en globalere kerndoelen voor de onderbouw past in het streven de school meer vrijheid te geven bij de invulling van het onderwijsprogramma. Daarnaast wil de overheid één doorlopende leerlijn realiseren van het primair onderwijs naar de bovenbouw van vmbo, havo en vwo met de kerndoelen voor de onderbouw als tussenschakel. In de eerste twee leerjaren moeten scholen tweederde van de tijd aan de 58 kerndoelen besteden. De overige tijd kunnen zij naar eigen keuze invullen. Deze vrije ruimte geeft scholen meer ruimte voor eigen leerstofkeuzes zodat zij leerlingen onderwijs op maat kunnen bieden.

Kerdoelen beschrijven in grote lijnen wat de inhoud van het onderwijs zou moeten zijn, maar niet op welke wijze de scholen deze leerstof moeten aanbieden. Dat laat de overheid geheel over aan de professionaliteit van de scholen. De nieuwe kerndoelen bevatten daarom geen suggesties voor de pedagogisch-didactische en organisatorische vormgeving van het onderwijs. De nieuwe kerndoelen bieden dus meer vrijheid, maar ook minder houvast. Voor degenen die daar behoefte aan hebben, heeft SLO zogeheten voorbeelduitwerkingen uitgebracht, zie <http://ko.slo.nl/00001/>. Bij elk kerndoel zijn voorbeelden van leerstofinhouden, didactische aanwijzingen en lesactiviteiten opgenomen. De voorbeelduitwerkingen zijn bedoeld als inspiratiebron om tot schooleigen keuzes te komen. Leraren kunnen er zelf mee aan de slag in hun lessen. Ter illustratie staat hierna een deel van de voorbeelduitwerking bij het kerndoel Nederlands 'De leerling leert zich mondeling en schriftelijk begrijpelijk uit te drukken'.

Mogelijke lesactiviteiten

Als veel leerlingen het moeilijk vinden om te spreken voor de groep, kan een (wekelijks) kringgesprek een opstapje zijn, bij voorkeur met 10-15 leerlingen. Leerlingen vertellen dan in informele sfeer iets aan de klas over eigen belevenissen of over de aanpak van een opdracht. Ze krijgen wel aanwijzingen voor duidelijk spreken, voor begrippen omschrijven als ze woorden nog niet kennen, voor naar elkaar luisteren en met respect op elkaar reageren, maar er wordt geen beoordeling gegeven.

Een delegatie van twee leerlingen gaat bij de directie toestemming vragen voor iets wat de klas graag wil. In groepjes wordt dit gesprek voorbereid zodat het tweetal van allerlei tips kan worden voorzien. Het gesprek kan geoefend worden in de klas en na het gesprek wordt gezamenlijk gereflecteerd. Varianten van gesprekken die echt gevoerd kunnen worden: solliciteren op een baantje of stageplek of een klacht indienen in een winkel.

Meer vrijheid voor de scholen betekent tegelijkertijd meer verantwoordelijkheid. De school mag bijvoorbeeld niet zomaar kerndoelen uit het onderwijsprogramma schrappen. Scholen moeten tegenwoordig verantwoording afleggen van de geleverde kwaliteit. Met de publicatie van de nieuwe kerndoelen wil de overheid de scholen ook een referentiekader bieden voor kwaliteitszorg en publieke verantwoording. De kerndoelen vervullen echter geen innovatieve functie. Vernieuwingen worden tegenwoordig niet meer centraal door de overheid geïnitieerd en geregisseerd; in het voortgezet onderwijs vindt innovatie heden ten dage vooral van onderop plaats, door en vanuit de scholen zelf. Innovatie is in de praktijk vooral een interactief zoek- en leerproces naar kwaliteitsverbetering dat scholen op basis van vrijwilligheid uitvoeren.

Het Nederlandse leerplanbeleid met globale doelen en veel vrijheid voor eigen invulling vertoont kenmerken van centrale en decentrale sturing. Beide vormen van sturing hebben voor- en nadelen. Gedetailleerde en voorschrijvende onderwijsdoelen bieden scholen houvast. Docenten verkeren dan niet in onzekerheid over wat ze moeten onderwijzen en toetsen, wanneer dat moet gebeuren en hoeveel tijd ze aan de verschillende onderwerpen en vaardigheden moeten besteden. Een nadeel is dat docenten zich kunnen gaan voelen als uitvoerders van het centrale overheidsbeleid ('meteraflezers'). Meer vrijheid voor eigen invulling kan motiverend en kwaliteitsverhogend werken omdat het de professionele ontwikkeling en betrokkenheid van docenten stimuleert. Het gevaar is echter dat docenten nog meer dan voorheen houvast gaan zoeken bij de methode. Een ander gevaar is dat het lastig wordt om doelgericht aan grootschalige vernieuwing te werken. Meer informatie hierover is te vinden in de publicatie 'Leerplan in ontwikkeling' van SLO, zie www.slo.nl > Publicaties > Productencatalogus > Zoekopdracht: leerplan in ontwikkeling.

Tussendoelen Nederlands, Engels en wiskunde/rekenen

In het voorjaar van 2012 heeft de SLO voor de kernvakken Nederlands, Engels en wiskunde/rekenen tussendoelen gepubliceerd. Deze nog niet wettelijk vastgestelde tussendoelen zijn op verzoek van het Ministerie van OCW toch alvast openbaar gemaakt om het onderwijsveld de gelegenheid te geven om het onderwijsaanbod en methodes op de tussendoelen aan te passen en hier ervaringen mee op te doen. U kunt ze downloaden via <http://www.slo.nl/nieuws/00291/>

2.2 De examenprogramma's en syllabi voor de bovenbouw

Eindtermen en eindexamenprogramma's

Voor de bovenbouw van vmbo, havo en vwo gelden sinds 2006 nieuwe eindtermen.

Eindtermen zijn omschrijvingen van de kennis, inzicht en vaardigheden waarover leerlingen aan het eind van hun schoolloopbaan moeten beschikken. De eindtermen zijn beschreven in de examenprogramma's. Daarin is globaal aangegeven welke leerstof (exameneenheden, domeinen) in het centraal examen en het schoolexamen getoetst wordt.

De eindexamenprogramma's voor de verschillende vakken en onderwijssoorten zijn te vinden op www.examenblad.nl.

De eindtermen voor de bovenbouw van het vmbo zijn gegroepeerd in exameneenheden. Het volgende overzicht toont een gedeelte van de exameneenheid Spreek- en gespreksvaardigheid (NE/K/5) uit het Examenprogramma Nederlands vmbo.

NE/K/5 Spreek- en gespreksvaardigheid

De kandidaat kan:

- 1 strategieën hanteren ten behoeve van de spreek- of gespreksituatie (zich voorbereiden, informatie verwerven, informatie verwerken, informatie verstrekken, reflecteren op de eigen deelname);
- 2 compenserende strategieën kiezen en hanteren wanneer de eigen taalkennis tekortschiet (omschrijvingen gebruiken, parafraseringen gebruiken, non-verbale middelen benutten);
- 3 het spreek-/luisterdoel in de situatie tot uitdrukking brengen (informatie geven, informatie vragen, overtuigen, een mening geven, tot handelen aanzetten);
- 4 het spreek-/luisterdoel en taalgebruik richten op verschillende soorten publiek (taalgebruik: woordkeus, intonatie, tempo, articulatie, houding, publiek, directe omgeving, instanties, gesprekspartners met een hogere status);
- 5 het spreekdoel van anderen herkennen en de reacties van anderen inschatten;
- 6 het verschijnsel van in Nederland voorkomende taalvarianten onderkennen en er in spreek- en gespreksituaties op inspelen (taalvarianten: dialecten, groepstalen, talen van allochtone en andere minderheden, Nederlands als standaardtaal).

De eindtermen voor havo en vwo zijn verdeeld in domeinen, zoals vaktheorie en praktijk, en subdomeinen, zoals kunst intercultureel en burgerlijke cultuur van Nederland in de zeventiende eeuw. Onderstaand voorbeeld bevat de omschrijving van het domein Mondelinge taalvaardigheid uit het Examenprogramma Nederlands havo/vwo.

Domein B: Mondelinge taalvaardigheid

De kandidaat kan ten behoeve van een voordracht, discussie of debat (ter keuze van de school):

- relevante informatie verzamelen en verwerken;
- deze informatie adequaat presenteren met het oog op doel, publiek en gespreksvorm;
- adequaat reageren op bijdragen van luisteraars of gespreksdeelnemers.

De syllabi

Tegenwoordig zijn de eindtermen om dezelfde redenen als de kerndoelen in globale termen geformuleerd. Een voor docenten belangrijk document is de syllabus die in beginsel jaarlijks verschijnt. De syllabus bevat een nadere beschrijving van en toelichting op de eisen voor het centraal examen. Zo kan de docent zich een goed beeld vormen van wat er in dat examen wel en niet gevraagd kan worden. Een syllabus geeft echter geen volledige beschrijving van de inhoud van het examen. Het centraal examen kan ook leerstof toetsen die niet expliciet in de syllabus staat,

maar die naar het algemeen gevoelen daarvan in het verlengde ligt. Een syllabus is met andere woorden slechts een hulpmiddel voor docenten. Behalve voor docenten is het ook een belangrijke bron voor examenmakers, auteurs van methoden, onderwijsbegeleiders en opleiders van docenten.

De syllabi voor de verschillende vakken zijn te vinden op de betreffende examen- en vakpagina's op www.examenblad.nl. Het overzicht hierna is ontleend aan de Syllabus Nederlandse taal voor het vmbo. Het laat zien hoe de examenstof van de zogeheten papieren examens voor de basisberoepsgerichte leerweg (B) verdeeld is over het centraal examen (CE) en het schoolexamen (SE).

Exameneenheden	B	CE	Moet op SE	Mag op SE
NE/K/1 Oriëntatie op leren en werken	X		B	
NE/K/2 Basisvaardigheden	X		B	
NE/K/3 Leervaardigheden voor het vak Nederlands	X	B	B	
NE/K/4 Luister- en kijkvaardigheid	X		B	
NE/K/5 Spreek- en gespreksvaardigheid	X		B	
NE/K/6 Leesvaardigheid	X	B		B
NE/K/7 Schrijfvaardigheid	X	B	B	B
NE/K/8 Fictie	X		B	

Behalve de globale verdeling van de examenstof over CE en SE kan een syllabus informatie bevatten over een of meer van de volgende onderwerpen: specificaties van examenstof, begrippenlijsten, bekend veronderstelde onderdelen van domeinen of exameneenheden die verplicht zijn op het schoolexamen, bekend veronderstelde voorkennis uit de onderbouw, bijzondere vormen van examinering (zoals computerexamens), voorbeeldopgaven, toelichting op de vraagstelling en toegestane hulpmiddelen. Onderstaand overzicht toont de specificatie van de eindterm Leervaardigheden voor het vak Nederlands (NE/K/3) uit de Syllabus Nederlandse taal voor het vmbo.

NE/K/3 Leervaardigheden voor het vak Nederlands

De kandidaat beheerst een aantal strategische vaardigheden die bijdragen tot de ontwikkeling van het eigen leervermogen.

De kandidaat kan:

- 1 strategieën kiezen en hanteren die afgestemd zijn op het bereiken van verschillende lees-, schrijf-, luister- en kijk- en spreek- en gespreksdoelen;
- 2 strategieën kiezen die het eigen taalleerproces bevorderen;
- 3 compenserende strategieën kiezen en hanteren wanneer de eigen taal- of communicatieve kennis tekortschiet:
 - informatie afleiden uit de context;
 - woordenboek gebruiken;
 - vragen wat iets betekent;
 - omschrijvingen gebruiken;
 - non-verbale middelen benutten.

Algemene vaardigheden in de bovenbouw van het vmbo

Behalve de exameneisen per vak bevatten de examenprogramma's voor het vmbo ook zes algemene onderwijsdoelen die voor alle vakken gelden. Ze bouwen voort op de kerndoelen voor de onderbouw en bereiden voor op het middelbaar beroepsonderwijs. De algemene vaardigheden waarover alle vmbo-leerlingen moeten beschikken, zijn opgenomen in de zogeheten preambule van het ministerie van Onderwijs, Cultuur en Wetenschap die aan het begin van ieder examenprogramma voor het vmbo is opgenomen, zie www.rijksoverheid.nl > Documenten en publicaties > Brochures > Trefwoord: examenprogramma vmbo. Alle docenten die lesgeven in de bovenbouw van het vmbo leveren dus een bijdrage aan het bereiken van deze vak- en sectoroverstijgende doelen. De zes algemene vaardigheden uit de preambule zijn:

- Werken aan vakoverstijgende thema's. In het kader van een brede en evenwichtige oriëntatie op mens en samenleving leert de leerling enig zicht te krijgen op relaties met de persoonlijke en maatschappelijke omgeving.
- Leren uitvoeren. De leerling leert in zoveel mogelijk herkenbare situaties, mede met gebruikmaking van ICT, een aantal schoolse vaardigheden verder te ontwikkelen.
- Leren leren. Mede met gebruikmaking van ICT leert de leerling zoveel mogelijk eigen kennis en vaardigheden op te bouwen. Daartoe leert hij onder andere een aantal strategieën die het leer- en werkproces kunnen verbeteren.
- Leren communiceren. Mede via een proces van interactief leren, leert de leerling een aantal sociale en communicatieve vaardigheden verder te ontwikkelen.
- Leren reflecteren op het leer- en werkproces. Door te reflecteren op het eigen cognitief en emotioneel functioneren, leert de leerling zicht te krijgen op en sturing te geven aan het eigen leer- en werkproces.
- Leren reflecteren op de toekomst. Door te reflecteren op het eigen cognitief en emotioneel functioneren, leert de leerling zicht te krijgen op de eigen toekomstmogelijkheden en interesses.

Tip

SLO heeft de algemene vaardigheid Leren communiceren nader uitgewerkt. Deze uitwerking heeft geen wettelijke status, maar geeft docenten wel wat meer steun bij het begrijpen van wat de wetgever hiermee zoal bedoeld zou kunnen hebben:

- samenwerking (samenwerken met collega's, met kritiek omgaan, conflicten hanteren en functioneren op de werkvloer);
- persoonlijke eigenschappen (zelfstandigheid, initiatief, verantwoordelijkheid, zelfvertrouwen/onafhankelijkheid, inlevingsvermogen, zelfkennis en zelfreflectie, betrouwbaarheid en presentatie);
- inzet en identificatie (prestatiebereidheid en identificatie met beroep);
- normen (loyaliteit en beroepscode naleven);
- veranderingsbereidheid (anticipatie, mobiliteit en opleidingsbereidheid).

De preambule is niet het enige document waarin de algemene vaardigheden voor de bovenbouw van het vmbo zijn opgenomen. Andere vaardigheden zijn per vak uitgewerkt en geïntegreerd in de examenprogramma's voor de verschillende vakken. Hierna staan enkele vakoverstijgende professionele en sociaal-communicatieve vaardigheden uit het Eindexamenprogramma verzorging. Meer informatie over het onderwijzen en toetsen van algemene vaardigheden in het voortgezet onderwijs is te vinden op de website van SLO, zie www.slo.nl > VO > Vmbo > Vakken > Beroepsgerichte programma's > Techniek > Elektrotechniek > Algemene karakteristieken > Bijdrage aan de algemene en specifieke (beroeps)vaardigheden > Algemene vaardigheden uit de preambule.

Professionele vaardigheden

De kandidaat kan:

- planmatig werken;
- reflecteren op het eigen handelen;
- op systematische wijze werkzaamheden uitvoeren;
- de Nederlandse taal functioneel gebruiken;
- tijdens de werkvoorbereiding en de werkuitvoering op basaal niveau berekeningen maken;
- economisch bewust omgaan met materialen en middelen;
- de werkzaamheden op een veilige wijze uitvoeren;
- milieubewust handelen;
- hygiënisch werken;
- zich aan- en inpassen in de bedrijfscultuur;
- samenwerken bij het uitvoeren van werkzaamheden;
- omgaan met verschillen op basis van culturele gebondenheid en geslacht;
- informatie beoordelen op bruikbaarheid, betrouwbaarheid en representativiteit.

Sociale en communicatieve vaardigheden

De kandidaat kan in een praktijksimulatie:

- cliënten ontvangen en afscheid van hen nemen;
- een zakelijk gesprek voeren;
- een onderhoudend gesprek voeren;
- sociale vaardigheden hanteren.

De kandidaat heeft kennis van het ontstaan en de invloed van vooroordelen.

Algemene vaardigheden in de bovenbouw van havo/vwo

De examenprogramma's voor de Tweede Fase kennen geen preambule met algemene vaardigheden die voor alle vakken gelden. De algemene vaardigheden zitten 'verstopt' in de examenprogramma's en syllabi voor de verschillende vakken. Dit bemoeilijkt de schoolbrede afstemming van het onderwijs in de algemene vaardigheden. Het Algemeen Pedagogisch Studiecentrum (APS) heeft de vakoverstijgende vaardigheden uit de verschillende examenprogramma's daarom naast elkaar gelegd. Zij vonden daarbij veel dubbeling, geen of weinig afstemming, onduidelijk gebruik van begrippen en verschillende accenten per vak. APS heeft uitgezocht wat de grootste gemene deler van alle vakken is, de zogeheten kernvaardigheden. Dat resulteerde in vijf hoofdgroepen:

- vaardigheden over voorbereiden en oriëntatie;
- vaardigheden over informatieverwerven;
- vaardigheden over informatieverwerken en -verwerpen;
- vaardigheden over informatieverstrekken;
- reflectieve en metacognitieve vaardigheden.

Een nadere uitleg van deze vijf algemene vaardigheden is te vinden op de website van APS, zie www.aps.nl > Voortgezet onderwijs en voer de zoekopdracht 'algemene vaardigheden' in. Het schoolbreed onderwijzen en toetsen van algemene vaardigheden in de bovenbouw havo/vwo is niet eenvoudig. Voor praktische tips, valkuilen en artikelen met achtergrondinformatie ga naar dezelfde website en volg hetzelfde zoekpad.

2.3 Referentieniveaus voor het voortgezet onderwijs

De kerndoelen zijn streefdoelen met betrekking tot het leerstofaanbod. Het zijn inhoudelijke standaarden en geen prestatiestandaarden die aangeven wat de leerlingen op cruciale momenten in hun schoolloopbaan precies moeten kennen en kunnen. Naast de kerndoelen en de eindtermen zijn er in opdracht van het ministerie van OCW referentieniveaus opgesteld voor Nederlandse taal en rekenen. De referentieniveaus zijn prestatiestandaarden. In de referentieniveaus is nauwkeurig vastgelegd wat de leerlingen op verschillende drempelmomenten (zoals eind po, eind vmbo, etc.) moeten kennen en kunnen op het gebied van taal en rekenen (zie onderstaand overzicht).

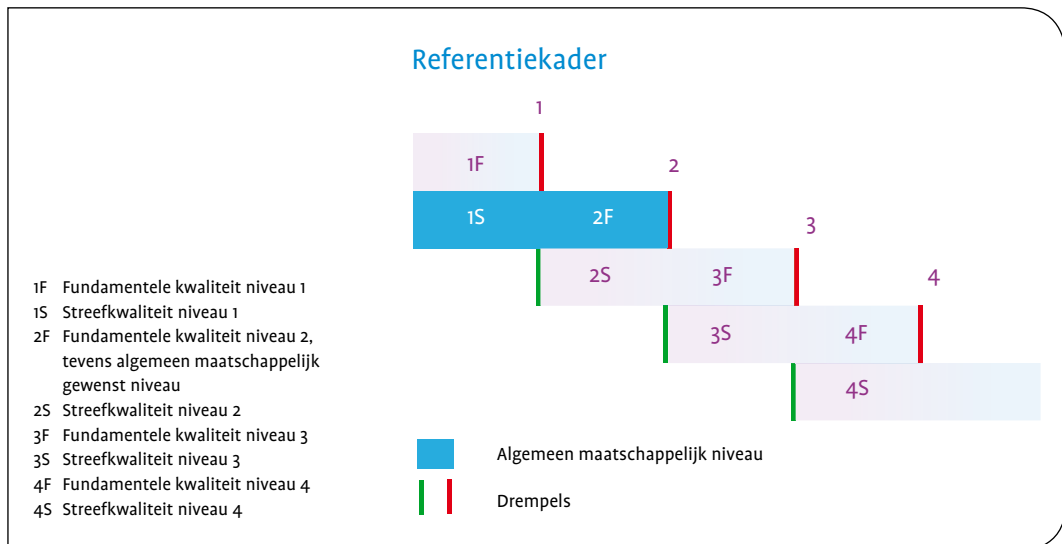
Niveau	Fundamentele kwaliteit	Drempel
1F	Eind primair onderwijs	Van po naar vo
2F	Eind vmbo	Van vo fase 1 naar vo fase 2 / van vmbo naar mbo
3F	Eind mbo-4 en havo	Van vo en mbo naar ho
4F	Eind vwo	Van vo naar wo

Met de referentieniveaus wil de overheid de taal- en rekenvaardigheden van leerlingen verhogen, docenten meer houvast geven bij de invulling van het onderwijs en de aansluiting tussen de verschillende sectoren van het onderwijs verbeteren. De referentieniveaus zijn ook bedoeld als referentiepunt bij de overgang van de ene onderwijssoort naar de andere. Het prestatieniveau van de leerlingen wordt uitgedrukt in de kennis, inzicht en vaardigheden die van belang zijn voor de doorstroom naar het vervolgonderwijs en het maatschappelijk en beroepsmatig functioneren.

Per 1 augustus 2010 zijn de referentieniveaus taal en rekenen wettelijk van kracht en vindt de invoering gefaseerd plaats. De referentieniveaus gelden voor het primair onderwijs, het voortgezet onderwijs en het middelbaar beroepsonderwijs.

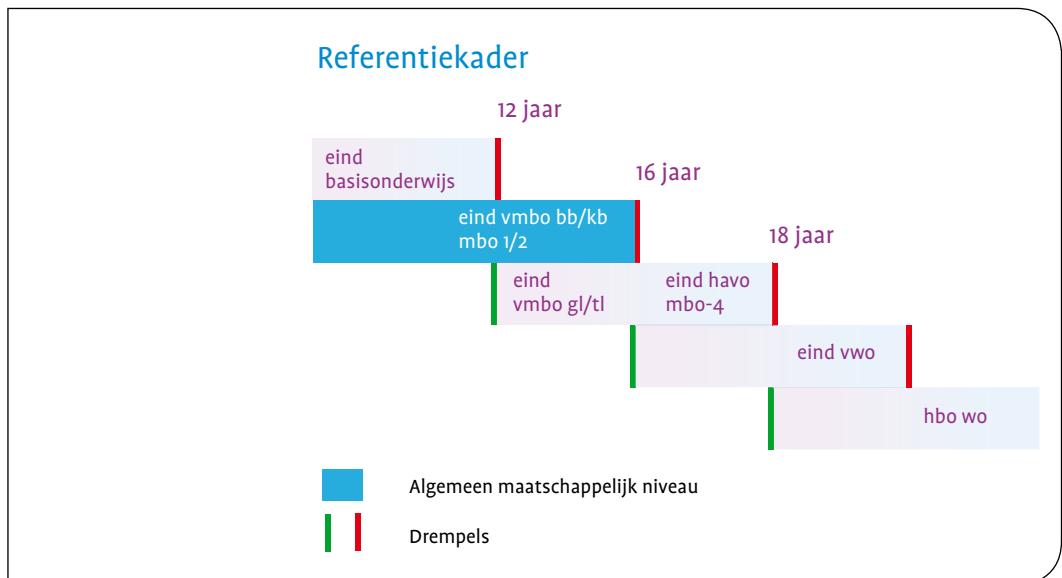
Referentieniveaus

Er zijn vier fundamentele niveaus en vier streefniveaus (1F t/m 4F en 1S t/m 4S) over alle onderwijssectoren (po, so, vo en mbo) heen. Niveau 1F en 1S zijn bijvoorbeeld van toepassing op het einde van het basis- en speciaal (basis)onderwijs, zie figuur 2.1. Hierbij wordt opgemerkt dat niveau 4F en 4S voor rekenen niet zijn ingevuld.



Figuur 2.1: Opbouw referentieniveaus

De koppeling van het referentiekader aan het onderwijsstelsel is weergegeven in figuur 2.2.



Figuur 2.2: Koppeling referentiekader aan het onderwijsstelsel

Voor meer uitleg bij figuur 2.1 en 2.2 dan in het bestek van deze publicatie mogelijk is, verwijzen wij naar het hoofdrapport van de Expertgroep Doorlopende Leerlijnen getiteld ‘Over de drempels met taal en rekenen’ en de website www.taalenrekenen.nl.

Referentieniveaus Nederlandse taal

De referentieniveaus zijn voor verschillende domeinen uitgewerkt. Voor taal zijn dat mondelinge taalvaardigheid, luistervaardigheid, leesvaardigheid, schrijfvaardigheid en taalverzorging en taalbeschouwing. Voor Nederlandse taal zijn er vier niveaus van toenemende beheersing, aangeduid als 1F, 2F, 3F en 4F. De niveaus geven een opklimmende moeilijkheidsgraad in basiskennis en -vaardigheden aan. Elk fundamenteel niveau omvat het voorgaande niveau. Bij het behalen van een fundamenteel niveau kan het volgende niveau gezien worden als een streefniveau. Voor docenten in de onderbouw zijn vooral de niveaus 1F en 2F van belang.

De niveaubeschrijvingen voor de referentieniveaus van het domein taal zijn:

- Mondelinge taalvaardigheid, met de drie subdomeinen gespreksvaardigheid, luistervaardigheid en spreekvaardigheid.
- Leesvaardigheid, met de twee subdomeinen: lezen van zakelijke teksten en lezen van fictionele, narratieve en literaire teksten.
- Schrijfvaardigheid. Schrijven wordt ruim opgevat, het gaat om het produceren van creatieve en zakelijke teksten.
- Begrippenlijst en taalverzorging. In de begrippenlijst staan termen en concepten die leerlingen en docenten nodig hebben om over taalvaardigheid van gedachten te wisselen. Bij taalverzorging gaat het om zaken die in dienst staan van een verzorgde schriftelijke taalproductie.

De vier niveaus zijn omschreven aan de hand van descriptoren en uitgewerkt in 'ik kan stellingen'. Onderstaand voorbeeld bevat de beschrijving van de vier beheersingsniveaus voor het subdomein Gespreksvaardigheid/Gesprekken, een onderdeel van Mondelinge taalvaardigheid.

	Gesprekken			
	Niveau 1F	Niveau 2F	Niveau 3F	Niveau 4F
Algemene omschrijvingen	Kan eenvoudige gesprekken voeren over vertrouwde onderwerpen in het dagelijks leven op en buiten school.	Kan in gesprekken over alledaagse en niet alledaagse onderwerpen uit leefwereld en (beroeps)opleiding uiting geven aan persoonlijke meningen, kan informatie uitwisselen en gevoelens onder woorden brengen.	Kan op effectieve wijze deelnemen aan gesprekken over onderwerpen uit de (beroeps)opleiding en van maatschappelijke aard.	Kan in alle soorten gesprekken de taal nauwkeurig en doeltreffend gebruiken voor een breed scala van onderwerpen uit (beroeps)opleiding en van maatschappelijke aard.

De referentieniveaus taal volstaan niet met een algemene beschrijving van hetgeen leerlingen aan het einde van een onderwijssoort moeten kennen en kunnen. Ze bevatten meestal ook een omschrijving van de concrete taken die de leerling op het betreffende niveau moet kunnen uitvoeren en de eisen waaraan de taakuitvoering moet voldoen. Het volgende voorbeeld toont de omschrijving en bijbehorende prestatie-eisen voor de twee taken van het subdomein Gesprekken, te weten Deelnemen aan discussie en overleg en Informatie uitwisselen.

	Taken			
	Niveau 1F	Niveau 2F	Niveau 3F	Niveau 4F
Deelnemen aan discussie en overleg	<ul style="list-style-type: none"> Kan de hoofdpunten volgen en kan de eigen mening verwoorden en onderbouwen met argumenten. Kan kritisch luisteren naar meningen en opvattingen en een reactie geven. 	<ul style="list-style-type: none"> Kan bespreken wat er gedaan moet worden en bijdragen aan de planning. Kan tijdens een discussie of overleg (op beleefde wijze) een probleem verhelderen, een overtuiging of mening, instemming of afkeuring uitdrukken en commentaar geven op de visie van anderen. 	<ul style="list-style-type: none"> Kan actief deelnemen aan discussies, debatten en overleg in kleinere of grotere groepen. 	<ul style="list-style-type: none"> Kan in een geanimeerde discussie of debat uitgesproken en overtuigende argumenten geven. Kan in een overleg een probleem helder schetsen, speculeren over oorzaken of gevolgen en voor- en nadelen van verschillende benaderingen afwegen.
Informatie uitwisselen	<ul style="list-style-type: none"> Kan in gesprekken binnen en buiten school informatie geven en vragen en kan kritisch luisteren naar deze informatie. Kan informatie beoordelen en een reactie geven. 	<ul style="list-style-type: none"> Kan informatie vragen en geven aan instanties binnen en buiten school. Kan informatie verzamelen en verwerken via het houden van een vraaggesprek. 	<ul style="list-style-type: none"> Kan actief deelnemen aan gesprekken in het kader van werk/ beroepsvoorbereiding, een project of (maatschappelijke) stage en de evaluatie daarvan. 	<ul style="list-style-type: none"> Kan deelnemen aan informatieve, meningsvormende, beschouwende en besluitvormende gesprekken over complexe onderwerpen.

Tot slot zijn de prestatieniveaus en -eisen nog verder verduidelijkt aan de hand van extra kenmerken waaraan de taakuitvoering op het betreffende niveau moet voldoen. Het volgende voorbeeld betreft een verduidelijking van de vier prestatieniveaus voor het beoordelingsaspect Beurten nemen en bijdragen aan samenhang uit het subdomein Gesprekken.

	Kenmerken van de taakuitvoering			
	Niveau 1F	Niveau 2F	Niveau 3F	Niveau 4F
Beurten nemen en bijdragen aan samenhang	<ul style="list-style-type: none"> Kan een kort gesprek beginnen, gaande houden en beëindigen. 	<ul style="list-style-type: none"> Kan de juiste frase gebruiken om aan het woord te komen. Kan een reactie uitstellen totdat hij de bijdrage van de ander geïnterpreteerd en beoordeeld heeft. 	<ul style="list-style-type: none"> Kan op doeltreffende wijze de beurt nemen. Kan standaardzinnen gebruiken (bijvoorbeeld: 'Dat is een moeilijk te beantwoorden vraag') om tijd te winnen en de beurt te behouden. 	<ul style="list-style-type: none"> Kan een passende frase kiezen om eigen opmerkingen op de juiste wijze aan te kondigen en de beurt te krijgen, of om tijd te winnen en de beurt te houden tijdens het nadenken.

Referentieniveaus rekenen

Bij rekenen is gekozen voor een combinatie van drie fundamentele niveaus die meer op toepassing gericht zijn (1F, 2F en 3F) en drie streefniveaus die voorbereiden op de meer abstracte wiskunde (1S, 2S en 3S).

De fundamentele niveaus betreffen basale kennis en inzichten en zijn gericht op een meer toepassinggerichte benadering van rekenen. Daarbij geldt dat een volgend fundamenteel niveau een eerder niveau omvat (2F omvat 1F en 3F omvat 2F). De streefniveaus bereiden voor op de meer abstracte wiskunde. Bij rekenen is er - in tegenstelling tot taal - geen invulling gegeven aan het vierde niveau, omdat men dan geheel in het domein van de wiskunde komt. Ook hier geldt dat een volgend streefniveau een eerder niveau omvat (2S omvat 1S en 3S omvat 2S), maar daarbij omvat het eerste streefniveau (1S) ook nog het eerste fundamentele niveau (1F). 1S is daarbij ook een verdieping van kennis en vaardigheden ten opzichte van 1F.

Voor rekenen zijn er vier domeinen beschreven, te weten:

- Getallen
- Verhoudingen
- Meten en meetkunde
- Verbanden

Elk domein is bij rekenen opgebouwd uit de onderdelen:

- notatie, taal en betekenis, waarbij het gaat om de uitspraak, schrijfwijze en betekenis van getallen, symbolen en relaties en om het gebruik van wiskundetaal;
- met elkaar in verband brengen, waarbij het gaat om het verband tussen begrippen, notaties, getallen en dagelijks spraakgebruik;
- gebruiken, waarbij het gaat om rekenvaardigheden in te zetten bij het oplossen van problemen.

Elk van deze drie onderdelen is steeds opgebouwd uit drie typen kennis en vaardigheden:

- paraat hebben: kennis van feiten en begrippen, reproduceren, routines, technieken;

- functioneel gebruiken: kennis van een goede probleemaanpak, het toepassen, het gebruiken binnen en buiten het schoolvak;
- weten waarom: begrijpen en verklaren van concepten en methoden, formaliseren, abstraheren en generaliseren, blijk geven van overzicht.

Onderstaand voorbeeld geeft een indruk van de prestatieniveaus en -eisen voor het onderdeel Notatie, taal en betekenis uit het domein Getallen. De rechter kolom geeft een nadere toelichting aan de hand van enkele voorbeelden.

Getallen				
	Niveau 1F	Niveau 2F	Niveau 3F	Niveau 3F
	Paraat hebben	Paraat hebben	Paraat hebben	Voorbeelden
Notatie, taal en betekenis: • Uitspraak, schrijfwijze en betekenis van getallen, symbolen en relaties • Wiskundetaal gebruiken	<ul style="list-style-type: none"> • 5 is gelijk aan (evenveel als) 2 en 3 • de relaties groter/kleiner dan • 0,45 is vijftienveertig honderdsten • breuknotatie met horizontale streep • teller, noemer, breukstreep 	<ul style="list-style-type: none"> • schrijfwijze negatieve getallen: -3°C, -150 m • symbolen zoals $<$ en $>$ gebruiken • gebruik van worteltekens, machten 	<ul style="list-style-type: none"> • uitspraak, schrijfwijze en betekenis van negatieve getallen (ook op de rekenmachine) zoals ze voorkomen in situaties met bijvoorbeeld temperatuur, schuld en tekort en hoogte 	<ul style="list-style-type: none"> • het vriest 8 graden kan ook worden weergegeven als: het is -8°C en uitgesproken als 'min 8' of '8 graden onder 0' • tekorten en schulden kunnen weergegeven worden met een minteken • in een tabel de betekenis van positieve (overschotten) en negatieve verschillen (tekorten) aflezen en interpreteren • op de rekenmachine bijvoorbeeld $-5,23$ en $-7,81$ correct intypen

De referentieniveaus beschrijven zoals gezegd wat leerlingen aan het einde van een onderwijssoort moeten kennen en kunnen. Ze geven geen tussendoelen of -niveaus per leerjaar. Op basis van het referentiekader heeft het Freudenthal Instituut in samenwerking met SLO en KPC Groep een doorlopende leerlijn voor rekenen ontwikkeld. De rekenlijn is ontwikkeld voor leerkrachten basisonderwijs en voortgezet onderwijs. De nadruk ligt op de overgang van po naar vo die plaatsvindt in de periode van tien tot veertien jaar. De rekenlijn biedt een overzicht van cruciale leermomenten - de zogeheten kernen - die zich voordoen bij het rekenen in het basisonderwijs en de eerste twee leerjaren van het voortgezet onderwijs. In de vier domeinen Getallen, Verhoudingen, Meten en meetkunde en Verbanden zijn verschillende leerstoflijnen in beeld gebracht. De kernen zijn uitgewerkt aan de hand van toetsopgaven, leerlingenwerk en achtergrondinformatie. Zie www.fi.uu.nl/rekenlijn voor meer informatie.

Referentietoetsen

De overheid heeft zogeheten diagnostische taal- en rekentoetsen in het vooruitzicht gesteld die gebaseerd zijn op de referentieniveaus. Hiermee kunnen scholen vaststellen hoe de leerlingen presteren in vergelijking met de referentieniveaus. Zo nodig kunnen dan gerichte maatregelen worden genomen om het taal- en rekenonderwijs te verbeteren. De school kan de diagnostische toetsen naar verwachting in elk leerjaar afnemen. De afname is vooralsnog niet verplicht en er zijn geen kosten aan verbonden. Meer informatie over deze toetsen is te vinden in hoofdstuk 1 van deze publicatie.

Referentieniveaus en de eindexamens

De referentieniveaus hebben ook consequenties voor de eindexamens in het voortgezet onderwijs. Met ingang van het schooljaar 2013-2014 zullen de eindexamens op de referentieniveaus geijkt zijn. De referentieniveaus Nederlandse taal worden opgenomen in de examenprogramma's en vertaald naar examenopgaven en toetsen. Aan het einde van het basisonderwijs en het voortgezet onderwijs zullen de volgende niveau-eisen gelden:

- niveau 1F: eind basisonderwijs;
- niveau 2F: eind vmbo/mbo 2;
- niveau 3F: eind havo 5 en mbo 4;
- niveau 4F: vwo 6.

Rekentoets VO

Verder zullen de eindexamens vanaf het schooljaar 2013-2014 een rekentoets bevatten.

De zogeheten Rekentoets VO is bestemd voor alle leerlingen, dus ook voor de leerlingen die geen wiskunde in hun pakket hebben. Naar verwachting zal het merendeel van de leerlingen die instromen in het voortgezet onderwijs (ten minste 75%) niveau 1F hebben. Uitgangspunt voor de rekentoets voor vmbo is referentieniveau 2F en voor havo/vwo is dat 3F.

Meer informatie over de referentietoetsen is te vinden in hoofdstuk 1 van deze publicatie en op de website van Cito, www.cito.nl > Onderwijs > Voortgezet onderwijs > Rekentoets VO.

De minister is van plan de rekentoets vanaf het examenjaar 2013-2014 op te nemen in de uitslagregeling. Als de plannen doorgaan, zullen havo- en vwo-leerlingen op de rekentoets ten minste een vijf moeten halen. Vanaf 2015-2016 mogen zij hooguit één onvoldoende (minimaal een vijf) halen voor de kernvakken Nederlands, Engels, wiskunde en de rekentoets. Vmbo-leerlingen moeten vanaf 2013-2014 om te slagen zowel voor Nederlands als de rekentoets minimaal een vijf halen en vanaf 2015-2016 wordt dit aangescherpt tot hooguit één vijf voor Nederlands en de rekentoets. Onderzoekers Zwitser en Béguin (2011) van Cito hebben de toekomstige uitslagregels vergeleken met de huidige situatie en komen tot de conclusie dat het aantal ten onrechte gezakte leerlingen niet of nauwelijks zal toenemen.

Europees Referentiekader (ERK)

De referentieniveaus voor taal en rekenen zijn niet het enige referentiekader voor onderwijzen en toetsen in het voortgezet onderwijs. Een ander referentiekader is het Europees Referentiekader (ERK) voor de moderne vreemde talen (zie ook hoofdstuk 4). ERK is de vertaling van het 'Common European Framework of Reference' (CEFR of CEF). Het kent zes prestatieniveaus, van beginner via gevorderde tot near-native. De zes niveaus (A1, A2, B1, B2, C1 en C2) zijn omschreven aan de hand van descriptoren en uitgewerkt in 'ik kan stellingen'. Onderstaand voorbeeld beschrijft de zes beheersingsniveaus van spreekvaardigheid:

Descriptoren	
Niveau A1	<ul style="list-style-type: none">Ik kan eenvoudige uitdrukkingen en zinnen gebruiken om mijn woonomgeving en de mensen die ik ken te beschrijven.
Niveau A2	<ul style="list-style-type: none">Ik kan een reeks uitdrukkingen en zinnen gebruiken om in eenvoudige bewoordingen mijn familie en andere mensen, leefomstandigheden, mijn opleiding en mijn huidige of meest recente baan te beschrijven.
Niveau B1	<ul style="list-style-type: none">Ik kan uitingen op een simpele manier aan elkaar verbinden, zodat ik ervaringen en gebeurtenissen, mijn dromen, verwachtingen en ambities kan beschrijven.Ik kan in het kort redenen en verklaringen geven voor mijn meningen en plannen.Ik kan een verhaal vertellen of de plot van een boek of film weergeven en mijn reacties beschrijven.
Niveau B2	<ul style="list-style-type: none">Ik kan duidelijke, gedetailleerde beschrijvingen presenteren over een breed scala van onderwerpen die betrekking hebben op mijn interessegebied.Ik kan een standpunt over een actueel onderwerp verklaren en de voordelen en nadelen van diverse opties uiteenzetten.
Niveau C1	<ul style="list-style-type: none">Ik kan duidelijke, gedetailleerde beschrijvingen geven over complexe onderwerpen en daarbij subthema's integreren, specifieke standpunten ontwikkelen en het geheel afronden met een passende conclusie.
Niveau C2	<ul style="list-style-type: none">Ik kan een duidelijke, goedlopende beschrijving of redenering presenteren in een stijl die past bij de context en in een doeltreffende logische structuur, zodat de toehoorder in staat is de belangrijke punten op te merken en te onthouden.

Het ERK biedt docenten moderne vreemde talen een kader voor de beoordeling van de taalvaardigheid van de leerlingen. Behalve de omschrijvingen van de prestatieniveaus biedt het richtlijnen voor de beoordeling en aanwijzingen voor het geven van feedback en ondersteuning. Leerlingen weten wat zij al kunnen en wat zij moeten doen om hun taalniveau te verhogen. Met het Europees Taalportfolio kunnen leerlingen hun ervaringen met het leren van vreemde talen documenteren. Het bestaat uit een talenpaspoort, een taalbiografie en een dossier. Voor meer informatie zie www.europeestaalportfolio.nl.

Cito heeft voor de niveaus A1, A2, B1 en B2 ERK-toetsen Duits, Engels en Frans ontwikkeld. Leerlingen ontvangen een certificaat als zij voor de toets slagen. Voor meer informatie, lees hoofdstuk 1 en 4 van deze publicatie en zie www.erk.nl/docent/toetsing.

Tip

Bronnen voor verdere studie

- www.slo.nl
- www.aps.nl
- www.cinop.nl
- www.cito.nl
- www.examenblad.nl
- www.kerndoelenonderbouw.slo.nl
- www.taalenrekenen.nl
- www.erk.nl
- www.europeestaalportfolio.nl
- www.webcef.eu
- www.cito.nl > Onderwijs > Voortgezet onderwijs > Rekentoets VO
- www.referentiesets.nl
- www.cito.nl > Onderwijs > Voortgezet onderwijs > Cito Volgstelsysteem.

3 Het centraal examen en het schoolexamen

3 Het centraal examen en het schoolexamen

Hans Kuhlemeier

Aan het einde van het voortgezet onderwijs leggen de leerlingen een examen af. Een examen is een door een bevoegde instantie ingestelde toetsing van leerresultaten waaraan de kandidaat bepaalde rechten of bevoegdheden kan ontleen. Het examen voortgezet onderwijs vervult een dubbele functie: afsluiting en toelating. Als de kandidaat voor dat examen slaagt, ontvangt hij of zij een diploma. Dat vormt het bewijs dat de kandidaat het onderwijs op de school met goed gevolg heeft afgesloten (de afsluitingsfunctie). Daarnaast geeft het diploma toegang tot een opleiding voor vervolgonderwijs (de toelatingsfunctie). Zo geeft het vwo-diploma recht op toelating tot het wetenschappelijk onderwijs. Het bezit van het diploma verhoogt ook de kans op het vinden van een baan. Het diploma voortgezet onderwijs heeft nog steeds een grote persoonlijke, maatschappelijke en economische waarde. Dat wordt door vervolgoopleidingen, werkgevers, docenten, ouders en kandidaten ook als zodanig erkend. Dit noemen we het civiele effect van het diploma. Het examen beschermt leerlingen tegen ondeugdelijk onderwijs en geeft de samenleving zekerheid over wat er is geleerd. In dit hoofdstuk zetten we uiteen wat een (beginnende) docent ten minste van het examen moet weten om de leerlingen en zichzelf daar goed op te kunnen voorbereiden. We beginnen met een algemene beschrijving van het examensysteem van het voortgezet onderwijs. Daarna bespreken we het centraal examen en het schoolexamen. Vervolgens gaan we in op het gebruik van examengegevens voor andere doeleinden dan afsluiting en toelating.

3.1 Het examensysteem in het voortgezet onderwijs

De Nederlandse examens bewegen zich in het spanningsveld tussen de verantwoordelijkheid van de overheid en de vrijheid van de school. Aan de ene kant stelt de overheid vanuit het oogpunt van kwaliteitsbewaking eisen aan de inhoud en organisatie van de examens. Aan de andere kant krijgen scholen binnen deze eisen ruimte om het examen naar eigen onderwijsinhoudelijke en levensbeschouwelijke inzichten in te richten. De huidige examens in het voortgezet onderwijs bestaan uit een centraal examen (CE) en een schoolexamen (SE). De verantwoordelijkheid voor de afname en beoordeling van het CE en het SE heeft de wetgever in handen van de school gelegd. Hiermee geeft de overheid aan veel vertrouwen te hebben in de professionaliteit en integriteit van de school en de examinatoren. In dit opzicht nemen de Nederlandse examens in de wereld een unieke plaats in. Hoe uitzonderlijk het Nederlandse examensysteem is, wordt duidelijk als men het aan buitenlandse toetsdeskundigen probeert uit te leggen. De eerste reacties zijn altijd die van onbegrip en ongeloof. Kenmerkende reacties zijn: 'Nou, dat zou bij ons niet werken', 'Weet je wel zeker dat het werkt?', 'En hoe weet je dat het werkt?'

De examens zijn niet altijd geweest zoals ze er nu uitzien. Tot halverwege de jaren zeventig van de vorige eeuw bestond het examen uit een schriftelijk en een mondeling gedeelte. Het mondelinge examen werd afgenomen onder leiding van onafhankelijke deskundigen en gecommiteerden die door de overheid waren aangesteld. Voor bepaalde vakken telden rapportcijfers mee in het eindcijfer. De mondelinge examens waren lastig te organiseren. Tegenwoordig kennen we de situatie van een extern CE en een intern SE. Het huidige SE is ontstaan als vervanger van het mondeling examen en de rapportcijfers. Geruime tijd bestond een groot deel van de

centrale examens alleen uit meerkeuzevragen. Het SE was bedoeld om leerstof te examineren die niet met meerkeuzevragen getoetst kon worden. Daarnaast zou het SE leerlingen als het ware een extra kans moeten bieden, naast het centraal examen. Pas later werd het SE ook gezien als een mogelijkheid voor scholen om hun eigen identiteit te benadrukken. Het wettelijke kader voor de huidige examens is beschreven in het Examenbesluit, zie www.examenblad.nl en geef zoekopdracht 'Examenbesluit'.

Tip

Elke school is wettelijk verplicht een van de personeelsleden als examensecretaris aan te stellen. De examensecretaris is als geen ander op de hoogte van de planning, de voorbereiding, de afname, de correctie en beoordeling, de vaststelling van het resultaat, het bekendmaken van de uitslag en het beheren en bewaren van het examenwerk. Beginnende docenten kunnen met vragen over het CE of SE bij hem of haar terecht.

Staatsexamens

Naast het CE en SE zijn er nog de staatsexamens. In het voortgezet onderwijs is het staatsexamen de enige mogelijkheid om zonder regulier onderwijs te volgen examen te doen en een diploma te behalen. Deze landelijke examens worden niet door de scholen georganiseerd, maar door de overheid. Het staatsexamen bestaat uit de centraal schriftelijke examens en mondelinge commissie-examens die door twee onafhankelijke beoordelaars afgenomen en beoordeeld worden. De staatsexamens zijn onder andere bedoeld voor leerlingen die graag een diploma willen halen maar niet naar school kunnen of willen, bijvoorbeeld omdat ze in het buitenland wonen. Ook leerlingen die alleen een certificaat voor een bepaald vak willen halen kunnen een staatsexamen aanvragen. In 2007 legden in totaal ongeveer 3.500 kandidaten een staatsexamen af. Zie voor meer informatie www.duo.nl en zoek op 'staatsexamens'. Lees ook het artikel 'Een gelijkwaardig diploma voor bijzondere kandidaten. Staatsexamens voor voortgezet onderwijs'. Het is geschreven door Wim Kleijne en staat in nummer 4 van het tijdschrift Examens uit 2008.

3.2 Het centraal examen

Het voortgezet onderwijs kent afzonderlijke eindexamens voor de basisberoepsgerichte leerweg, kaderberoepsgerichte leerweg, gemengde leerweg, theoretische leerweg, havo en vwo. Examens zijn een grootschalig gebeuren. In 2011 waren er 103.000 vmbo-kandidaten, 56.700 havo-kandidaten en 40.000 vwo-kandidaten.

Het College voor Toetsen en Examens (CvTE) heeft de opdracht om namens de Nederlandse overheid de kwaliteit van de centrale examens in Nederland te waarborgen én te zorgen voor een vlekkeloze afname van deze examens. Cito maakt de centrale examens op basis van een constructieopdracht van het CvTE. In de constructieopdracht zijn de productspecificaties voor het uiteindelijke examen opgenomen, zoals de verdeling van de vragen over de examenstof, de keuze van teksten en contextmateriaal, de toe te passen vraagvormen (bijv. gesloten versus open vragen) en de inhoud en vorm van het correctievoorschrift. De opgaven zelf worden doorgaans geschreven door docenten met ervaring in examenklassen. De uiteindelijke samenstelling van het conceptexamen is in handen van een Cito-toetsdeskundige. De conceptexamens worden beoordeeld en zo nodig van commentaar voorzien door de vakcommissies van het CvTE. Daarnaast worden examenconcepten onder de verplichting van strikte geheimhouding ook voor commentaar voorgelegd aan niet-schoolgebonden deskundigen, zoals universitaire deskundigen voor speciale onderwerpen, vakmensen bij de beroepsgerichte vakken en, voor de moderne

vreemde talen, deskundigen die de desbetreffende taal als moedertaal spreken. De CvTE-vakcommissie stelt ten slotte de opgaven en het correctievoorschrift in hun definitieve vorm vast.

De normering is een belangrijk onderdeel van de constructie van een examen. Normeren gaat over het waarderen van de prestaties die examenkandidaten hebben geleverd. Gelijke prestaties en gelijke scores op een examen moeten met eenzelfde cijfer gewaardeerd worden. Normering heeft te maken met de wijze waarop de grens tussen voldoende en onvoldoende bepaald wordt en de scores in cijfers worden omgezet. Sinds 2000 wordt genormeerd volgens de 'CvTE-methode' waarbij de zogeheten normeringsterm wordt vastgesteld, ook wel N-term genoemd. De N-term moet ervoor zorgen dat de exameneisen in verschillende jaren en tijdvakken even zwaar zijn. De N-term varieert van 0,0 tot 2,0. Bij een N-term van 0,0 moet een kandidaat iets meer dan 60% van de scorepunten halen om een voldoende te halen. Bij een N-term van 2,0 heeft een kandidaat al een voldoende bij iets minder dan 40% van de maximale score. De tabellen voor de omzetting van scores naar cijfers worden elk jaar direct na de normering gepubliceerd op de websites www.examenblad.nl en www.cito.nl. Voor meer informatie over de normering wordt verwezen naar de ToetsSpecial 'Normering' op www.toetswijzer.nl.

3.2.1 Soorten centrale examens

Voor het voortgezet onderwijs maakt Cito jaarlijks meer dan vijfhonderd examens. De belangrijkste soorten centrale examens zijn: het Centraal Schriftelijk Examen (CSE) dat op een vast tijdstip wordt gehouden, beeldschermexamens die op een vast of flexibel tijdstip worden gehouden (CBT en Compex), het Centraal Praktisch Examen (CPE) en het Centraal Schriftelijk en Praktisch Examen (CSPE) waarin de theorie en de praktijk van de beroepsgerichte vakken geïntegreerd getoetst worden. Voor het eerste tijdvak van het examen 2011 heeft Cito in totaal 227 examens ontwikkeld (met ruim 1.3 miljoen afnames bij individuele kandidaten). Tabel 3.1 laat zien hoe deze 227 examens zijn verdeeld naar type examen en onderwijssoort.

Tabel 3.1: Verdeling van de examens voor het eerste tijdvak naar type examen en onderwijssoort

	VMBO BB	VMBO KB	VMBO GL	VMBO GL/TL	HAVO	VWO	Totaal
CSE	9	12	7	19	28	30	105
CBT	11	13	0	5	2	2	33
COMPEX	0	0	0	1	0	0	1
CPE	0	0	0	4	0	0	4
CSPE	37	39	8	0	0	0	84
Totaal	57	64	15	29	30	32	227

Schriftelijke examens op een vast tijdstip

Bij de meeste vakken is het papieren examen nog steeds de meest gangbare afnamevorm. De scholen ontvangen het examen enkele weken voor de afname in een gesealde envelop. Op de dag van de afname worden de enveloppen overal in Nederland op hetzelfde tijdstip open gemaakt, waarna de kandidaten met het examen beginnen. Het nakijken gebeurt door de 'eigen' docent die de kandidaten voor het examen heeft opgeleid en door een tweede corrector van een andere school. De Dienst Uitvoering Onderwijs (DUO) zorgt voor de koppeling van de scholen.

Beeldschermexamens


Beeldschermexamens hebben vele voordelen boven papieren examens. Kandidaten vinden een digitaal examen doorgaans aantrekkelijker. Daarnaast neemt de computer de correctie geheel of gedeeltelijk van de docent over. Ook biedt de computer meer mogelijkheden tot adaptief toetsen waarbij de moeilijkheid van de vragen wordt afgestemd op het vaardigheidsniveau van de leerling. Adaptief toetsen voorkomt niet alleen dat leerlingen gefrustreerd raken of zich te weinig uitgedaagd voelen omdat zij te moeilijke of te makkelijke vragen moeten beantwoorden, het komt ook de betrouwbaarheid van de scores ten goede waardoor het examen korter kan zijn en minder afnametijd vergt. In 2010 hebben ongeveer 800 scholen één of meer centrale examens afgenomen waarbij de computer een rol speelde. Bij de zogeheten Compex-examens maken de kandidaten de opgaven op de computer, maar hun antwoorden noteren zij nog op papier. Op circa 550 scholen hebben de kandidaten één of meer examens afgelegd als CBT-examen (Computer Based Testing). Dit is een examen waarbij de kandidaat de vragen op het scherm ziet en de antwoorden via de computer geeft. Van adaptief toetsen wordt tot op heden in de examens voor het voortgezet onderwijs nog geen gebruik gemaakt.

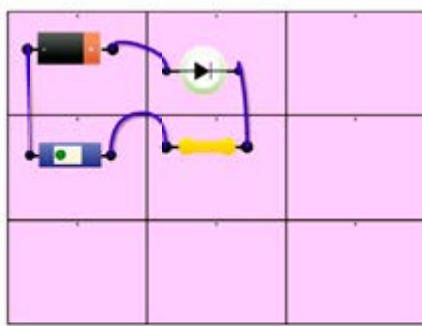
Interactieve beeldschermexamens

Voor enkele vakken is een begin gemaakt met zogeheten interactieve beeldschermexamens. Flash-applicaties voegen audiovisuele effecten en animaties aan het beeldscherm toe en bieden nieuwe mogelijkheden tot interactie met de kandidaat. Daardoor kunnen kennis en vaardigheden op een realistische en levendige manier geëxamineerd worden, bijvoorbeeld via simulatie. In figuur 3.1 presenteren we een karakteristieke opgave uit het experimentele beeldschermexamen NaSk. Het betreft een simulatie waarbij de kandidaten de spanning van een batterij meten met een multimeter.

SPANNING METEN

Multimeter





Je ziet een schakeling van een batterij, een LED, een schakelaar en een weerstand.
Ernaast staat een voltmeter afgebeeld.
Met je muis kun je de pennen van de voltmeter verplaatsen.

Meet de spanning van de batterij.
Vul de waarde in.

 volt

Figuur 3.1: Een interactieve opgave uit het experimentele beeldschermexamen NaSK

Voor meer informatie over experimentele interactieve beeldschermexamens, lees het artikel van Wesseling en Kuhlemeier (2010).

Centraal Schriftelijk en Praktisch Examen (CSPE)

Het Centraal Schriftelijk en Praktisch Examen, ofwel CSPE, is de belangrijkste examenvorm voor de beroepsgerichte vakken in het vmbo. In het CSPE worden de theorie en de praktijk van het beroepsgerichte vak geïntegreerd getoetst. Omdat de praktische opdrachten de nodige

voorbereiding vergen, krijgen de scholen de opgaven en correctievoorschriften lang voor de afname toegezonden. Bij de voorbereiding, de afname en het herkansingenbeleid biedt de overheid de scholen veel ruimte voor eigen invulling. Zo mogen de scholen het moment van afname binnen een vastgestelde periode zelf bepalen. Ook mag de school zelf bepalen hoe vaak kandidaten het praktijkexamen mogen herkansen en of zij ook mogen herkansen om een hoger cijfer te halen, bijvoorbeeld ter compensatie van een slecht gemaakt algemeen vormend vak.

Onderdeel van het CSPE zijn de zogeheten minitoetsen. Minitoetsen zijn korte meerkeuze-toetsen die geheel via de computer worden gemaakt en nagekeken. Scholen kunnen, binnen een door de overheid vastgestelde periode, zelf bepalen wanneer de afname plaatsvindt. Daartoe zijn er per examenvak meestal meerdere toetsvarianten beschikbaar.

3.2.2 De voorbereiding op het centraal examen

Docenten hebben als taak hun leerlingen zo goed mogelijk op het examen voor te bereiden. Uiteraard is de beste voorbereiding op het examen goed onderwijs. Docenten worden geacht de verplichte examenstof in hun klas te onderwijzen. Het is zaak daarbij geen belangrijke leerstof over het hoofd te zien.

De verplichte examenstof is omschreven in het examenprogramma en de syllabus. Deze documenten zijn te vinden op www.examenblad.nl. Het examenprogramma bevat voor elk vakgebied een globale omschrijving van de examenstof en de verdeling daarvan over CE en SE. De syllabus bevat de gedetailleerde eisen voor het CE, maar kan ook andere informatie bevatten, zoals voorbeeldopgaven, toelichtingen op de vraagstelling, begrippenlijsten, overzichten van kennis uit de onderbouw of het SE die als bekend verondersteld wordt, informatie over bijzondere vormen van examinering, bijvoorbeeld beeldschermexamens en de hulpmiddelen, zoals een woordenboek, die kandidaten bij het examen mogen gebruiken. Meer informatie over de inhoud van de examens is te vinden in hoofdstuk 2 van deze publicatie.

Tip

De meeste methoden bevatten veel meer leerstof dan een docent in één schooljaar kan behandelen. Daarnaast bevatten de methoden tegenwoordig grote hoeveelheden op zich leerzame leerstof die voor het behalen van het diploma niet echt nodig is. Als de docent de methode op de voet volgt, kan dat ertoe leiden dat belangrijke examenstof onbehandeld blijft en leerlingen gedupeerd worden. De syllabus wordt elk jaar vernieuwd. Bij aanvang van het nieuwe schooljaar doen docenten er goed aan om de syllabus nauwgezet te bestuderen. Door de stof uit de syllabus te vergelijken met die in de methode kan worden nagegaan welke leerstof meer en minder essentieel is voor de voorbereiding op het examen.

De beroepsgerichte examens voor het vmbo, de CSPE's, kennen uitgebreide richtlijnen voor de voorbereiding van het examen. Zo moet de examiner het correctievoorschrift en de instructie voor de examiner voor de kandidaten geheim houden. Evenmin is het geoorloofd om de examenopdrachten van tevoren met de kandidaten te oefenen. Andere aanwijzingen betreffen de voorbereiding van de lastig te organiseren praktijkopdrachten waarvoor meestal veel materiaal, gereedschap en hulpmiddelen nodig zijn. De examenopdrachten en correctievoorschriften worden daarom al in februari - lang vóór het eerst mogelijke afnamemoment - aan de examinatoren ter beschikking gesteld. Een evident nadeel van de vroegtijdige verspreiding van het examen is de mogelijkheid tot oneigenlijk gebruik van deze informatie. De Inspectie heeft onderzoek gedaan naar de praktijk van de beroepsgerichte examens in het vmbo, zie www.onderwijsinspectie.nl en geef zoekopdracht 'Praktijk

beroepsgerichte examens vmbó'. Uit dat onderzoek is gebleken dat nog niet alle examinatoren zich aan de regels voor de geheimhouding houden. Examinatoren mogen de kandidaten in algemene zin informeren over de opdrachten uit het komende praktijkexamen. Zij mogen de opdrachten uit het examen echter niet vooraf instrueren of met de leerlingen oefenen. Een op de drie examinatoren uit het Inspectieonderzoek zegt dat echter toch gedaan te hebben. Daarnaast gebruiken velen 'oud' examenmateriaal dat sterk lijkt op de opdrachten uit het komende praktijkexamen of hebben zelf vergelijkbaar materiaal ontwikkeld. Deze praktijk staat op gespannen voet met het uitgangspunt dat een examen voor iedere kandidaat dezelfde moeilijkheidsgraad moet hebben. Kandidaten die specifieke instructie en oefening in de opdrachten uit het komende examen ontvangen, zouden in het voordeel kunnen zijn in vergelijking met kandidaten van scholen waar dat niet of veel minder gebeurt. CvTE en Cito hebben diverse maatregelen genomen om de geheimhouding van de CSPE's beter te waarborgen. In de zogeheten Septembermededelingen is extra aandacht gevraagd voor de geheimhoudingsinstructies waarbij precies is aangegeven wat de examinator van te voren bekend mag maken. Een andere maatregel is de zogeheten 'grijze vlekken procedure' bij enkele examens die al in februari naar de scholen verzonden worden. Om de praktische opdrachten te kunnen voorbereiden moet de examinator het opgavenboekje kunnen inzien. Om misbruik van voorkennis van het examen te voorkomen zijn de tekstdelen uit het opgavenboekje die niet nodig zijn voor de organisatorische voorbereiding van de praktische opdrachten onleesbaar gemaakt. Verder heeft Cito cursussen en trainingen verzorgd in het professioneel en integer beoordelen van prestaties op de praktijkopdrachten uit de CSPE's. Tot slot hebben VO-raad en AOC-raad de protocollen voor de geheimhouding aangescherpt.

3.2.3 De afname van het centraal examen

Voor het vertrouwen in het diploma is het van belang dat alle scholen het examen op vergelijkbare wijze afnemen. Om dit te bewerkstelligen heeft het CvTE richtlijnen voor de afname opgesteld. Deze richtlijnen zijn wettelijk verankerd en daarmee bindend voor scholen en docenten. Om scholen te helpen bij de voorbereiding, afname en beoordeling van het examen hebben de VO-raad en de AOC Raad protocollen voor het centraal examen ontwikkeld, zie

www.vo-raad.nl >Thema's > Examens > Protocollen en checklisten. Er zijn protocollen voor de geheimhouding, het omgaan met verhinderde of te laat komende kandidaten, het omgaan met kandidaten die tijdens het examen ziek worden, het omgaan met examenwerk, examenfraude, de eerste en tweede correctie, en het CSPE. De protocollen beschrijven mogelijke en soms voorgeschreven handelwijzen met betrekking tot het centraal examen. Ze dragen ertoe bij dat de voorbereiding, afname en beoordeling van de centrale examens procedureel zuiver verlopen en dat de geheimhouding niet geschonden wordt. De 'Instructie voor de examinator' bevat de richtlijnen waarmee de docent bij de afname van zijn of haar examen rekening moet houden. Dit document wordt voorafgaand aan het examen aan de scholen verzonden, samen met de opgaven en het correctievoorschrift.

De kwaliteit van een examen en de waarde van het diploma is mede afhankelijk van de manier waarop scholen en docenten ermee omgaan. Docenten moeten de richtlijnen voor de afname kennen en nauwgezet opvolgen. Uit het Inspectieonderzoek naar de praktijk van de beroepsgerichte examens is gebleken dat nog maar weinig docenten de richtlijnen voor de afname van het CSPE volledig naleven. Om aan te geven hoe het niet moet, beschrijven we hieronder vier voorbeelden van afnamepraktijken uit het CSPE die in strijd zijn met de afnamevoorschriften, maar die toch nog regelmatig voorkomen:

- Examinatoren dienen alle onderdelen van het examen integraal en ongewijzigd af te nemen. Zij moeten de kandidaten de opdrachten laten uitvoeren zoals aangegeven in het examenboekje en de 'Instructie voor de examinator'. Zij mogen geen opdrachten toevoegen, weglaten, vervangen of aanpassen zoals bijvoorbeeld het weglaten van 'moeilijke'

opdrachten. Ongeveer een derde van de examinatoren zegt de regels voor een integrale en ongewijzigde afname van de opdrachten af en toe of regelmatig te overtreden. Een voorbeeld is de Kassaopdracht uit het examen Handel en Verkoop. Sommige docenten pasten de prijs van een artikel aan zodat de kandidaat het wisselgeld gemakkelijker kon teruggeven. Anderen lieten muntstukken van twee cent weg met als argument dat deze vrijwel nergens meer gebruikt worden.

- Tenzij anders is aangegeven, moeten scholen de materialen, gereedschappen en hulpmiddelen gebruiken die zijn opgenomen in het overzicht van de voorgeschreven middelen uit de 'Instructie voor de examinator'. Ongeveer een derde van de examinatoren zegt af en toe of regelmatig toch andere middelen te gebruiken. Een voorbeeld is afkomstig uit het examen Elektrotechniek. Veiligheid speelt binnen de elektrotechniek een belangrijke rol en het naleven van de veiligheidsvoorschriften is dan ook een belangrijk beoordelingsaspect tijdens het examen. Een van de docenten had de kandidaten niet onderwezen in het gebruik van veilige kleding en schoeisel, maar had hen voor dit beoordelingsaspect toch de maximale score toegekend.
- Kandidaten mogen geen hulp krijgen van de examinator. De kandidaat moet de praktische opdrachten geheel zelfstandig uitvoeren. Uiteraard mogen examinatoren de leerlingen wel aansporen hun best te doen en hen moed inspreken. Het geven van aanwijzingen, uitleg of hints die het examen voor de kandidaten makkelijker maken, is echter niet toegestaan. Slechts een kwart van de examinatoren geeft de kandidaten geen andere dan als aansporing en bemoediging bedoelde hulp. De overigen doen dat af en toe, regelmatig of vaak tot zeer vaak. Voorbeelden zijn opmerkingen en hints als: 'Moet je hier niet nog even naar kijken?', 'Weet je zeker dat dit goed is wat je nu doet?', 'Kan die olie wel in een keer bij de mayonaise?', 'Lees het eerst nog maar een keer voordat je verder gaat', 'Ja, dat bedoelen ze inderdaad met die vraag', 'Jazeker, het water moet koken bij blancheren', 'Vergeet je niet iets?' en 'Weet je zeker dat je klaar bent?'
- Kandidaten mogen tijdens het examen niet met elkaar overleggen of elkaar helpen. De kandidaat mag uitsluitend aan de examinator om hulp vragen, en niet aan medekandidaten. Volgens een op de vijf examinatoren vindt bij hun examen af en toe overleg plaats tussen kandidaten en/of helpen de kandidaten elkaar. Een voorbeeld uit het examen Elektrotechniek is een examinator die niet ingreep toen een leerling zijn installatie in spiegelbeeld maakte omdat hij bij zijn medeleerling afkeek in plaats van op zijn eigen installatietekening te kijken.

3.2.4 De correctie van het CSPE

De correctie van het centraal examen wordt in Nederland uitgevoerd door de 'eigen' docent die de kandidaten heeft opgeleid. De VO-raad heeft samen met de AOC-raad een protocol opgesteld met aanwijzingen voor de eerste en tweede correctie, zie www.vo-raad.nl >Thema's > Examens > Protocollen en checklisten examens.

Correctievoorschrift en antwoordmodel

Uit onderzoek is bekend dat beoordelaars ertoe neigen soepeler te beoordelen als zij degenen die beoordeeld wordt persoonlijk kennen en zelf hebben opgeleid. Daarnaast is bekend dat de ene docent soepeler beoordeelt dan de ander. Mede om deze redenen heeft de overheid bindende richtlijnen voor de correctie ter beschikking gesteld. De richtlijnen die de docent bij het nakijken van het examenwerk moet toepassen staan in het correctievoorschrift. Dat bevat onder meer algemene en vakspecifieke regels voor de beoordeling en een antwoordmodel dat ook wel beoordelingsmodel wordt genoemd.

Enkele belangrijke algemene regels voor de beoordeling uit het correctievoorschrift zijn:

- Indien een prestatie gedeeltelijk juist is, wordt indien het antwoordmodel dit toelaat, een deel van de te behalen scorepunten toegekend in overeenstemming met het

antwoordmodel.

- Indien een andere prestatie is geleverd dan aangegeven in het antwoordmodel en deze is aantoonbaar vakinhoudelijk juist of gedeeltelijk juist worden scorepunten toegekend naar analogie van het antwoordmodel.
- Indien de examinerator voor het begin van het examen meent dat in het examen of in het antwoordmodel een fout of onvolkomenheid zit, deelt hij dit onverwijld mee aan het College voor Toetsen en Examens.
- Indien een vermeende fout of onvolkomenheid pas tijdens de afname blijkt, beoordeelt de examinerator het werk van de kandidaten alsof het examen en antwoordmodel juist zijn. Hij of zij kan de fout of onvolkomenheid dan alsnog mededelen aan het College voor Toetsen en Examens.
- Het is niet toegestaan zelfstandig af te wijken van het antwoordmodel. Met een eventuele fout wordt bij de definitieve vaststelling van de normering van het examen rekening gehouden.

In het antwoordmodel is per vraag of opdracht aangegeven hoe de examinerator de antwoorden, de handelingen of de producten van de kandidaten moet beoordelen. Onderstaand voorbeeld is afkomstig uit het examen Verzorging voor de basisberoepsgerichte leerweg 2010. In een van de praktijkopdrachten speelt de kandidaat de rol van een stagiair in een verpleeghuis en verricht daar allerlei activiteiten. Onder meer moet de kandidaat een bewoner met een rollator naar de wachtruimte begeleiden, als baliemedewerker een telefoongesprek voeren en reflecteren over de manier waarop de taken zijn uitgevoerd. Hieronder staan drie antwoordmodellen aan de hand waarvan de examinerator de kwaliteit van de communicatie met de bewoner moet beoordelen. In de vormgeving is getracht de oorspronkelijke lay-out van het antwoordmodel zoveel mogelijk intact gelaten. Het eerste antwoordmodel bevat twee inhoudselementen die beide in voldoende mate aanwezig moeten zijn om één punt te mogen toekennen. Het tweede antwoordmodel heeft de vorm van een beoordelingsschaal waarbij de examinerator 0, 1 of 2 punten toekent. Het derde antwoordmodel kent vijf inhoudselementen met een maximumscore van 3 punten waarbij de examinerator per onjuiste of niet uitgevoerde handeling 1 punt moet aftrekken.

		Max Score
1b	Communicatie/gesprek openen en afscheid nemen De kandidaat: <ul style="list-style-type: none"> • begroet meneer Boekelo vriendelijk en vraagt zelf waarmee zij/hij meneer Boekelo kan helpen • neemt zelf afscheid door te vragen of alles naar wens is en door bijvoorbeeld een fijne dag te wensen, tot ziens of tot een volgende keer te zeggen 	
	Alleen 1 punt toekennen indien beide beoordelingspunten juist zijn	1

1c	Omgang met meneer Boekelo De kandidaat gaat tijdens de hele opdracht:	
	<ul style="list-style-type: none"> • op een beleefde en professionele manier om met meneer Boekelo (kijkt meneer Boekelo regelmatig aan, spreekt in u-vorm, praat duidelijk, reageert op wat meneer Boekelo zegt) 	2
	<ul style="list-style-type: none"> • bijna steeds op een beleefde en professionele manier om met meneer Boekelo, maar er is één onvolkomenheid 	1
	<ul style="list-style-type: none"> • herhaaldelijk niet op een beleefde en professionele manier om met meneer Boekelo, er is meer dan één onvolkomenheid 	0

1d	De kandidaat: <ul style="list-style-type: none"> • vertelt meneer Boekelo dat hij in de wachtruimte kan wachten • begeleidt meneer Boekelo naar de wachtruimte • vraagt of zij/hij meneer Boekelo ergens mee kan helpen • helpt bij het uittrekken van de jas • hangt de jas op 	
	Per onjuiste of niet uitgevoerde handeling	-1 3

Kandidaten hebben recht op een professionele, objectieve en rechtvaardige beoordeling. Idealiter zou het niet mogen uitmaken wie het examen nakijkt. In de praktijk zijn er echter verschillen tussen correctoren in de strengheid of soepelheid van de correctie. Dat kan ertoe leiden dat een kandidaat met eenzelfde vaardigheidsniveau bij de ene docent gemakkelijker een voldoende behaalt dan bij de andere. Docenten worden geacht het correctievoorschrift goed te bestuderen en strikt na te leven. Uit het Inspectieonderzoek naar de praktijk van de beroepsgerichte examens in het vmbo blijkt dat praktijkdocenten sterk verschillen in de wijze waarop zij de kandidaten beoordelen. Ter illustratie van hoe het niet moet, beschrijven we hieronder vier voorbeelden van beoordelingspraktijken uit het CSPE waarbij docenten de voorschriften voor de beoordeling niet of niet volledig toepasten:

- In het correctievoorschrift zijn de aspecten waar de examinerator bij het beoordelen op moet letten precies aangegeven. Examineratoren worden geacht alle genoemde aspecten in hun beoordeling te betrekken. Ongeveer een derde van de examineratoren uit het Inspectieonderzoek zegt dat niet altijd te doen. Een voorbeeld is afkomstig uit het examen Verwerking agrarische producten. Bij een van de opdrachten moesten de kandidaten friet snijden en vervolgens bakken. Aardappels moeten in de lengterichting worden gesneden en de kleur mag na het bakken natuurlijk niet te bleek of te donker zijn. Sommige examineratoren bleken ten onrechte punten toe te kennen voor aardappels die overdwars waren gesneden of hadden de kleur van de frietjes niet in de beoordeling betrokken.
- In het correctievoorschrift is aangegeven op welke beoordelingsaspecten de examinerator precies moet letten. Examineratoren mogen alleen de genoemde aspecten in hun beoordeling betrekken. Bijna de helft van de examineratoren zegt af en toe tot zeer vaak andere dan de genoemde aspecten in de beoordeling te betrekken. Een voorbeeld uit het examen Handel en verkoop is het ten onrechte toekennen van punten voor handelingen die op school zijn aangeleerd, maar die volgens het correctievoorschrift niet beoordeeld worden, zoals de vraag of een pakje symmetrisch is ingepakt of de manier waarop de kandidaat geld teruggeeft bij een kassaopdracht.
- In het correctievoorschrift is aangegeven hoeveel punten de examinerator voor een zwakke, middelmatige of goede prestatie moet toekennen. Examineratoren mogen kandidaten niet meer punten geven dan is voorgeschreven. Ruim de helft van de examineratoren uit het Inspectieonderzoek zegt dat af en toe tot regelmatig toch te doen. Zo mochten kandidaten Elektrotechniek een installatie onbepaald opnieuw testen als deze niet bleek te werken, zonder dat de examinerator daarvoor punten in mindering bracht.
- Om de kwaliteit van de correctie te waarborgen heeft de wetgever de tweede correctie ingesteld. Bij de praktische opdrachten van de CSPE's stellen de eerste en tweede examinerator de score in onderling overleg vast. Uit het Inspectieonderzoek blijkt dat er lang niet altijd een tweede examinerator in de examenruimte aanwezig is en dat de score vaak niet in onderling overleg wordt vastgesteld.

Kwaliteitsverbetering van het CSPE

In het Onderwijsverslag 2007/2008 (2009) concludeert de Inspectie dat de praktijk van het CSPE 'niet door de beugel kan' (pag. 252). Nog niet alle docenten nemen het praktijkexamen volledig en integraal af, gebruiken geen ander dan het voorgeschreven materiaal, geven de kandidaten geen andere dan als aansporing en bemoediging bedoelde hulp en zorgen ervoor dat kandidaten niet overleggen of elkaar helpen. Een deel van de examinatoren interpreteert de beoordelingsaspecten nog op eigen wijze, betreft eigen beoordelingsaspecten in de beoordeling, let op minder of andere beoordelingsaspecten dan het correctievoorschrift aangeeft en/of beoordeelt te streng of te soepel. Daarbij viel op dat een te soepele beoordeling veel vaker voorkwam dan een te strenge beoordeling. Alhoewel begrijpelijk vanuit het standpunt van de examinerator als degene die de kandidaten heeft opgeleid, staat deze examenpraktijk op gespannen voet met een objectieve beoordeling waarbij kandidaten met dezelfde vaardigheid ongeacht de school waarop zij zitten een gelijke kans hebben om voor het examen te slagen. Dit kan de rechtsgelijkheid van kandidaten en de geloofwaardigheid van het examen in gevaar brengen. Het is om deze reden dat de overheid de richtlijnen voor het CSPE in het schooljaar 2009-2010 heeft aangescherpt.

3.2.5 De correctie van het centraal schriftelijk examen

Bij de centraal schriftelijke examens wordt het examenwerk behalve door de 'eigen' docent ook nagekeken door een corrector van een andere school, de tweede corrector.

De eerste correctie

Correctoren van het centraal schriftelijk examen moeten hun kandidaten beoordelen volgens het correctievoorschrift. Dat bestaat onder meer uit algemene en vakspecifieke regels voor de beoordeling en een antwoordmodel (zie paragraaf 3.2.4). Het beoordelen van de antwoorden van de kandidaten op open vragen is niet altijd eenvoudig. Behalve de algemene en vakspecifieke regels voor de beoordeling moet de corrector het antwoordmodel grondig bestudeerd hebben. Daarnaast moet de docent over voldoende vakinhoudelijke kennis beschikken. Bij de beoordeling van de antwoorden op open vragen zijn beoordelingsverschillen tussen correctoren niet volledig te vermijden. Reeds in 1983 concludeerde Sanders dat 'ondanks het gebruik van antwoordmodellen de verschillen tussen beoordelaars bij bepaalde vakken niet acceptabel zijn' (pag. 170). Bij sommige vakken is het vaak niet mogelijk om alle geheel of gedeeltelijk goede antwoorden die kandidaten kunnen geven in het antwoordmodel op te nemen. Het antwoordmodel volstaat dan noodgedwongen met een beperkt aantal voorbeelden van goede antwoorden. Naarmate het antwoordmodel de verzameling van mogelijke antwoorden minder goed dekt, heeft de corrector minder houvast bij het beoordelen en ontstaan er meer mogelijkheden voor eigen interpretaties en verschillen in soepelheid. Het correctievoorschrift bevat daarom een aantal algemene beoordelingsregels die de corrector moet toepassen als het antwoordmodel geen of weinig sturing biedt. Zo moet de corrector als een vraag gedeeltelijk goed beantwoord is, zelf bepalen welk deel van de te behalen scorepunten hij of zij toekent. En als een antwoord op een open vraag niet in het antwoordmodel voorkomt en dit antwoord op grond van aantoonbare, vakinhoudelijke argumenten als juist of gedeeltelijk juist aangemerkt kan worden, dan moeten scorepunten worden toegekend naar analogie of in de geest van het antwoordmodel. Naast de algemene beoordelingsregels zijn er bij sommige vakken nog vakspecifieke regels. Zo kent het examen geschiedenis havo 2009, aanvullend op de hiervoor genoemde algemene regel, de eis dat 'Vakinhoudelijke argumenten afkomstig moeten zijn uit gezaghebbende, wetenschappelijke publicaties'. Als een antwoord op een open vraag niet in het antwoordmodel voorkomt, mag de corrector dus alleen scorepunten toekennen naar analogie of in de geest van het antwoordmodel als hij of zij de volledige of gedeeltelijke juistheid van het antwoord aan de hand van gezaghebbende, wetenschappelijke publicaties kan aantonen. De beoordelingspraktijk leert dat correctoren in situaties waarin het antwoordmodel niet of onvoldoende

voorziet vaak verschillende scores toekennen, zonder dat altijd onmiddellijk duidelijk is wie het bij het rechte eind heeft. Hoe lastig de beoordeling van open vragen kan zijn, illustreren we aan de hand van twee voorbeelden uit het examen geschiedenis havo 2009. De gepresenteerde antwoorden van kandidaten zijn ‘echte’ antwoorden waarbij de oorspronkelijke formulering onveranderd is overgenomen. Deze twee voorbeelden maken ook duidelijk dat het beoordelen van de antwoorden op open vragen een zware wissel trekt op de vakinhoudelijke deskundigheid van de docent.

Voorbeeld 1

Vraag 20 uit het examen geschiedenis havo 2009 is een zogeheten noemvraag waarbij de kandidaat gevraagd wordt drie voordelen te noemen van de Spaanse verovering van Antwerpen voor de Hollandse nijverheid. De inleiding op de vraag vermeldt dat de Hollandse nijverheid en handel profiteerden van de verovering van Antwerpen door het Spaanse leger. De precieze formulering is als volgt:

De Hollandse nijverheid en handel profiteerden van de verovering van Antwerpen door het Spaanse leger.
3p **20** Noem drie voordelen voor de Hollandse nijverheid van de Spaanse verovering van Antwerpen.

Het antwoordmodel bij vraag 20 is van het type ‘Een voorbeeld van een juist antwoord is’. De corrector kan 0, 1, 2 of 3 punten toekennen. Het antwoordmodel geeft vier voorbeelden van juiste voordelen waarbij de kandidaat per juist voordeel 1 punt krijgt met een maximum van 3 punten. Het antwoordmodel is hieronder onverkort weergegeven:

20 maximumscore 3

Voorbeeld van een juist antwoord is (drie van de volgende):

Door de verovering van Antwerpen door de Spanjaarden:

- kwamen geschoolde arbeidskrachten uit Antwerpen/de Zuidelijke Nederlanden naar Holland.
- verdween de Antwerpse nijverheid als concurrent.
- konden de opstandige gewesten de Schelde afsluiten/de toegang tot Antwerpen blokkeren.
- vestigden zuidelijke immigranten zich in Holland waar zij met hun kennis/bedrijfskapitaal de nijverheid versterkten.

per juist voordeel

1

Overeenkomstig de vraagstelling moet de kandidaat drie voordelen opnoemen van de Spaanse verovering van Antwerpen voor de Hollandse nijverheid. Antwoorden die niet specifiek over nijverheid gaan, moeten volgens de maker van het examen fout gerekend worden. Kandidaten blijken bij deze vraag regelmatig voordelen voor de handel te noemen in plaats van voor de nijverheid, zie de voorbeelden hierna van ‘echte’ antwoorden van kandidaten die deelnamen aan het examen geschiedenis havo 2009.

De Hollandse nijverheid profiteerde van de Spaanse verovering van Antwerpen omdat ze nu minder concurrentie hadden, de handelspositie van Antwerpen nu naar Holland verschoof en dus beter werd en omdat veel mensen naar Holland trokken waardoor het dichter bevolkt werd en door die redenen dus het middelpunt van de handel werd.

Drie voordelen voor de Hollandse nijverheid van de Spaanse verovering van Antwerpen zijn:

De val van Antwerpen. Hierdoor kwam de handel naar Amsterdam.

Rijke kooplieden-regenten vertrokken uit Antwerpen naar Holland en namen daarbij hun kapitaal en kennis mee.

De Schelde werd afgesloten waardoor er geen schepen meer naar Antwerpen konden gaan en de handel nog meer in Holland gevestigd werd.

- De kooplieden en handelaren uit Antwerpen kwamen naar Holland, daardoor nam de kennis toe in Holland.*
- Door de afsluiting van de Schelde vestigde zich in Holland een stapelmarkt.*
- Doordat de Schelde was afgesloten en hier meer handel kwam, namen de arbeidsplaatsen toe. Dat was een voordeel.*

- 1. De Antwerpse haven was niet veilig dus gingen ze naar Hollandse havens voor nijverheid producten.*
- 2. Door de verovering van Antwerpen vluchtte veel mensen naar Holland. Deze mensen gingen verder hun werk doen in Holland dus er was een overvloed aan werknemers.*
- 3. Er waren ook mensen die vluchtte, omdat ze heel rijk waren en veel kennis hadden deze hielpen mee aan de financiering en vernieuwing van de nijverheid.*

Strengere correctoren kennen aan voordelen voor de handel terecht geen punten toe, terwijl hun soepelere collega's dat wel doen. Laatstgenoemden beargumenteren de toegekende punten door erop te wijzen dat nijverheid en handel sterk met elkaar verbonden zijn en dat voordelen voor de nijverheid ook ten goede komen aan de handel. Overigens had dit beoordelingsprobleem wellicht verminderd kunnen worden door het woordje handel uit de inleiding op de vraag te schrappen en aan het antwoordmodel een opmerking toe te voegen dat voordelen voor de handel fout gerekend moeten worden.

Voorbeeld 2

Het tweede voorbeeld betreft opgave 1 uit het examen geschiedenis havo uit 2009. De inleiding op deze vraag vermeldt dat de Franse regering in de coalitieoorlogen de dienstplicht invoerde. De vraag is een zogeheten uitlegvraag en luidt 'Leg uit dat zij hiermee de betrokkenheid van de Franse burgers bij de staat kon vergroten'.

De maximumscore bij deze vraag is 2 punten. Het antwoordmodel beschrijft een voorbeeld van een juist antwoord dat uit twee varianten bestaat. Volgens de eerste variant is de juiste uitleg,

van het gegeven dat de Franse regering door de dienstplicht in te voeren de betrokkenheid van de burgers bij de staat kon vergroten, dat een groot aantal (jonge) mannen meer in aanraking kwam met de idealen van de Franse revolutie of het Franse nationalisme. Bij de tweede variant is de juiste uitleg dat een groot aantal (jonge) mannen meer onder invloed van politieke commissarissen kwam. Deze tweede variant is op verzoek van het CvTE aan het antwoordmodel toegevoegd. Het antwoordmodel is hieronder integraal opgenomen.

1 maximumscore 2

Voorbeeld van een juist antwoord is:

Door de invoering van de dienstplicht kwam een groot aantal (jonge) mannen meer in aanraking met de idealen van de Revolutie/het Franse nationalisme / onder invloed van politieke commissarissen.

Hieronder staan vijf voorbeelden van ‘echte’ antwoorden van kandidaten die in de praktijk aanleiding gaven tot grote scoreverschillen tussen eerste correctoren.

De dienstplicht houdt in dat alle mannen tussen de 18 en 25 jaar (of ouder) in het leger moesten dienen. Doordat deze mannen het leger in moesten kregen ze meer mee van de oorlogen in de tijd & leerden ze vechten voor hun vaderland. De betrokkenheid van de burgers werd vergroot doordat mannen hun vrouwen en kinderen en ouders achter lieten. Deze leefden met de oorlog mee omdat een geliefde van hen het leger in ging en oorlog ging voeren.

De dienstplicht was ingesteld voor mannen van 18-25 jaar. De betrokkenheid van de Franse burgers wordt hierdoor vergroot doordat de burgers nu zelf ten strijd moeten gaan.

Door de dienstplicht was iedereen vanaf een bepaalde leeftijd verplicht mee te vechten in de oorlog. Hierbij werden dus de burgers betrokken bij de beslissingen die de staat nam (bij deze dus een oorlog). Bijna alle mannen van een bepaalde leeftijd (meestal vanaf 18) werden opgeroepen dat ze moesten vechten en wel elke vrouw had een vader, broer, man of zoon die ging vechten.

De burger werd betrokkener bij de staat omdat er de kans was dat die in het leger moest en dus afhankelijk was van de beslissingen van de staat op bv. militair gebied. Ook kende iedereen wel iemand die in het leger zat of zou gaan. Ze trokken het lot van die mensen aan en gingen dus ook de beslissingen van de staat volgen, hierdoor werden de burgers meer betrokken tot de staat.

Vrijwel alle mannen in Frankrijk die oud genoeg waren gingen in dienstplicht. Hierdoor was een enorm deel van de bevolking direct bij oorlogen betrokken.

Waarom zijn deze vijf antwoorden zo verschillend beoordeeld?

- Een eerste reden is dat de antwoorden op geen enkele wijze zijn terug te vinden in het antwoordmodel. Zo verwijst geen van de kandidaten in zijn of haar uitleg naar de idealen van de Franse revolutie, het Franse nationalisme of de invloed van politieke commissarissen die in het antwoordmodel genoemd worden. Voor zover kandidaten andere historische verklaringen geven, zijn deze meestal niet (door docenten) terug te vinden in gezaghebbende wetenschappelijke publicaties.
- Een tweede reden heeft te maken met de aard van het antwoordmodel. Het bestaat uit een beoordelingschaal met de toegestane scores 0, 1 en 2. Het antwoordmodel geeft alleen een voorbeeld van een volledig juist antwoord. Omdat alleen de maximumscore aan de hand van een voorbeeld omschreven is, biedt het antwoordmodel de corrector weinig houvast bij het toekennen van 1 of 0 punten. Het herkennen en waarderen van half en geheel foute antwoorden wordt met andere woorden aan de vakinhoudelijke deskundigheid van de corrector overgelaten.
- Een derde reden is dat kandidaten vaak zuiver psychologische verklaringen geven die op zich niet onlogisch zijn, maar die niet door gezaghebbende wetenschappelijke publicaties gestaafd worden (en dus fout gerekend zouden moeten worden). Zo verwijzen veel kandidaten ernaar dat de Franse overheid de burgers bij de staat wist te betrekken doordat iedere soldaat wel een familielid heeft dat met hem meeleeft. Verschillen in soepelheid ontstaan waar strenge correctoren een verwijzing naar termen uit het antwoordmodel of de vakliteratuur eisen en soepele correctoren genoeg nemen met een niet-historisch antwoord.

Tip

Verschillende correctoren blijken de algemene regels voor de beoordeling en het antwoordmodel verschillend te interpreteren. Het is dan ook belangrijk dat docenten niet in hun eentje met de correctie van een nieuw examen aan de slag gaan. Alvorens met de correctie te beginnen, is het verstandig het antwoordmodel samen met de collega's goed door te nemen, de eerste antwoorden van de kandidaten gezamenlijk te beoordelen en eventuele verschillen in interpretatie uit te discussiëren. Ook is het zinvol deel te nemen aan de bespreking van de opgaven en het antwoordmodel die veel vakinhoudelijke verenigingen direct na de afname van het centraal schriftelijk examen organiseren.

De tweede correctie

Bij de centraal schriftelijke examens wordt het examenwerk behalve door de 'eigen' docent een tweede keer nagekeken door een corrector van een andere school. Na het examen nagekeken te hebben, stuurt de eerste corrector het examenwerk op naar de tweede corrector. Vrijwel altijd bevat het examenwerk de annotaties van de eerste corrector, dit wil zeggen de toegekende punten per vraag en soms ook een aanduiding van eventuele onvolkomenheden in het antwoord. Vaak beschikt de tweede corrector behalve over de totaalscore van de eerste corrector ook over het cijfer van de kandidaat op het schoolexamen dat voor de afname van het centraal examen bekend moet zijn. Als de tweede corrector 'klaar' is, moet hij of zij contact opnemen met de eerste corrector. De eerste en tweede corrector moeten de uiteindelijke scores in onderling overleg vaststellen. Het overleg tussen eerste en tweede corrector vindt meestal plaats via de telefoon. Wat kunnen eerste en tweede corrector doen als zij van mening verschillen over de toegekende scores? Als de tweede corrector vindt dat er sprake is van grote onzorgvuldigheid, aperte fouten of verkeerde interpretatie van de correctievoorschriften dient hij of zij er eerst in overleg met de eerste corrector uit te komen. Als dat niet lukt, kan het scoreverschil worden gemiddeld. De Inspectie van het Onderwijs ziet middelen echter als een zwaktebod dat niet past bij professioneel handelende vakdeskundigen. Als de eerste en/of tweede corrector niet willen middelen, kan de tweede corrector zich tegenwoordig melden bij

zijn eigen bevoegd gezag die dan contact kan opnemen met het bevoegd gezag van de eerste corrector. Als beiden het niet eens kunnen worden, melden zij dit bij de Inspectie en deze kan vanuit haar toezichthoudende taak bij de examens optreden. Dit kan betekenen dat de Inspectie besluit tot de inzet van een derde onafhankelijke corrector. Uiteraard kan deze procedure ook worden toegepast bij klachten over het werk van de tweede corrector.

Volgens het Examenbesluit moeten tweede correctoren het werk van de kandidaten integraal nakijken, alsof het gaat om een eerste correctie. Dit wil zeggen dat de tweede corrector alle werken nakijkt en per kandidaat het volledige examenwerk met alle opgaven. De Inspectie constateert regelmatig dat lang niet alle correctoren het examenwerk zorgvuldig genoeg nakijken. Een op de drie tweede correctoren is van mening dat de eerste corrector tekortschiet. Veertig procent van de tweede correctoren ontdekt dat er fouten zijn gemaakt tijdens de eerste correctieronde. Maar ook de tweede correctoren laten steken vallen, ook zij kijken het werk niet zorgvuldig genoeg na. Als een eerste corrector wordt gecorrigeerd door een tweede corrector die onvoldoende nauwkeurig nakijkt dan is de kans groot dat het examenwerk niet naar behoren wordt beoordeeld. De kandidaten krijgen in dat geval niet het cijfer dat ze verdienen. De Inspectie concludeert dat 'gezien het percentage correctoren dat fouten maakt en onvoldoende nauwkeurig alle opgaven controleert, kan het niet anders dan dat een aantal kandidaten een niet correcte eindscore krijgt'. In 2008 onderstreepte de toenmalige staatssecretaris nogmaals het belang van een volledige tweede correctie met als argument dat de leerling recht heeft op een objectieve en professionele correctie. Tegenwoordig is volledige en integrale tweede correctie dan ook wettelijk verplicht.

Voor lang niet iedereen staat de noodzaak van integrale tweede correctie echter vast. De VO-raad is bijvoorbeeld van mening dat docenten de examens tijdens de eerste correctieronde goed nakijken en dat het zeer goed mogelijk is dat een goed uitgevoerde steekproef voldoende garantie biedt. Ook volgens het Platform Vakinhoudelijke Verenigingen Voortgezet Onderwijs (VVVO) is steekproefsgewijze tweede correctie in de gegeven situatie de beste oplossing. Op basis van eigen onderzoek wijst de VVVO op de hoge taakbelasting die integrale tweede correctie voor docenten met zich meebrengt; docenten blijken de integrale tweede correctie alleen ten koste van hun eigen vrije tijd en nachtrust te kunnen uitvoeren. Als de overheid aan een integrale tweede correctie wil vasthouden, aldus het Platform VVVO, zal de voor de correctie benodigde tijd 'zichtbaar en geormerkt in de taakbelasting van de betrokken docenten opgenomen moeten worden'.

Hoe 'erg' zijn verschillen tussen correctoren?

Onderzoek heeft laten zien dat onafhankelijke correctoren gemiddeld beduidend lagere scores toekennen dan de 'eigen' docent. Daarnaast blijken eerste correctoren vaak sterk te verschillen in de strengheid of soepelheid van de correctie. Zo vond Sanders al in 1983 dat 'het voor leerlingen bijzonder veel kan uitmaken wanneer zij niet door een milde beoordelaar, maar door een strenge beoordelaar beoordeeld worden; jammer genoeg (of gelukkig maar) weten zij, noch hun beoordelaars, niet of zij bevoordeeld of benadeeld worden' (pag. 171). Hij concludeerde dat 'de huidige procedure, waarbij de eigen docent door een tweede corrector gecontroleerd wordt, geen oplossing biedt voor mogelijke bevoordeling of benadeling' (pag. 171).

Verbetering van de correctie van het centraal schriftelijk examen

Oorzaken van verschillen tussen correctoren kunnen liggen in het examen (bijv. het vak, de vraag, het antwoordmodel en de lengte en complexiteit van de antwoorden), de corrector (bijv. vakkennis, opleiding, training, neiging tot inaccuraat, soepel en opportunistisch beoordelen) en de context waarin de correctie plaatsvindt. Hieronder noemen we enkele verklaringen die een rol zouden kunnen spelen. We maken daarbij een onderscheid in het examen, de docent en de context.

Het examen

- Verschillen tussen correctoren kunnen mogelijk gedeeltelijk verklaard worden vanuit kenmerken van het examen. De ervaring leert dat de verschillen tussen eerste correctoren toenemen naarmate de vraagstelling meer open is en het antwoordmodel minder sturing biedt. Een mogelijk aandachtspunt is aanpassing van de antwoordmodellen bij de open vragen, zodat eerste correctoren minder vaak zelfstandig een beroep hoeven te doen op de algemene en vakspecifieke scoringsregels. Bij dit laatste kan men dan vooral denken aan de regel voor het zelfstandig toekennen van punten aan antwoorden die niet in het antwoordmodel voorkomen en het zelfstandig toekennen van punten aan gedeeltelijk juiste antwoorden.

De corrector

- Een toegeeflijke beoordeling is deels begrijpelijk als men bedenkt dat de eerste corrector tevens degene is die de kandidaat heeft opgeleid (en dat tweede correctoren de kandidaten van de eerste corrector doorgaans niet veel strenger zullen beoordelen dan zij hun 'eigen' kandidaten hebben beoordeeld). Docenten kiezen het beroep van leraar omdat zij jongeren willen helpen. Zij zijn begaan met het lot van hun leerlingen en een deel van de docenten neigt ertoe de prestaties van de kandidaten hoger te waarderen dan gerechtvaardigd is. Hoewel niet expliciet onderzocht, lijkt de persoonlijkheid van de corrector hierbij een belangrijke rol te spelen. Correctoren verschillen in de mate van slordigheid, de neiging om eigen kandidaten het voordeel van de twijfel te geven en de wil om de verleiding van opportunistisch-strategisch beoordelingsgedrag te weerstaan.
- Docenten zijn meestal niet geschoold in het examineren en beoordelen van kandidaten. Zij kunnen zich bijvoorbeeld nog niet als examinator laten certificeren.

De context waarbinnen de correctie plaatsvindt

- Het openbaar maken van examengegevens, het beoordelen van de onderwijskwaliteit van scholen (mede) op basis van hun examenprestaties en het gebruik van examencijfers voor personeelsbeoordeling kan de druk op hoge examenprestaties en de neiging tot toegeeflijk beoordelen verhogen.
- Het Platform VVVO heeft erop gewezen dat veel docenten het examenwerk in een korte periode onder tijdsdruk moeten nakijken. De als niet optimaal ervaren randvoorwaarden dragen er mogelijk toe bij dat de tweede correctie lang niet altijd integraal wordt uitgevoerd.
- De positie van de tweede corrector is relatief zwak in vergelijking met die van de eerste corrector. Zo moeten tweede correctoren hun gelijk zien te halen bij de eerste corrector in plaats van andersom. Daarnaast moet de tweede correctie vaak onder een nog grotere tijdsdruk worden uitgevoerd dan de eerste correctie.

Om de correctie van het CSE op alle scholen in overeenstemming te brengen met de wettelijke vereisten en kwaliteitscriteria zijn maatregelen voorgesteld. In het advies 'Examinering: draagvlak en toegankelijkheid' (2006) heeft de Onderwijsraad voorgesteld om de correctievolgorde om te draaien. Eerst kijkt de docent van de andere school het examen na en komt met een scorevoorstel. Pas daarna is de 'eigen' docent aan de beurt die moet aangeven in hoeverre hij of zij het met het voorstel eens is. Als de correctievolgorde wordt omgedraaid, zal de eigen docent, in de rol van tweede corrector, het werk van de eerste corrector zeer nauwkeurig controleren. Het betreft immers het examenwerk van de eigen leerlingen. In hetzelfde advies vraagt de Onderwijsraad de scholen om in het taak- en vergoedingenbeleid rekening te houden met het werk dat de correctie met zich meebrengt.

Een alternatief voor het omkeren van de correctievolgorde is dat de examenwerken na gescand te zijn digitaal naar de eerste en tweede correctoren worden verzonden. De eerste en tweede correctie kunnen dan 'tegelijktijd' plaatsvinden. De tweede corrector hoeft dan niet meer te

wachten tot de 'eigen' docent klaar is. Overigens kan men zich afvragen hoe groot deze tijdswinst in de praktijk zal zijn. Vrijwel alle eerste correctoren zijn immers ook tweede corrector. Wel voorkomt het gelijktijdig nakijken de tijdsdruk die optreedt als de eerste corrector het werk pas zeer laat naar de tweede corrector opstuurt. Een tweede voordeel is dat alle correctoren het examenwerk integraal moeten nakijken, dus ook de tweede correctoren. Ook op dit voordeel valt wat af te dingen. De tweede correctoren moet het examenwerk namelijk ook integraal nakijken als de 'eigen' docent het werk zeer zorgvuldig en objectief heeft nagekeken en een steekproefsgewijze controle voldoende garantie zou hebben geboden. Omdat integrale tweede correctie tegenwoordig wettelijk verplicht is, is dit tegenargument formeel gezien niet erg sterk en alleen van praktische importantie. Een derde voordeel van gelijktijdige eerste en tweede correctie is dat alle correctoren een blanco examen onder ogen krijgen. Daardoor wordt de tweede corrector niet beïnvloed door de scores en aantekeningen van eerste corrector. Over de vraag of het wenselijk is dat de tweede corrector kennis neemt van de annotaties van de eerste corrector zijn de meningen echter verdeeld. De aantekeningen van de eerste corrector op het examenwerk laten zien hoe de score tot stand is gekomen. Deze informatie biedt de tweede corrector steun bij het beoordelen; bovendien kan het goed van pas komen in het overleg tussen eerste en tweede corrector waarin de uiteindelijke score wordt vastgesteld. Het lijkt dan ook begrijpelijk dat tweede correctoren de annotaties van de eerste corrector over het algemeen sterk op prijs stellen.

3.3 Het schoolexamen

Het SE gaat in de tijd vooraf aan het CE. De school moet de SE-cijfers vaststellen voordat het CE begint. Het SE biedt scholen een grote mate van vrijheid om te bepalen wanneer zij welke vakken toetsen, welke toetsvormen zij daarbij gebruiken en hoe zij de prestaties van de kandidaten nakijken en beoordelen. De docenten moeten zich daarbij wel houden aan het globale examenprogramma. Binnen dit algemene kader kunnen de scholen de inhoud van het SE zelf invullen. De toetsen, werkstukken en praktische opdrachten voor het SE maakt of kiest de school geheel zelf. Ook de afname en beoordeling zijn volledig in handen van de school. Hiermee beoogt de overheid recht te doen aan de professionaliteit van docenten en kan de school aansluiten bij de gekozen didactische aanpak. Anders dan bij de centraal schriftelijke examens wordt het examenwerk meestal alleen nagekeken door de docent van de eigen school.

3.3.1 De inhoud en toetsvormen van het schoolexamen

Het SE toetst ten minste dat deel van het examenprogramma dat niet in het CE geëxamineerd wordt. Zoals uiteengezet in hoofdstuk 2, zijn de examenprogramma's in 2007 geglobaliseerd. Dit wil zeggen dat er minder onderwijsdoelen zijn die in nog globalere termen geformuleerd zijn. Dat geeft de scholen die dat wensen meer ruimte voor eigen invulling van de leerstof. Verder zijn de zogeheten vormvoorschriften voor de toetsing uit de wet verdwenen. Daardoor kunnen de scholen tegenwoordig zelf bepalen met welke toetsvormen ze de kennis en vaardigheden van de leerlingen toetsen. Een praktijkdocent uit het vmbo kan er bijvoorbeeld voor kiezen om geen praktische opdrachten in het SE op te nemen (al moeten er in het SE nog wel vaardigheden getoetst worden). Tot slot kunnen scholen tegenwoordig naar eigen goeddunken relevante onderwerpen aan het SE toevoegen. Zo mogen verplichte examenonderdelen van het CE ook in het SE getoetst worden. Maar de school kan ook geheel eigen programma's of onderdelen aan het SE toevoegen. Het onderwijzen en examineren van niet-verplichte onderdelen valt echter af te raden als dat zou leiden tot overlappendheid van het onderwijsprogramma of tijdgebrek en stress tijdens het SE.

Hieronder staan drie tips voor de constructie van opgaven en opdrachten voor het SE. De opgaven en opdrachten van het SE moeten aan verschillende kwaliteitseisen voldoen.

Het gaat niet alleen om de vakinhoudelijke correctheid, maar ook om een juiste moeilijkheidsgraad, helder taalgebruik, duidelijke instructie, functionele vormgeving en objectieve correctie en beoordeling.

Tip

Uit onderzoek van de Inspectie is naar voren gekomen dat de kwaliteit van het toetsmateriaal dat docenten zelf maken nogal eens te wensen overlaat. Voor vuistregels en tips voor het maken van gesloten en open vragen wordt verwezen naar respectievelijk hoofdstuk 6 en 7 van het algemene deel van Toetsen op School.

Tip

Voor het beoordelen van praktische vaardigheden is de beoordelingsschaal het meest geëigende middel. Zie voor vuistregels en tips voor het ontwikkelen van beoordelingsschalen hoofdstuk 8 van het algemene deel van Toetsen op School.

Tip

Uit onderzoek van de Inspectie blijkt dat de schoolexamens veel toetsmateriaal en correctie-modellen van educatieve uitgevers bevatten. Omdat dit materiaal niet geheim is, kan het aanleiding geven tot examenfraude, bijvoorbeeld door huiswerkinstituten. Voorkom dat opgaven uit het leerboek het schoolexamen te sterk bepalen. Hanteer daarom een bovengrens voor de hoeveelheid opgaven uit een bundel, zowel in totaal als per schoolexamen.

Producten voor het schoolexamen

Uit onderzoek van de Inspectie blijkt dat extern ontwikkeld toetsmateriaal voor het SE vaak van hogere kwaliteit is dan het toetsmateriaal dat docenten geheel zelfstandig ontwikkelen. Dat geldt onder meer voor moeilijk toetsbare vakonderdelen zoals kijk- en luistervaardigheid Nederlands en de moderne vreemde talen (zie ook hoofdstuk 1 van deze publicatie). Voor een handig overzicht van alle Cito-producten voor het schoolexamen, zie www.cito.nl > Voortgezet onderwijs > Schoolexamens.

Workshop kwaliteitsmonitor schoolexamens

Kwaliteitszorg en schoolzelfevaluatie worden steeds belangrijker. Vooral de kwaliteit van het schoolexamen staat daarbij centraal. Cito heeft op verzoek van de overheid een instrument ontwikkeld waarmee scholen de kwaliteit van de schoolexamens in kaart kunnen brengen: de kwaliteitsmonitor schoolexamens. In de workshop kwaliteitsmonitor schoolexamens krijgt men deskundige begeleiding bij het gebruik van de kwaliteitsmonitor en het vaststellen en zo nodig verbeteren van de kwaliteit van het schoolexamen. De workshop wordt gegeven bij Cito dan wel ter plekke op de school.

Website

Ook de website www.schoolexamensvo.nl biedt docenten de mogelijkheid om de kwaliteit van schoolexamens te monitoren en waar mogelijk te verbeteren. Het is de plek om informatie en hulpmiddelen te vinden. Het is ook dé ontmoetingsplek voor docenten en schoolleiders waar men elkaar met vragen, ervaringen en opmerkingen over schoolexamens stimuleert.

3.3.2 Examenreglement, PTA en andere documenten

Voor docenten belangrijke documenten zijn: het examenreglement, het programma van toetsing en afsluiting (PTA), de handreikingen van SLO en de checklists van de VO-raad.

Examenreglement en PTA

Scholen moeten het examenreglement en het PTA jaarlijks vóór 1 oktober aan de Inspectie toezenden en aan leerlingen verstrekken. Het examenreglement bevat onder meer de regels met betrekking tot de organisatie van het eindexamen, de gang van zaken tijdens het eindexamen, de herkansing van het schoolexamen en maatregelen die scholen bij ziekte of onregelmatigheden moeten nemen.

Het PTA is eveneens een schooldocument. Het is een belangrijke informatiebron voor docenten, ouders, leerlingen en de Inspectie. Het PTA bevat voor elk examenvak onder meer de volgende onderdelen:

- welke onderdelen uit het examenprogramma op welke momenten in het SE worden getoetst;
- wat de leerstof/inhoud is per toets/opdracht;
- in welke vorm deze leerstof/inhoud wordt getoetst, bijvoorbeeld: schriftelijk, mondeling of praktische opdracht;
- wat de herkansingsmogelijkheden zijn;
- wat het gewicht van de toets/opdracht is;
- hoe het SE-cijfer tot stand komt.

Handreikingen SLO

De handreikingen van SLO zijn bedoeld om docenten ideeën over de inhoud en de toetsvormen van het SE aan te reiken. De documenten vormen een belangrijke inspiratiebron voor degenen die de nieuwe mogelijkheden van het geglobaliseerde examenprogramma optimaal willen benutten. De handreikingen voor de verschillende examenvakken en onderwijssoorten zijn te downloaden van de website van SLO, www.slo.nl.

Checklists VO-raad

De VO-raad heeft samen met de AOC Raad checklists ontwikkeld, zie www.vo-raad.nl > Thema's > Examens > Protocollen en checklisten examens, om scholen te helpen bij de voorbereiding, afname en beoordeling van het schoolexamen. Er zijn checklists voor het examenreglement, het programma van toetsing en afsluiting, de kwaliteit van het schoolexamen, verandering of te laat komen van kandidaten, fraude door kandidaten en de afronding van de schoolexamencijfers. Met behulp van de checklists kunnen scholen controleren of ze voldoen aan criteria die noodzakelijk of gewenst zijn voor een zorgvuldige uitvoering van het schoolexamen. Elke checklist is voorzien van een toelichting en van een verwijzing naar de actuele regelgeving. In een aantal gevallen is ook de tekst van de betreffende regelgeving opgenomen.

3.3.3 Kwaliteitseisen voor het schoolexamen

Ten behoeve van de analyse van de kwaliteit van de schoolexamens heeft Cito samen met de Inspectie een conceptueel kader opgesteld. Het zogeheten toetsingskader bevat onder meer een beschrijving van de standaarden en indicatoren aan de hand waarvan men de algemene kwaliteit van het SE en de specifieke kwaliteit per vakonderdeel kan beoordelen. De zes standaarden en de negentien indicatoren zijn samen met een korte omschrijving weergegeven in het overzicht hierna.

Standaard	Indicator	Omschrijving
1 Inhoud	1 Verplichte onderdelen SE	Het SE voldoet aan de inhoudelijke specificaties uit het eindexamenprogramma
2 Niveau	2 Conform niveauspecificaties	Het SE voldoet aan de niveauspecificaties uit het eindexamenprogramma
	3 Afnamemoment gerelateerd aan niveau	Het toetsmateriaal wordt afgenomen op een moment dat de kandidaten het vereiste niveau redelijkerwijs bereikt kunnen hebben
	4 Moeilijkheid toetsen en opdrachten	De toetsen en opdrachten zijn noch te moeilijk noch te gemakkelijk (in vergelijking met het door de eindtermen geïmpliceerde niveau en de moeilijkheid van de toetsen en opdrachten uit het centraal examen)
	5 Soepelheid/strengheid beoordelingsmiddelen	De beoordelingsmiddelen zijn noch te streng noch te soepel
	6 Becijfering kandidaten	De toegekende cijfers zijn noch te hoog noch te laag
3 Toetsconstructie	7 Conform producteisen	Het SE is zorgvuldig samengesteld en voldoet aan relevante producteisen
	8 Vergelijkbaarheid opgaven	Als met toetsvarianten wordt gewerkt, is de moeilijkheidsgraad ervan vergelijkbaar
	9 Verplichte toetsinhoud in overleg	De vakcollega's stellen de inhoud van het SE in overleg vast
4 Beoordeling	10 Beoordelingsaspecten relevant	Alle beoordelingsaspecten zijn relevant
	11 Beoordelingsaspecten volledig	Alle relevante beoordelingsaspecten worden in de beoordeling betrokken
	12 Gebruik antwoord- c.q. beoordelingsmodel	De beoordeling vindt plaats met een antwoord- c.q. beoordelingsmodel met vermelding van de beoordelingscriteria en het gewicht ervan
	13 Cijfers voor toetsen en opdrachten	Voor de toetsen en handlingsdelen worden cijfers respectievelijk de kwalificaties 'naar behoren' of 'onvoldoende' toegekend
	14 Beoordeling in overleg	De beoordeling geschiedt door ten minste twee docenten
5 Normering	15 Scorepunten en/of cesuur vooraf vastgelegd	De scorepunten en indien van toepassing ook de cesuur worden vóór de afname van het toetsmateriaal schriftelijk vastgesteld
	16 Scorepunten vroegtijdig meegedeeld	De scorepunten worden vóór de afname van het toetsmateriaal aan de kandidaten kenbaar gemaakt
	17 Bijstelling scorepunten/ cesuur beargumenteerd	De scorepunten en indien van toepassing de cesuur worden bijgesteld op basis van argumenten die vóór of tijdens het SE zijn/worden vastgelegd
	18 Normering in overleg	De vakcollega's bepalen de normering in onderling overleg
6 Weging	19 Gewicht	Het gewicht bij de bepaling van het SE-cijfer is voor elke toets aangegeven

Hieronder lichten we de indicatoren aan de hand van het schoolexamen Engels havo kort toe.

1 Verplichte onderdelen SE

Het SE moet voldoen aan de inhoudelijke specificaties uit het eindexamenprogramma. Voor het vak Engels moet het examendossier in havo 5 bijvoorbeeld toetsen bevatten voor kijk- en luistervaardigheid, gespreksvaardigheid en schrijfvaardigheid.

2 Conform niveauspecificaties

Het SE moet voldoen aan de niveauspecificaties uit het eindexamenprogramma. Voor Engels havo moeten in het SE ten minste de domeinen aan bod komen die in het Centraal Examen niet geëxamineerd worden, dat wil zeggen kijk-/luistervaardigheid, gespreksvaardigheid en schrijfvaardigheid. De eindtermen en specificaties voor het schoolexamen Engels havo zijn opgenomen in het Examenprogramma moderne vreemde talen, zie www.examenblad.nl. Uit bijlage 1 van dit document kan men bijvoorbeeld opmaken dat kandidaten in het subdomein Gesprekken voeren, moeten voldoen aan de volgende eisen: a) adequaat moeten kunnen reageren in sociale contacten met doeltaalgebruikers, b) informatie kunnen vragen en verstrekken, c) uitdrukking kunnen geven aan gevoelens, d) zaken of personen kunnen beschrijven en standpunten en argumenten kunnen verwoorden en e) strategieën kunnen toepassen om een gesprek voortgang te doen vinden. Voor de handreiking voor het schoolexamen Engels havo verwijzen wij naar www.slo.nl > VO > Tweede fase > (School) examen > Handreikingen.

3 Afnamemoment gerelateerd aan niveau

Het toetsmateriaal moet worden afgenomen op een moment dat de kandidaten het vereiste niveau redelijkerwijs bereikt kunnen hebben. Als een school bijvoorbeeld een vakonderdeel al in de vierde klas afsluit, is het de vraag of alle kandidaten voldoende de gelegenheid hebben gehad om zich het vereiste niveau eigen te maken.

4 Moeilijkheid toetsen en opdrachten

De toetsen en opdrachten mogen noch te moeilijk noch te gemakkelijk zijn. Als het toetsmateriaal stevast te moeilijk of te gemakkelijk is en docenten dat niet compenseren door een soepele respectievelijk strenge beoordeling (zie indicator 5 en 6), kunnen kandidaten in vergelijking met leeftijdsgenoten op andere scholen worden benadeeld respectievelijk bevoordeeld.

5 Soepelheid/strengheid beoordelingsmiddelen

De beoordelingsmiddelen moeten noch te streng noch te soepel zijn. Onder beoordelingsmiddelen verstaan we hier bijvoorbeeld een antwoordmodel voor het scoren van open vragen of een beoordelingsschaal voor het beoordelen van gespreksvaardigheid. De scorings- en beoordelingsmiddelen laten zien hoe de school scores (punten) toekent aan zwakke, middelmatige en goede prestaties. Verschillen tussen scholen in de strengheid van de scorings- en beoordelingsmiddelen kunnen ertoe leiden dat een kandidaat met een vergelijkbaar prestatieniveau op de ene school een grotere kans heeft om te slagen dan op de andere school.

6 Becijfering kandidaten

De toegekende cijfers moeten noch te hoog noch te laag zijn. Samen met de toetsen en opdrachten en de scorings- en beoordelingsmiddelen geven de scores en cijfers die de examinerator op de leerlingenwerken heeft genoteerd inzicht in de billijkheid van de beoordeling. Bij het toekennen van cijfers kunnen docenten compenseren voor te moeilijke of te gemakkelijke toetsen of voor te strenge of te soepele beoordelingsmiddelen. Examinatoren kunnen kandidaten uiteindelijk strenger of soepeler beoordelen dan men op grond van de

toetsen en opdrachten en de gehanteerde beoordelingsmiddelen zou verwachten. Dit is alleen na te gaan aan de hand van de uiteindelijk toegekende scores en cijfers.

7 Conform producteisen

Het SE moet zorgvuldig zijn samengesteld en voldoen aan relevante producteisen. Het gaat hier om elementaire eisen van algemene kwaliteit, validiteit, betrouwbaarheid en efficiëntie. Bij de algemene kwaliteit gaat het om eisen zoals de aansluiting bij de eindtermen, de eenduidigheid van de vragen en opdrachten en de vakinhoudelijke correctheid. Ten aanzien van de betrouwbaarheid van het SE moeten de toetsen voldoende lang zijn om een betrouwbare toetsing mogelijk te maken. Een schrijfopdracht van een kwartier kan bijvoorbeeld niet bijzonder betrouwbaar zijn. De producteis van efficiëntie verlangt bijvoorbeeld dat de toetsen niet langer zijn dan voor een betrouwbare meting strikt noodzakelijk is.

8 Vergelijkbaarheid opgaven

Als met toetsvarianten wordt gewerkt, bijvoorbeeld bij inhaaltoetsen of herkansingen, moet de moeilijkheidsgraad ervan vergelijkbaar zijn. Scholen werken met toetsvarianten als zij grote aantallen kandidaten moeten examineren en er vanwege de geheimhouding niet steeds dezelfde toets of opdracht kan worden gebruikt. Ook bij de herkansing dienen dezelfde toetsen en opdrachten te worden gebruikt. Omwille van een rechtvaardige examinering is het gewenst dat de varianten van een vergelijkbare moeilijkheidsgraad zijn.

9 Verplichte toetsinhoud in overleg

De vakcollega's moeten de inhoud van het SE in overleg vaststellen. In de situatie dat aan een kleine school voor het desbetreffende examenvak slechts één vakdocent verbonden is, moet deze de inhoud van het SE vaststellen met ten minste één vakcollega uit de onderbouw, met een vakdocent van een verwant vak of met een vakdocent van een andere school. Dit betekent overigens niet dat de collega zelf actief betrokken moet zijn bij het ontwikkelen van het examen. Hij of zij moet echter wel de gelegenheid hebben gehad om het conceptmateriaal van commentaar te voorzien.

10 Beoordelingsaspecten relevant

Alle beoordelingsaspecten moeten relevant zijn. Bij de beoordeling van de kandidaten mogen alleen aspecten worden betrokken die relevant zijn voor de getoetste kennis en vaardigheden. Bij gespreksvaardigheid ligt het voor de hand te letten op de inhoud van het taalgebruik, de vlotheid en de uitspraak. Bij schrijfvaardigheid zijn het de inhoud, opbouw en correctheid van het taalgebruik. Voor luistervaardigheid zou de stel- en spelvaardigheid van de kandidaat bijvoorbeeld geen rol mogen spelen.

11 Beoordelingsaspecten volledig

Alle relevante beoordelingsaspecten moeten in de beoordeling worden betrokken. Zo moeten er bij schrijfvaardigheid Engels ook andere aspecten worden beoordeeld dan uitsluitend grammatica (zoals doel- en publiekgerichtheid).

12 Gebruik antwoord- c.q. beoordelingsmodel

De beoordeling moet plaatsvinden met een antwoord- c.q. beoordelingsmodel met vermelding van de beoordelingscriteria en het gewicht ervan. Als de school bijvoorbeeld de luister- of schrijftoetsen Engels van Cito gebruikt, kan men ervan uitgaan dat voor deze vakonderdelen aan deze eis voldaan is.

13 Cijfers voor toetsen en opdrachten

Voor de toetsen Engels havo 5 moeten cijfers worden gegeven. Volgens de oude

exameneisen mochten voor het handelingsdeel alleen de kwalificaties ‘naar behoren’ of ‘onvoldoende’ worden toegekend. In de nieuwe examenprogramma’s komen de handelingsdelen niet meer voor.

14 Beoordeling in overleg

De beoordeling geschiedt door minimaal twee docenten. Bij de mondelinge toetsing van gespreksvaardigheid Engels kan men bijvoorbeeld eisen dat de score in onderling overleg tussen twee examinatoren wordt vastgesteld. Is er in de bovenbouw slechts één vakdocent Engels werkzaam, dan moet deze een vakcollega uit de onderbouw inschakelen of bijvoorbeeld een collega Duits of Frans of een docent Engels van een andere school.

15 Scorepunten en/of cesuur vooraf vastgelegd

De scorepunten en indien van toepassing ook de cesuur dienen vóór de afname van het toetsmateriaal schriftelijk te worden vastgelegd. Bij schrijfvaardigheid Engels kan er bijvoorbeeld van worden uitgegaan dat aan deze indicator is voldaan als in de toets is aangegeven hoeveel punten de kandidaat met de hele opgave of met elke deelopgave kan verdienen.

16 Scorepunten vroegtijdig meegedeeld

De scorepunten worden vóór of tijdens de afname van het toetsmateriaal aan de kandidaten kenbaar gemaakt. Bij schrijfvaardigheid Engels moet de kandidaat bijvoorbeeld vroegtijdig geïnformeerd zijn over de te behalen maximumscore en het gewicht van de beoordelingsaspecten in het cijfer en de cesuur.

17 Bijstelling scorepunten/cesuur beargumenteerd

De scorepunten en indien van toepassing de cesuur dienen te worden bijgesteld op basis van argumenten die vóór of tijdens het SE zijn/worden vastgelegd.

18 Normering in overleg

De vakcollega’s dienen de normering in onderling overleg te bepalen. Is er in de bovenbouw slechts één vakdocent Engels werkzaam, dan moet deze een vakcollega uit de onderbouw inschakelen of een docent van een andere school.

19 Gewicht

Het gewicht bij de bepaling van het SE-cijfer dient voor elke toets te zijn aangegeven. De kandidaat moet niet alleen geïnformeerd zijn over de te behalen maximumscore en de puntentoekening per vraag maar ook over het gewicht van de toetsen in het eindcijfer voor het SE.

3.3.4 De kwaliteit van de schoolexamens

Hoe is het gesteld met de kwaliteit van de schoolexamens? In opdracht van de Inspectie van het Onderwijs heeft Cito een onderzoek uitgevoerd naar de kwaliteit van de schoolexamens in het voortgezet onderwijs. Aan het onderzoek hebben 74 havo- en 84 vmbo-scholen meegedaan.

Het onderzoek is uitgevoerd bij de vakken Engels en biologie in de havo en Nederlands en wiskunde in het vmbo. De kwaliteit van het SE is onder leiding van toetsdeskundigen van Cito vastgesteld door in totaal twaalf beoordelaars (drie per vak). Allen waren ervaren en volledig bevoegde docenten die daarnaast betrokken waren bij de constructie van de centrale examens. Na een training in het beoordelen, hebben de beoordelaars de kwaliteit van het SE op de negentien indicatoren beoordeeld die in voorgaande paragraaf omschreven zijn. Belangrijke aandachtspunten waren daarbij de inhoudelijke dekking van het eindexamenprogramma, de kwaliteit van de opgaven, de moeilijkheid van de toetsen, de strengheid/soepelheid van de

beoordelingsmiddelen en de becijfering, de volledigheid en relevantie van de beoordelingsaspecten, het vooraf vastleggen en vroegtijdig meedelen van de scorepunten, het aangeven van het gewicht bij de bepaling van het cijfer, het moment van toetsing en de mate waarin docenten de toetsen, beoordeling en normering in onderling overleg vaststellen. Hieronder vatten we de belangrijkste conclusies van het onderzoeksverslag 'Het schoolexamen in het voortgezet onderwijs' kort samen.

De belangrijkste conclusie uit het onderzoek is dat de kwaliteit van het schoolexamen naar het oordeel van de beoordelaars op de overgrote meerderheid van de onderzochte scholen in orde is. Althans waren naar de mening van de beoordelaars de geconstateerde tekortkomingen vrijwel nooit ernstig genoeg om de kwalificatie 'onvoldoende' te rechtvaardigen. De percentages scholen met een SE van voldoende kwaliteit bedroegen volgens de gehanteerde criteria voor de vakken Engels, Nederlands, biologie en wiskunde respectievelijk 78%, 100%, 96% en 94%.

Een tweede conclusie is dat, ondanks het overwegend positieve oordeel over de algehele kwaliteit van het SE, de verschillen tussen scholen volgens de beoordelaars aanzienlijk zijn. Zo gaven de beoordelaars de school met het 'slechtste' SE Engels op vijf indicatoren een voldoende; de dertien scholen met het 'beste' SE Engels hadden hun schoolexamen op alle negentien indicatoren in orde. Ook bij de vakken biologie en wiskunde verschilden scholen in de kwaliteit van het SE, zij het dat de verschillen minder uitgesproken waren dan bij Engels. Het vak Nederlands vormt hierop een uitzondering in de zin dat de beoordelaars, ondanks dat zij wel kwaliteitsverschillen constateerden, de kwaliteit van het SE Nederlands volgens de gehanteerde criteria van geen enkele school als onvoldoende aanmerkten.

Een derde conclusie was dat de kwaliteit van het schoolexamen naar de mening van de beoordelaars niet voor alle indicatoren even hoog is. Zo waren de verplichte onderdelen van het SE Engels op een derde van de scholen in onvoldoende mate in het examen vertegenwoordigd; daarentegen maakten nagenoeg alle scholen in voldoende mate gebruik van antwoord- of beoordelingsmodellen. Bij biologie werden bijvoorbeeld de toetsen en opdrachten op één op de zes scholen als te moeilijk of te gemakkelijk aangemerkt; daarentegen voldeden de beoordelingsaspecten op nagenoeg alle scholen in redelijke mate aan de eis van relevantie. Bij wiskunde werden bijvoorbeeld de scorepunten naar het oordeel van de beoordelaars op ruim een derde van de scholen niet of te laat aan de kandidaten meegedeeld; daarentegen namen alle scholen de toetsen in voldoende mate af op een moment dat de kandidaten het vereiste niveau redelijkerwijs bereikt konden hebben. Wederom vormt het vak Nederlands een uitzondering: de beoordelaars waren de mening toegedaan dat alle onderzochte scholen aan alle indicatoren van de kwaliteit van het SE voldeden.

Een vierde conclusie was dat de kwaliteit van het schoolexamen niet voor alle vakonderdelen gelijk is. Zo bleken de beoordelaars Engels de kwaliteit van het onderdeel luistervaardigheid veel hoger te waarderen dan van het onderdeel schrijfvaardigheid. Bij wiskunde werd de kwaliteit van de vakonderdelen Algebraïsche verbanden, Rekenen, meten en schatten en Meetkunde op nagenoeg alle scholen als voldoende aangemerkt, maar van het onderdeel Statistiek was de kwaliteit op ruim een vijfde van de scholen niet in orde. Het vak Nederlands vormt hierop een uitzondering omdat de beoordelaars hier van mening waren dat de kwaliteit van alle beoordeelde vakonderdelen voldoende was.

Al met al waren de beoordelaars van mening dat de schoolexamens in het voortgezet onderwijs op de overgrote meerderheid van de scholen in voldoende mate voldoen aan redelijke kwaliteitseisen.

3.3.5 Het verschil tussen de cijfers op het CE en SE

Scholen hebben de laatste jaren meer ruimte gekregen om het SE naar eigen inzicht in te richten. De verantwoordelijkheid voor het SE ligt nu meer dan ooit bij de school zelf. De keerzijde van de medaille is dat de school zich moet verantwoorden, onder meer tegenover de Inspectie die toeziet op de kwaliteit van het SE. Het Toezichtkader van de Inspectie, dat wil zeggen het 'Toezichtkader po/vo 2009' (met een aanpassing in 2011), eist dat de school de kwaliteit van het schoolexamen waarborgt en verantwoordt tegenover belanghebbenden. De Inspectie toetst onder meer of de cijfers op het SE niet te hoog zijn in vergelijking met die van het CE, dat wil zeggen, niet hoger zijn dan wat redelijkerwijs van de leerlingen verwacht mag worden. De Inspectie heeft geconstateerd dat de SE-cijfers op sommige scholen te hoog zijn in vergelijking met de CE-cijfers. Overdreven hoge SE-cijfers staan op gespannen voet met het streven om kandidaten onafhankelijk van de school gelijke kansen te bieden. Voor een individueel vak is een SE-CE-verschil soms goed te verklaren, bijvoorbeeld als de school in het SE andere leerstof toetst dan het CE en de school aannemelijk kan maken dat de kandidaten de SE-stof beter beheersen dan die van het CE. Is het SE-cijfer echter over een periode van drie jaar gemiddeld meer dan een half punt hoger dan het CE-cijfer, dan kan de Inspectie de school onder verscherpt toezicht plaatsen. De school moet dan een verbeterplan opstellen en laten zien dat er vooruitgang wordt geboekt. Op dit moment is verscherpt toezicht voor de Inspectie de enige mogelijkheid. Met ingang van het schooljaar 2013-2014 kan scholen de bevoegdheid tot het organiseren van examens worden ontnomen. De kandidaten zijn dan aangewezen op de staatsexamens. Omdat dit ernstig gezichtsverlies betekent, is de verwachting dat scholen het niet zover zullen laten komen.

3.3.6 Eindcijfers en einduitslag

Voor de meeste vakken geldt dat het diplomacijfer een gemiddelde is van het SE-cijfer en het CE-cijfer. Hierbij tellen beide cijfers meestal even zwaar mee. Sommige vakken worden alleen met een SE afgesloten. Daartoe behoren maatschappijleer, de kunstvakken (inclusief CKV) en lichamelijke opvoeding. Voor deze vakken is het cijfer op het SE direct het eindcijfer.

Als de eindcijfers van alle vakken, eventueel na een herkansing, zijn vastgesteld, kan de uitslag worden bepaald. Hoe dat in zijn werk gaat, staat in de zogeheten slaag-zakregels. Deze zijn niet voor alle onderwijssoorten gelijk. De slaag-zakregels zijn gepubliceerd in artikel 49 van het Examenbesluit, zie www.examenblad.nl. De regels zijn met ingang van het schooljaar 2011-2012 aangescherpt. Hiermee beoogt de overheid de 'zesjescultuur' tegen te gaan en de doorstroming naar het hoger onderwijs te verbeteren. De verscherpte exameneisen worden gefaseerd ingevoerd en omvatten vooralsnog drie maatregelen:

Met ingang van het schooljaar 2011-2012 geldt de eis dat een kandidaat voor de centrale examens gemiddeld een voldoende moet hebben. Is het gemiddeld cijfer op het CE lager dan 5,5, dan is de kandidaat gezakt, ook al staat daar een zeer hoog cijfer op het SE tegenover.

Met ingang van het schooljaar 2011-2012 krijgt het SE in de basisberoepsgerichte leerweg, net als in de andere onderwijssoorten al het geval was, hetzelfde gewicht als het CE.

Met ingang van het schooljaar 2012-2013 mogen kandidaten in havo en vwo als eindcijfer (het gemiddelde van CE en SE) voor de vakken Nederlands, Engels of wiskunde nog maar één vijf behalen; voor leerlingen zonder wiskunde geldt dat hooguit één vijf voor Nederlands of Engels behaald mag worden. Deze laatste aanscherping geldt overigens nog niet voor het vmbo.

De minister wil hierover pas een besluit nemen nadat de referentieniveaus in het vmbo zorgvuldig zijn ingevoerd. Voor meer informatie over de eventuele verdere aanscherping van de exameneisen in het vmbo, raadpleeg het artikel van Beuving en Waterreus (2010).

3.4 De betrouwbaarheid van examens

Hoofdstuk 3 van het algemene deel van Toetsen op School is gewijd aan de betrouwbaarheid van toetsscores. Voor een goed begrip van wat hierna besproken wordt, is het gewenst dat de lezer kennis heeft genomen van dit hoofdstuk. In deze paragraaf presenteren we eerst de betrouwbaarheid van de scores van afzonderlijke centrale examens (van het eerste tijdvak 2011). Daarna gaan we in op de betrouwbaarheid van de scores voor examenvakken, dat wil zeggen de betrouwbaarheid van centrale examens in combinatie met schoolexamens. Ten slotte komt de betrouwbaarheid van de slaag-zakbeslissing die genomen wordt op basis van meerdere examens aan de orde, denk bijvoorbeeld aan het havodiploma.

Betrouwbaarheid van de scores voor een afzonderlijk centraal examen

De betrouwbaarheden van de centrale examens voor het eerste tijdvak worden jaarlijks gepubliceerd op de website van Cito, zie www.cito.nl > Voortgezet onderwijs > Centrale examens > Examenverslagen. De betrouwbaarheid van de scores blijken voor het ene vak en schooltype hoger dan voor het andere. De Commissie Testaangelegenheden Nederland (COTAN) heeft normen opgesteld voor het evalueren van de betrouwbaarheid van toetsen. Geïnteresseerde lezers verwijzen we naar hoofdstuk 10 van het algemene deel van Toetsen op School. Als het om belangrijke beslissingen zoals examens gaat, beoordeelt de COTAN een betrouwbaarheid van minimaal 0,90 als goed, tussen de 0,80 en 0,90 als voldoende en onder de 0,80 als onvoldoende. Als we de norm van de COTAN als maatstaf zouden nemen, zou een flink aantal afzonderlijke centrale examens als onvoldoende betrouwbaar gekwalificeerd moeten worden. De vraag is echter of de COTAN-norm hier wel van toepassing is. De beslissing voldoende of onvoldoende voor bijvoorbeeld het vak Duits havo wordt immers niet alleen genomen op basis van het centraal examen. Voor het nemen van die laatste beslissing is ook het cijfer op het havo-schoolexamen Duits van belang. Hierna zullen we zien dat de betrouwbaarheid van een examenvak, dat wil zeggen de combinatie van een centraal examen en een schoolexamen, veel hoger is dan de betrouwbaarheid van alleen een centraal examen of alleen een schoolexamen.

Waarom is het ene examen betrouwbaarder dan het andere?

Eerder constateerden we dat de betrouwbaarheid van het CSE sterk van examen tot examen verschilt. Aan betrouwbaarheidsverschillen tussen examens kunnen verschillende oorzaken ten grondslag liggen. Een eenduidige verklaring waarom het ene examen betrouwbaarder is dan het andere is dan ook niet eenvoudig te geven. Hieronder geven we een aantal mogelijke verklaringen:

- Betrouwbaarheidsverschillen kunnen het gevolg zijn van beoordelaarsfouten. Zo doen de CSPE's voor de praktijkvakken in het vmbo een groter beroep op het subjectieve menselijke oordeelsvermogen dan bijvoorbeeld de centraal schriftelijke examens wiskunde. Zoals uit onderzoek blijkt, zijn de CSPE's dan ook gevoeliger voor beoordelaarsfouten. In dit verband dient wel opgemerkt te worden dat bij het bepalen van de betrouwbaarheid van de examens geen rekening gehouden wordt met een onzorgvuldige correctie of verschillen in soepelheid tussen correctoren. De reden voor dat laatste is dat het verzamelen van de beoordelingen van de vele beoordelaars die bij de beoordeling van de verschillende examens betrokken zijn, logistiek ingewikkeld en dan ook zeer kostbaar is. Degenen die de betrouwbaarheid van de examens uitrekenen, gaan er noodgedwongen vanuit dat alle examenkandidaten als het ware door een en dezelfde beoordelaar beoordeeld zijn. Verschillen tussen correctoren zullen de betrouwbaarheid van een examen negatief beïnvloeden – hoe groter die verschillen hoe negatiever die invloed.
- Betrouwbaarheidsverschillen kunnen het gevolg zijn van verschillen tussen vakgebieden. Een betrouwbaar examen is gemakkelijker te realiseren bij een homogeen vakdomein dan bij een heterogeen vakdomein. Deskundigen lijken het erover eens te zijn dat de domeinen/

examenprogramma's van de beroepsgerichte vakken in het vmbo over het algemeen smaller en homogener zijn dan die van de algemeen vormende vakken in havo en vwo.

- Betrouwbaarheidsverschillen kunnen het gevolg zijn van vaardigheidsverschillen van examenkandidaten. In het vmbo zijn de vaardigheidsverschillen tussen de kandidaten over het algemeen groter dan in meer homogene populaties van havo en vwo. Naarmate de geëxamineerde kandidaten sterker verschillen in vaardigheid, behoort een hogere betrouwbaarheid van het examen meer tot de mogelijkheden.
- Betrouwbaarheidsverschillen kunnen het gevolg zijn van vraagvormen die gebruikt worden. Uit onderzoek in 2003 van Kuhlemeier, Steentjes en Kleintjes bij het vak wiskunde in de onderbouw is bijvoorbeeld gebleken dat examens bestaande uit open vragen betrouwbaarder zijn dan examens bestaande uit meerkeuzevragen. Althans, als de meerkeuzevraag en de open vraag betrekking hebben op precies dezelfde leerstof en de alternatieven van de meerkeuzevraag ook in het antwoordmodel bij de overeenkomstige open vraag terugkomen. De hogere betrouwbaarheid van open vragen zit hem daarbij niet zozeer in het vraagtype maar meer in het gegeven dat open vragen doorgaans meer genuanceerd gescoord worden dan meerkeuzevragen. Bij meerkeuzevragen heeft de corrector immers vaak alleen de keuze tussen goed (score 1) en fout (score 0), terwijl de corrector bij open vragen doorgaans een keuze uit veel meer scores kan maken. Met een open vraag kan derhalve een beter onderscheid worden gemaakt tussen goede en minder goede kandidaten dan met een enkelvoudige meerkeuzevraag mogelijk is. Dit verklaart wellicht waarom de betrouwbaarheid van een examen met open vragen die lastig te beoordelen zijn toch niet onder hoeft te doen voor een objectief scorebaar examen met alleen meerkeuzevragen.
- Betrouwbaarheidsverschillen kunnen het gevolg zijn van verschillen in de moeilijkheidsgraad van examens. Uit onderzoek blijkt dat de betrouwbaarheid van toetsen minder wordt naarmate de gemiddelde moeilijkheidsgraad sterker afwijkt van 50% van de maximale score. Een extreem moeilijke of gemakkelijke toets zal dan ook minder betrouwbaar zijn dan een toets waarbij de leerlingen de helft van de opgaven goed maken. Zo wordt het lastig om een betrouwbare toets te maken als het onderwijs effectief is, leerlingen de stof zeer goed beheersen en bijgevolg bijna alle examenvragen met succes beantwoorden.
- Betrouwbaarheidsverschillen kunnen het gevolg zijn van verschillen in aantallen vragen. In hoofdstuk 3 van het algemene deel van Toetsen op School is uiteengezet dat een hoge betrouwbaarheid het best gerealiseerd kan worden door veel vragen in het examen op te nemen. Gezien de relatieve onbetrouwbaarheid van sommige centrale examens valt te overwegen het aantal vragen waar dat nodig is te verhogen. De keerzijde van de medaille is echter een toename van de duur van het examen voor de kandidaten en een toename van de correctielast voor de docenten. Uit het onderzoek van Kuhlemeier en Kremers (in voorbereiding) naar de praktijk van de eerste en tweede correctie blijkt dat veel docenten de correctie als tijdsintensief en belastend ervaren. Met name docenten van examens met een hoge correctielast zoals geschiedenis en Nederlands zullen een nog langer examen doorgaans niet op prijs stellen.

Voor gedetailleerde informatie over de centrale examens kan men terecht op de website van Cito. De zogeheten examenverslagen zijn bereikbaar via www.cito.nl > Voortgezet onderwijs > Centrale examens > Examenverslagen.

De rubriek Examenverslagen bevat de volgende onderdelen:

- Examenverslag: een rapport dat voor vmbo, havo en vwo inzicht geeft in de aantallen kandidaten, de productie van de examens, de verzameling van afnamegegevens en de normering.
- De resultaten per schooltype/leerweg: een overzicht van de landelijke resultaten per schooltype.
- De Excel-applicaties Resultaten per examen en Betekenis geven aan cijfers: bestanden die

gedetailleerde kwantitatieve informatie over de examens geven. Voor toelichting en uitleg van deze applicaties verwijzen we naar hoofdstuk 3 van het algemene deel van Toetsen op School.

- Waardering per examen: een overzicht van de resultaten van een korte enquête die gehouden is onder correctoren.

De betrouwbaarheid van de scores voor een examenvak

Hiervoor werd opgemerkt dat beslissingen, zoals een voldoende of onvoldoende cijfer, bij examens genomen worden op basis van het centrale examen in combinatie met het school-examen. Hoewel informatie over de betrouwbaarheid van schoolexamens niet beschikbaar is, is het soms wel mogelijk om de betrouwbaarheid van schoolexamens te schatten. Uit onderzoek van Van Rijn, Béguin en Verstralen (2009a, 2009b) blijkt dat de betrouwbaarheid van het schoolexamen veelal ongeveer gelijk zal zijn aan de betrouwbaarheid van het centrale examen. Dit laatste betekent dat we de betrouwbaarheid van een examenvak kunnen schatten met de Spearman-Brown formule voor testverlenging die in hoofdstuk 3 van het algemene deel van Toetsen op School besproken is. Toepassing van deze formule houdt bijvoorbeeld in dat met een betrouwbaarheid van het centrale examen die gelijk is aan 0,66, de betrouwbaarheid van het examenvak gelijk is aan 0,80. Dit betekent dat we op basis van een examenvak veel nauwkeuriger beslissingen kunnen nemen dan op basis van alleen het centrale examen of alleen het schoolexamen. Het voorgaande betekent dat de betrouwbaarheid van de meeste examenvakken voldoende of goed zal zijn.

De betrouwbaarheid van de slaag-zakbeslissing

Van Rijn, Béguin en Verstralen (2009a, 2009b) hebben onderzoek gedaan naar de betrouwbaarheid van de slaag-zakbeslissing op basis van meerdere examens, dat wil zeggen het vmbo-, havo- en vwo-diploma. Ondanks dat de betrouwbaarheden van de verschillende examenvakken erg hoog zijn, schatten zij dat voor bijvoorbeeld het vwo-diploma het percentage misclassificaties, dat wil zeggen het percentage kandidaten dat ten onrechte zakt of slaagt, toch nog gelijk is aan 6%. Wel dient opgemerkt te worden dat bij die schatting van 6% geen rekening gehouden is met het effect van eventuele herkansingen. Het percentage kandidaten dat ten onrechte slaagt of ten onrechte zakt, is met name afhankelijk van de betrouwbaarheid van de afzonderlijke examenvakken en de gehanteerde slaag-zakregels. Hoe hoger de betrouwbaarheid van de afzonderlijke examenvakken, hoe minder examenkandidaten ten onrechte hun diploma (niet) krijgen.

3.5 Het gebruik van examens voor andere doeleinden

Behalve de afsluitings- en toelatingsfunctie vervullen examens nog vele andere belangrijke functies. Hieronder bespreken we het gebruik van het examen voor pedagogische doeleinden, de vernieuwing van het onderwijs, het beoordelen van docenten en het beoordelen van scholen.

Pedagogische functie

Hoewel een examen niet bedoeld is om leerlingen te motiveren, speelt het motivatie-element wel degelijk een belangrijke rol, zeker in het voortgezet onderwijs. Veel docenten in het voortgezet onderwijs zullen op de vraag waarom ze leerlingen toetsen voorleggen als antwoord geven: 'Mijn leerlingen doen alleen maar iets als ze er een cijfer mee kunnen verdienen. Bovendien zijn het zesjesleerlingen: als ze zich er met een minimale inspanning van af kunnen maken, dan doen ze dat.' Het examen wordt vaak gebruikt als stok achter de deur om de leerlingen aan het leren te krijgen. Calculerende leerlingen hebben, meer dan leerlingen die intrinsiek gemotiveerd zijn voor een bepaald vak, een grote behoefte aan informatie over de eisen waaraan hun prestatie moet voldoen. Ze willen precies weten welke leerstof ze moeten

bestuderen om te slagen, ook al is het met de hakken over de sloot. Het examen motiveert leerlingen met andere woorden tot inzet en leren. In het advies 'Examinering in ontwikkeling' rangschikt de Onderwijsraad deze pedagogische functie onder de kwaliteitsfunctie van de examens. Daartoe behoren aldus de Onderwijsraad ook de volgende functies van examens:

- examens bepalen mede de inhouden van programma's en de handelwijzen van studenten en docenten (*backwash* effecten);
- examens prikkelen de 'talentontwikkeling' van leerlingen;
- examens geven informatie aan ouders en verantwoording aan derden;
- examens bevorderen de communicatie met het afnemend veld en tussen docenten onderling.

Vernieuwing van het onderwijs

Een maatschappelijke functie van de examens is het bijdragen aan de vernieuwing van het onderwijs. Examens zijn niet alleen het sluitstuk van het onderwijs, maar kunnen ook het beginpunt van vernieuwingen zijn. Examens zijn een belangrijk middel om nieuwe onderwijsdoelen en leerstof die nog maar mondjesmaat onderwezen worden ingang te doen vinden in de onderwijspraktijk. De sturende werking van eindexamens wordt ook wel aangeduid met de eerder genoemde term *backwash*.

Backwash effecten van examens zijn niet altijd positief. Van een negatief teruglageffect is sprake als de docent zich in het laatste examenjaar beperkt tot het geven van examentraining en andere belangrijke onderwijsdoelen verwaarloost.

In onderstaand kader staat een voorbeeld van hoe toetsing en examinering de vernieuwing kan ondersteunen bij de beoordelingspraktijk op een school in het oosten van het land.

Onze school werkt al langere tijd aan competentiegericht onderwijs. Van alle docenten wordt geëist dat zij uit een scala van competentiegerichte toetsvormen er ten minste twee hanteren bij de beoordeling van de leerresultaten. Door zo te werken wordt de keuze voor een bepaalde toetsvorm een hulpmiddel in het vernieuwingstraject. Immers, door het toetsen en de examenprogramma's mede te enten op of te laten aansluiten bij nieuwe opvattingen van leren, krijgen leerlingen geen tegenstrijdige signalen meer. Zo staat het toets- en examenprogramma de implementatie van nieuwe vormen van leren niet langer in de weg, maar zal het daartoe zelfs een krachtige impuls geven.

Centrale examens kunnen vernieuwingen op gang helpen, maar ook in de weg staan. Innovatieve scholen die serieus werk maken van vakkenintegratie ervaren het examineren van 'losse' vakken doorgaans als niet passend bij hun onderwijs. Een van de hindernissen is dat docenten regelmatig afzien van deelname aan experimenten met vakintegratief leren omdat zij vrezen dat de leerlingen het niveau van het examen dan niet meer halen (lees: het boek niet uitkrijgen). Zo gezien houdt examinering van 'losse' vakken vakkenintegratie dus tegen. Een onderliggend probleem is dat de inhoud van de methoden het vereiste eindniveau meestal ver overstijgt. Er komt veel meer aan de orde dan gezien de eindtermen nodig is. Lang niet alle docenten zijn even vertrouwd met de eindtermen. Ze houden zich aan de methode en menen dat ze hun tijd hard nodig hebben om 'het boek uit te krijgen'. Een integratief centraal examen zou de nog maar moeilijk op gang komende integratie van algemene en beroepsgerichte vakken een krachtige stimulans kunnen geven.

Met name voor leerlingen in de basisberoepsgerichte leerweg vormen centrale examens een struikelblok, zegt Andries Meijer van de Onderwijscampus in Winschoten: 'Het reproduceren van theoretische kennis is voor deze leerlingen erg lastig. Daarom hebben ze een groot probleem met het centraal examen. De leerlingen leren tijdens hun opleiding (volgens het concept natuurlijk leren) juist vaardigheden om dingen die ze niet weten op te zoeken, waar ze dat moeten doen, en dergelijke. Dat vereist een ander examen. Het is lastig deze leerlingen in korte tijd klaar te stomen voor het centraal examen.'

Andere maatschappelijke functies van het examen (naast het faciliteren dan wel belemmeren van onderwijsvernieuwing) zijn volgens de Onderwijsraad:

- monitoren/beïnvloeden van het algemene onderwijspeil;
- beoordelen prestaties van het onderwijs in relatie tot kosten van het onderwijs;
- bijdragen aan nationale doelstellingen als gelijkheid van kansen en nationale identiteit;
- benchmarking in internationaal opzicht (diplomavergelijking).

Beoordelen van docenten

Een geheel nieuwe functie van de examens is het gebruik voor prestatiedifferentiatie, in de wandelgangen bekend als 'loon naar lesgeven'. Op dit moment wordt het salaris van een docent grotendeels bepaald door het aantal jaren dat hij of zij in het onderwijs werkzaam is. De behaalde examenresultaten vormen tegenwoordig een belangrijk onderwerp in de functioneringsgesprekken tussen docenten en leidinggevend. De uitkomst van deze gesprekken hebben echter nog maar zelden financiële consequenties voor de docent in kwestie. De financiële prikkel om als docent extra goed je best te doen is dus niet erg groot. Sommige politieke partijen pleiten er daarom voor om goede docenten meer salaris te geven. De hoogte van de examencijfers die de docent met zijn of haar leerlingen realiseert, lijkt dan een relatief hard criterium.

Het gebruik van examencijfers voor andere functies dan waarvoor ze oorspronkelijk ze bedoeld zijn, heeft ook keerzijden. De concurrentie tussen docenten om een salarisverhoging kan ten koste gaan van de onderlinge samenwerking. Een gebrek aan heldere en meetbare doelen en criteria kan leiden tot wantrouwen in de leiding van de school. Beloningsdifferentiatie op basis van examencijfers kan strategisch of frauduleus handelen in de hand werken. Als docenten zich te sterk op hoge examencijfers gaan concentreren, kunnen andere belangrijke onderwijsdoelstellingen en de sfeer in de klas daaronder te lijden hebben. Voorbeelden van frauduleus handelen zijn het geven van ontoelaatbare hulp aan kandidaten tijdens de afname van het examen en het welbewust geven van te hoge cijfers bij de correctie van examenwerk. Docenten die welbewust te soepel nakijken, bevoordelen hun eigen leerlingen (en wellicht zichzelf), maar duperen leerlingen van scholen waar wel integer beoordeeld wordt. Deze ongewenste neveneffecten kunnen op termijn de nu nog hoge waarde van het diploma aantasten.

Beoordelen van scholen

Examens worden niet alleen gebruikt voor afsluiting en toelating. Een andere functie is evaluatie en verbetering van het onderwijs op school. De grotere autonomie van scholen brengt met zich mee dat zij zich tegenwoordig publiekelijk moeten verantwoorden voor hun prestaties. Hiervoor worden in toenemende mate examenresultaten gebruikt (met name cijfers en slaag-/zakpercentages). Dag- en weekbladen zoals Trouw en Elsevier publiceren jaarlijks zogeheten ranglijsten waarin scholen op basis van hun examenresultaten op een rij worden gezet. De VO-raad heeft een website Vensters voor Verantwoording, zie

www.venstersvoorverantwoording.nl, met vergelijkende informatie over de kwaliteit van scholen maar geen ranglijsten van scholen.

De Inspectie maakt in het toezicht op de onderwijskwaliteit veelvuldig gebruik van examengegevens. Informatie over de examenresultaten en de doorstroom rapporteert de Inspectie in de zogeheten Opbrengstenkaart (voorheen Kwaliteitskaart genoemd). Dit overzicht laat zien hoe elke school het doet in vergelijking met andere scholen in de regio en met het landelijk gemiddelde. De Opbrengstenkaart en de beoordelingen van vrijwel alle scholen voor voortgezet onderwijs zijn te downloaden via <http://toezichtkaart.owinsp.nl/schoolwijzer>. Scholen en docenten kunnen deze gegevens gebruiken voor zelfevaluatie, om zichzelf een spiegel voor te houden en hun onderwijs te verbeteren.

Het openbaar maken van examenresultaten van scholen voorziet in een bepaalde behoefte. Burgers zijn doorgaans ook belastingbetalers die willen weten of scholen hun geld goed besteden. Ouders willen scholen kunnen vergelijken om een geschikte school voor hun kind te kunnen kiezen. In het kader van de marktwerking wil de overheid scholen prikkelen tot het leveren van meer kwaliteit. In technisch opzicht blijkt het echter lastig om aan de hand van examengegevens de kwaliteit van de individuele school of docent zuiver vast te stellen. Examens meten de kennis, inzicht en vaardigheden waarover leerlingen aan het einde van hun opleiding beschikken. Ze maken niet zichtbaar hoe goed het schoolbeleid is geweest of hoe goed de docenten hebben lesgegeven. De examenresultaten hangen namelijk ook af van factoren waar de school weinig invloed op heeft, maar die wel van grote betekenis zijn voor de onderwijsresultaten, zoals de intelligentie en de sociaal-economische achtergrond van de leerlingen. Zo is het voor een witte school uit Wassenaar veel eenvoudiger om de leerlingen tot hoge examenprestaties te brengen dan voor een zwarte school in een onderwijsstimuleringsgebied. Wil men scholen eerlijk met elkaar kunnen vergelijken, dan moet er rekening worden gehouden met de samenstelling van de leerlingenpopulatie op de school in kwestie. Onderzoekers hebben geprobeerd de 'netto' bijdrage van de school uit te rekenen. Als zij in hun statistische analyses rekening houden met de samenstelling van de leerlingenpopulatie, dan blijken de verschillen tussen scholen grotendeels te verdwijnen. Kennelijk verschillen de scholen voor voortgezet onderwijs - uitzonderingen daargelaten - veel minder in de kwaliteit van hun onderwijs dan de ranglijsten doen vermoeden.

Tip

Bronnen voor verdere studie

- www.examenblad.nl
- www.cvte.nl
- www.rijksoverheid.nl/ministeries/ocw > Onderwerpen > Voortgezet onderwijs
- www.cito.nl > Onderwijs > Voortgezet onderwijs
- www.laks.nl
- www.examenklacht.nl
- www.toetswijzer.nl
- www.schoolexamensvo.nl
- www.venstersvoorverantwoording.nl
- www.schoolvo.nl
- www.examenbundel.nl
- www.examen.nl
- www.eigenwijzer.nl
- www.eindexamens.nu
- <http://examen.pagina.nl>
- <http://eindexamen.3fm.nl>
- www.scholieren.com/eindexamens
- www.eindexamensite.nl
- www.elsevier.nl > Weekblad > Onderzoeken > De beste scholen
- <http://schoolprestaties.trouw.nl>

4 Toetsen met het ERK in de klas

4 Toetsen met het ERK in de klas

Alma van Til

In de vorige hoofdstukken is het Europees Referentiekader (ERK) al enkele malen ter sprake gekomen. In hoofdstuk 1 worden verschillende toetsen behandeld die gekoppeld zijn aan het ERK. In hoofdstuk 2 wordt uit de doeken gedaan wat het ERK inhoudt. Dit hoofdstuk is bedoeld voor docenten die al enigszins bekend zijn met het ERK en die hun toetsing daarop willen laten aansluiten. We behandelen achtereenvolgens de inhoud van ERK-toetsen, het opstellen van een toetsplan en een toetsmatrijs, het maken van opdrachten en een correctievoorschrift volgens ERK-criteria en het normeren van de resultaten.

Dit hoofdstuk is een bewerking van de publicatie 'Toetsen en beoordelen met het ERK' (2011), te downloaden via www.slo.nl > Publicaties > Recente uitgaven. Deze publicatie is het resultaat van een samenwerkingsverband tussen SLO en Cito en biedt naast theoretische achtergronden ook allerlei praktische tips en voorbeelden om als docent mee aan de slag te gaan.

4.1 De inhoud van ERK-toetsen

De inhoud van een ERK-toets wordt bepaald door het ERK. Simpel kan het niet, zou je zeggen. Toch schuilt er een addertje onder het gras. De door de Nederlandse Taalunie uitgegeven Nederlandse vertaling van het ERK, getiteld 'Gemeenschappelijk Referentiekader voor de Moderne Vreemde Talen: Leren, Onderwijzen, Beoordelen' telt 232 pagina's en 54 schalen, zie www.taaluniversum.org. Het gevaar dreigt al snel dat men door de bomen het bos niet meer ziet. Daarom is het goed om eerst de visie van het ERK nader onder de loep te nemen. Die visie wordt uit de doeken gedaan in hoofdstuk 2 van voornoemde publicatie. In dat hoofdstuk wordt gesteld dat het leren van het Engels, Frans, Duits of Spaans uiteindelijk tot doel heeft om de taalleerder te laten functioneren in het desbetreffende taalgebied. En of het nu gaat om het kopen van een brood in de winkel of het promoten van Nederlandse windmolens bij een buitenlandse overheid, uiteindelijk gaat het erom dat taalleerders bepaalde taken kunnen uitvoeren in de moderne vreemde taal. Dat ze dus de winkel uitgaan met dat brood of thuishkomen met die order voor nieuwe windmolens.

Voor docenten die ERK-toetsen willen maken, is het goed om deze visie in het achterhoofd te houden want het betekent dat toetsen die de beheersing van vaardigheden evalueren niet kunnen ontbreken. Degenen die zich willen beperken tot het evalueren van woordenschat en/of grammatica en andere ondersteunende vaardigheden kunnen tot op een bepaalde hoogte hun toetsen wel relateren aan het ERK. Er zijn namelijk schalen geformuleerd voor grammaticale correctheid en het bereik en de beheersing van woordenschat. Het zou echter te ver gaan om te beweren dat toetsen die geïsoleerd kennis van grammatica en woordenschat toetsen helemaal 'ERK-proof' zijn omdat ze voorbijgaan aan de visie van het ERK op taal.

Het toetsen van ondersteunende vaardigheden (woordenschat, grammatica, spelling, uitspraak, etc.) is in het Nederlandse vreemdetalenonderwijs een veelbesproken onderwerp. Veel docenten gebruiken dit soort toetsen als didactisch hulpmiddel. Leerlingen die weten dat ze een toets leesvaardigheid krijgen, zullen de neiging hebben om zich minimaal voor te bereiden 'omdat er toch niets te leren valt'. Bij woordenschat en grammatica is dat anders. De taal wordt als het ware aangeboden in hapklare brokken die de leerling eenvoudigweg tot zich kan nemen. De aanname is vervolgens dat leerlingen daardoor ook beter gaan presteren op vaardigheids-

toetsen. Of dit inderdaad zo is, is een discussie die we hier niet gaan voeren. Feit is dat het toetsen van ondersteunende vaardigheden in het Nederlandse taalonderwijs veel voorkomt en de vraag is hoe dit soort toetsing zich verhoudt tot het ERK.

Hierboven stelden we al dat evaluaties die zich inhoudelijk uitsluitend richten op ondersteunende vaardigheden niet ERK-proof zijn. Maar hoe staat het met een mix van toetsing waarbij enerzijds spreken, lezen, luisteren, etc. aan bod komen en anderzijds grammatica, woordenschat, uitspraak, etc.? Kunnen docenten die het toetsen van ondersteunende vaardigheden sec, dus niet in een communicatieve context, afwisselen met lees-, luister-, spreek- en schrijftoetsen wél claimen dat hun evaluatie strookt met de visie van het ERK? Het antwoord hangt af van het gewicht dat de verschillende onderdelen in de schaal leggen. Het is van belang dat de nadruk ligt op het toetsen van vaardigheden als lezen, luisteren, spreken, gesprekken voeren en schrijven. Is dat niet het geval, dan wordt het lastig om vol te houden dat men bij toetsing de ideeën heeft gevolgd die leidend waren bij de bedenkers van het ERK.

4.1.1 Vaardigheden apart of geïntegreerd toetsen

Het ERK onderscheidt vijf verschillende vaardigheden:

- lezen
- luisteren
- mondelinge productie-activiteiten (spreken)
- algemene mondelinge interactie (gesprekken voeren)
- schrijven

Het ligt voor de hand dat men bij toetsing dit onderscheid volgt en aparte toetsen maakt voor elk van deze vaardigheden. Bovendien is in het eindexamenprogramma een aparte eindterm beschreven voor deze vaardigheden. Voordeel van deze aanpak is ook dat het de mogelijkheid geeft om ‘deelkwalificaties’ te halen. Lees- en luistervaardigheid worden in het Nederlandse vreemdetalenonderwijs vaak op een hoger ERK-niveau beheerst dan de productieve vaardigheden. Het is bijvoorbeeld heel goed mogelijk dat een 6 vwo-leerling voor luistervaardigheid Duits presteert op niveau B2 en spreekvaardigheid Duits op niveau B1.

Toch zou het jammer zijn om toetsen waarbij meerdere vaardigheden geïntegreerd worden meteen af te keuren. Vooral omdat men met dit soort toetsen de realiteit soms dicht kan benaderen. Neem bijvoorbeeld een toets waarbij leerlingen eerst de telefoon op moeten nemen en een gesprek moeten voeren. Vervolgens moeten ze op basis van dat gesprek een notitie achterlaten voor iemand anders. Er is een kans dat leerlingen in de werkelijkheid ook met een soortgelijke situatie geconfronteerd zouden kunnen worden. Daarmee is deze opdracht, waarbij zowel gespreks- als schrijfvaardigheid een rol spelen, heel levensecht.

Gelukkig biedt het ERK ook een kapstok voor geïntegreerde toetsing in de vorm van een globale schaal. Deze geeft voor alle niveaus een algemeen beeld waarop taalleerders functioneren. En dit algemene beeld heeft betrekking op meerdere vaardigheden. Geïntegreerde toetsing op basis van het ERK is dus mogelijk. Wel moet men zich daarbij realiseren dat het bij slechte prestaties van leerlingen op de toets lastig te achterhalen is wat daarvan de oorzaak kan zijn. In het geval van hiervoor genoemd voorbeeld: hebben leerlingen slecht gescoord omdat ze moeite hebben met gespreksvaardigheid of omdat ze het lastig vinden om één en ander in de vreemde taal op papier te zetten.

Voordat men beslist over het wel of niet inzetten van geïntegreerde toetsen is het dus goed om eerst na te denken over het doel van de toets. Gaat het bijvoorbeeld om een assessment waarbij vastgesteld moet worden of leerlingen geschikt zijn voor het volgen van een bepaalde buitenlandse opleiding of om een toets die als doel heeft om na te gaan op welke fronten

leerlingen achterblijven? Het zal duidelijk zijn dat geïntegreerde toetsen zich beter lenen voor het eerste doel dan voor het tweede.

4.2 Toetsen met behulp van een toetsplan en een toetsmatrijs

Om te waarborgen dat de vaardigheden die het ERK onderscheidt voldoende aan bod komen, is het aan te bevelen om een toetsplan te maken. In zo'n plan wordt aangegeven wanneer welke vaardigheden worden getoetst en hoe zwaar deze meetellen. Daarnaast wordt in het toetsplan ook vermeld welke toetsvormen worden gebruikt. Zo kan schrijfvaardigheid de ene keer getoetst worden met behulp van een computer-based test (CBT) waarin een e-mail moet worden verstuurd van minstens 200 woorden, en de andere keer weer met een schriftelijke toets waarbij leerlingen steeds zinnen moeten aanvullen. Aan elke toetsvorm kleven voor- en nadelen. Door in het toetsplan een mix van toetsvormen op te nemen, kan men ervoor zorgen dat de zwakke kanten van de ene toetsvorm worden gecompenseerd door de andere toetsvorm. In hoofdstuk 4 van 'Toetsen en beoordelen met het ERK' zijn verschillende voorbeelden van toetsplannen opgenomen. Deze maken in één oogopslag duidelijk of:

- de vaardigheden op een evenwichtige manier aan bod komen in zowel de onder- als de bovenbouw;
- de toetsdruk over alle leerjaren heen even groot is;
- er sprake is van variatie in toetsvormen.

Als alle toetsen met behulp van het meerjarenplan zijn ingepland, kun je als docent aan de slag met de inhoud van de toetsen. Voordat je start met het bedenken van opdrachten is het echter aan te bevelen om een toetsmatrijs te maken. Hiermee kan men ervoor zorgen dat de toetsinhoud op een evenwichtige manier aan bod komt. Want een toets die uitsluitend bestaat uit opdrachten waarin gemaïld moet worden met partnerscholen is in de ogen van de leerlingen misschien leuk, maar inhoudelijk heel beperkt. Subvaardigheden als aantekeningen maken, berichten opstellen, formulieren invullen, verslagen en rapporten of vrij schrijven, komen in de toets immers niet aan de orde en ook de verschillende domeinen blijven onderbelicht.

De klassieke definitie van een toetsmatrijs is een tabel waarin aangegeven wordt hoe de opgaven worden verdeeld over twee dimensies: inhoudscategorieën en gedragscategorieën. Daardoor wordt het mogelijk om systematisch te werk te gaan bij het ontwikkelen van een toets. In hoofdstuk 2 van het algemene deel van Toetsen op School wordt aangegeven welke voordelen deze werkwijze heeft.

4.2.1 De inhoudscategorieën van een toetsmatrijs voor het ERK

Voor het bepalen van de inhoudscategorieën van een ERK-toetsmatrijs zou men zich kunnen richten op de domeinen die het ERK onderscheidt. Domeinen zijn contexten waarbinnen taalactiviteiten plaatsvinden. Elk domein kent een eigen soort taalgebruik. Van de ene naar de andere context schakelen, levert dan ook nogal eens problemen op. Iedereen kent wel voorbeelden van personen die uitstekend mee kunnen komen in de vreemde taal zolang er maar over het werk wordt gepraat. Neemt het gesprek opeens een andere wending en komt een huis-, tuin- en keukenonderwerp zoals bijvoorbeeld een mierenplaag aan de orde, dan staat diezelfde persoon met de mond vol tanden. Kortom: domeinen zijn heel bepalend voor het soort taalgebruik dat men nodig heeft en lenen zich daarom goed voor een indeling in inhoudscategorieën. Het ERK definieert er vier:

- het publieke domein;
- het persoonlijke domein (hierna ook wel aangeduid als 'dagelijks leven');
- het professionele domein (hierna ook wel aangeduid als 'werk');
- het educatieve domein (hierna ook wel aangeduid als 'opleiding').

4.2.2 De gedragscategorieën van een toetsmatrijs voor het ERK

De tweede dimensie van een klassieke toetsmatrijs betreft gedragscategorieën. Daarbij wordt meestal een onderscheid gemaakt tussen gedrag van een lagere orde, zoals bijvoorbeeld het reproduceren van kennis, en gedrag van een hogere orde, zoals bijvoorbeeld het schrijven van een essay (Roelofs & Visser, 2011; Pennewaard, 2011). Bij het toetsen van luisteren, lezen, gesprekken voeren, etc. gaat het vrijwel altijd om gedrag van een hogere orde. Werkt men met het ERK, dan ligt het voor de hand om bij het kiezen van gedragscategorieën uit te gaan van de schalen die zijn beschreven. Daarbij moeten echter wel keuzes worden gemaakt. Het aantal schalen is nu eenmaal te groot om ze allemaal apart te bevragen.

De schalen van het ERK kunnen worden opgesplitst in drie hoofdcategorieën:

- schalen voor vaardigheden zoals bijvoorbeeld spreken of schrijven;
- schalen voor strategieën zoals bijvoorbeeld planning of compensatie;
- schalen voor ondersteunende vaardigheden zoals bijvoorbeeld grammaticale correctheid of coherentie en cohesie.

Voor de vaardigheden geldt vervolgens dat ze zijn opgesplitst in subvaardigheden. Deze zijn in tabel 4.1 weergegeven. Door de subvaardigheden als uitgangspunt te nemen van toetsconstructie kan men zorgen voor een zo groot mogelijke dekking van de schalen uit de drie voorgaande hoofdcategorieën. Dat komt omdat bij het bevragen van de subvaardigheden de schalen voor strategieën en ondersteunende vaardigheden impliciet worden meegenomen. Om bijvoorbeeld informele gesprekken te begrijpen op niveau B1 moeten leerlingen immers ook strategieën beheersen op dat niveau. Datzelfde geldt voor een ondersteunende vaardigheid zoals bijvoorbeeld woordenschat.

Tabel 4.1: Overzicht van subvaardigheden volgens de publicatie 'Taalprofielen'

Vaardigheid	Subvaardigheden
Luisteren	<ul style="list-style-type: none"> • Gesprekken tussen moedertaalsprekers verstaan • Luisteren als lid van een live publiek • Luisteren naar aankondigingen en instructies • Luisteren naar tv, video- en geluidsopnames
Lezen	<ul style="list-style-type: none"> • Correspondentie lezen • Oriënterend lezen • Lezen om informatie op te doen • Instructies lezen
Spreken	<ul style="list-style-type: none"> • Monologen • Een publiek toespreken
Schrijven	<ul style="list-style-type: none"> • Correspondentie • Aantekeningen, berichten, formulieren • Verslagen en rapporten • Vrij schrijven
Gesprekken voeren	<ul style="list-style-type: none"> • Informele gesprekken • Bijeenkomsten en vergaderingen • Zaken regelen • Informatie uitwisselen

Uit de tabel kunnen we afleiden dat een ERK-toets voor schrijven in zou moeten gaan op vier verschillende subvaardigheden, te weten: 1) aantekeningen maken, berichten opstellen, formulieren invullen, 2) verslagen en rapporten, 3) vrij schrijven, en 4) correspondentie. Dit houdt in dat er voor elk van de vijf vaardigheden een aparte toetsmatrijs gemaakt moet worden. Overigens zijn er veel opdrachten te bedenken waarmee meerdere subvaardigheden worden gedekt. Bij een gespreksopdracht waarbij 'zaken moeten worden geregeld' is er bijvoorbeeld ook al snel sprake van 'informatie-uitwisseling'. Zie voor meer informatie de publicatie 'Taalprofielen' (2004) op www.erk.nl > Docent > De niveaubeschrijvingen.

4.2.3 Een concreet voorbeeld van een toetsmatrijs

Een voorbeeld van een toetsmatrijs voor schrijven wordt hierna gegeven. In de rijen zijn de subvaardigheden weergegeven. Deze zijn in dit voorbeeld nader uitgewerkt in *can do statements* voor niveau A2, met uitzondering van de subvaardigheid Verslagen en essays, omdat de opstellers van het ERK geen *can do statements* hebben kunnen bedenken die geschikt waren voor leerlingen op dit niveau. Per subvaardigheid worden meestal meerdere *can do statements* genoemd. Hieruit kan men bij het samenstellen van een toets een keuze maken. Het is dus niet verplicht om alle *can do statements* die bij een subvaardigheid genoemd worden, te bevragen. Ten slotte zijn ten behoeve van de domeinen verschillende kolommen aangemaakt. In het witte gedeelte van de tabel kan worden aangegeven welke opgave betrekking heeft op welke subvaardigheid en welk domein. Wil men een andere vaardigheid toetsen, of een ander niveau, dan moet de tabel worden aangepast. In hoofdstuk 5 van de publicatie 'Toetsen en beoordelen met het ERK' is voor elke vaardigheid op elk niveau een toetsmatrijs weergegeven waarbij de witte cellen leeg zijn gelaten zodat elke docent deze naar believen kan invullen.

Tabel 4.2: Voorbeeld van een toetsmatrix voor Schrijven op niveau A2

Vaardigheid: Schrijven		ERK-niveau: A2				
Subvaardigheid	Can do statement	Dagelijks leven	Publiek domein	Opleiding	Werk	Aantal opg. per subvaardigheid
Correspondentie	<ul style="list-style-type: none"> • Kan een eenvoudig persoonlijk briefje of e-mail schrijven. • Kan aan een contactpersoon een korte bevestiging (e-mail of fax) van gemaakte afspraken schrijven. • Kan aan een eenvoudige chatsessie deelnemen. 	Opgave 2			Opgave 3	2
Aantekeningen, berichten, formulieren	<ul style="list-style-type: none"> • Kan standaardformulieren invullen. • Kan eenvoudige notities en aantekeningen maken voor zichzelf. • Kan eenvoudige notities en aantekeningen maken voor anderen. • Kan korte, eenvoudige berichten schrijven over zaken van direct belang. 		Opgave 4			1
Vrij schrijven	<ul style="list-style-type: none"> • Kan in korte, eenvoudige zinnen vertrouwde zaken beschrijven. • Kan in korte, eenvoudige zinnen een persoon beschrijven. • Kan kort en eenvoudig een gebeurtenis of een ervaring beschrijven. 			Opgave 1		1
Aantal opgaven per domein		1	1	1	1	Eind-totaal: 4

De uitwerking van de toetsmatrijs in tabel 4.2 voorbeeld kan leiden tot de volgende schrijftoets:

Toets schrijfvaardigheid Frans, niveau A2

Deze toets bestaat uit vier opgaven. Bij elke opgave moet je een kort stukje tekst schrijven. Hiervoor gebruik je het antwoordblad dat bij deze toets hoort.

Opgave 1

Can do statement: Kan in korte, eenvoudige zinnen vertrouwde zaken beschrijven (vrij schrijven)

Domein: opleiding

Je volgt een cursus Frans in het buitenland. Tijdens de eerste les moeten alle cursisten iets over zichzelf schrijven. Schrijf een korte paragraaf over jezelf waarin je je voorstelt. Vertel hoe oud je bent en uit welk land je komt. Leg ook uit waarom je graag Frans wilt leren.

Opgave 2

Can do statement: Kan een eenvoudig persoonlijk briefje of e-mail schrijven (correspondentie)

Domein: dagelijks leven

Je schrijft een e-mail naar een Franse vriendin om haar uit te nodigen tijdens de zomervakantie. Vermeld een datum en stel voor wat jullie tijdens de vakantie kunnen gaan doen.

Opgave 3

Can do statement: Kan een eenvoudig persoonlijk briefje of e-mail schrijven (correspondentie)

Domein: werk

Je bent op zoek naar een baantje tijdens de zomervakantie. Toevallig hebben je ouders goed contact met een Franse campingbaas. Hij is op zoek naar vakantiekrachten en heeft je gevraagd om in juli of augustus op zijn camping te komen werken. Schrijf de campingbaas een e-mail waarin je positief op zijn verzoek reageert. Vermeld hoeveel weken je kunt komen en stel een vraag over je salaris.

Opgave 4

Can do statement: Kan standaardformulieren invullen (aantekeningen, berichten, formulieren)

Domein: publieke sector

Je hebt in een Franse trein een koffer laten staan. Je kijkt op de site van de Franse spoorwegen en vindt een formulier voor het aangeven van vermiste voorwerpen. Vul dit formulier in (zie antwoordblad). Vermeld je persoonlijke gegevens, geef aan wat je mist en geef een beschrijving ervan (kleur en merk).

De ervaring leert dat het bedenken van opdrachten gespreid over domeinen en subvaardigheden tot een heel gevarieerd en aantrekkelijk geheel leidt. Bovendien zorgt men er op deze manier voor dat het taalmateriaal waar leerlingen mee in aanraking komen een goede afspiegeling vormt van wat ze in het dagelijks leven over de grens nodig hebben. Voor leesvaardigheid betekent dit bijvoorbeeld dat men zich niet kan baseren op artikelen uit tijdschriften en de krant, want die vallen alleen onder het publieke domein. Om een goede spreiding te krijgen, moeten ook teksten worden toegevoegd uit de andere domeinen, zoals bijvoorbeeld een zakelijke brief, een e-mail met pakinstructies voor een vakantiereis en een brochure voor een zomercursus. Dit vraagt over het algemeen iets meer speurwerk. De kunst is

daarbij om de belevingswereld van de leerling niet te veel uit het oog te verliezen. Want het is niet de bedoeling om leerlingen vanaf de eerste toets al te confronteren met leveringsvoorwaarden van kantoormeubilair of aanwijzingen van de belastingdienst. Een vacaturetekst waarin studenten worden opgeroepen om bij te verdienen in een restaurant is dan een goed alternatief.

4.3 Het maken van opdrachten bij productieve vaardigheden

Als eenmaal met behulp van de matrix is vastgesteld welke domeinen en subvaardigheden bevestigd gaan worden, kan men aan de slag met de opgaven zelf. Een ERK-opgave voor spreek-, gespreks- of schrijfvaardigheid bestaat uit een aantal elementen:

- **Niveau-aanduiding**

Het is van belang om de leerlingen te laten weten welk ERK-niveau van hen wordt verwacht en wat dit niveau inhoudt. Dit kan behandeld worden in de les.

- **Situatiebeschrijving**

In de situatiebeschrijving wordt informatie gegeven over de context en de rol van de gesprekspartners of het publiek waarvoor de leerling spreekt of schrijft. Dit is van belang omdat de leerlingen hun taalgebruik hierop moeten afstemmen. Tegen een Duitse politieagent die hen op de bon wil slingeren, zullen ze immers een andere toon moeten aanslaan dan tegen een Duitse buurjongen die ongevraagd hun fiets heeft geleend.

Bij het maken van situatiebeschrijvingen valt het meestal niet mee om de hoeveelheid informatie goed te doseren. Men is al gauw geneigd om er een heel gedetailleerd verhaal van te maken, met als gevolg dat leerlingen zich eerst door hele lappen tekst heen moeten werken voordat ze überhaupt aan spreken of schrijven toe komen. Probeer daarom opdrachten zo beknopt mogelijk te formuleren en alleen die informatie te geven die nodig is voor het maken van de opdracht.

- **Opdracht**

In de opdracht wordt informatie gegeven over wat de leerling moet doen, zoals informatie geven, overtuigen, argumenten aandragen, etc.

- **Input van de native speaker (optioneel)**

Een opgave kan voorzien worden van een stukje tekst in de doeltaal. Dat kan een artikeltje zijn waarop gereageerd moet worden of een vraag van een native speaker. Dit helpt de leerling om als het ware de knop om te zetten naar de situatie waarin de vreemde taal wordt gesproken. Daarnaast maakt het de opdracht vaak ook authentiek.

4.3.1 Opdrachten wel of niet in de doeltaal?

De opdracht en situatiebeschrijving kunnen zowel in het Nederlands als in de doeltaal gegeven worden. Als de doeltaal gebruikt wordt, raden we aan om het taalgebruik eenvoudig te houden. De vuistregel is dat het taalniveau van de opdracht één ERK-niveau lager moet liggen dan het niveau waarvoor de toets is bedoeld. Dit laatste om te voorkomen dat het begrijpen van de opdracht een obstakel vormt voor het spreken of schrijven in de doeltaal. Ook de input van de native speaker mag geen belemmering vormen voor het uitvoeren van de taak. Het gebruik van Nederlands in de opdracht en de situatiebeschrijving heeft als voordeel dat er geen vocabulaire wordt weggegeven. Maar het nadeel is dat het al gauw leidt tot vertaalopdrachten. Die worden door leerlingen en docenten vaak als kunstmatig ervaren. De vraag is bovendien of met een vertaalopdracht echt spreek- of schrijfvaardigheid wordt getoetst. Om vertalen te vermijden,

kan een deel van de informatie uit de opdracht verwerkt worden in de situatiebeschrijving of de input van de native speaker. De opdracht hoeft dan niet meer zo specifiek te zijn, waardoor er niet meer woordelijk vertaald hoeft te worden. Ter illustratie volgen hieronder twee opgaven voor gespreksvaardigheid. In de eerste opgave is er sprake van vertalen terwijl dat in de tweede opgave al een stuk minder is.

Gespreksopgave 1, niveau B1

Situatiebeschrijving

Gisteravond ben je naar een feest gegaan met een Engelssprekende vriend(in). Daar hebben jullie wat gedronken. Nu voeren jullie een gesprekje.

Opdracht leerling 1

Jij begint. Zeg dat je gesprekspartner er bleek uit ziet.

Opdracht leerling 2

Zeg dat je hoofdpijn hebt en vraag de ander om een aspirientje.

Gespreksopgave 2, niveau B1

Situatiebeschrijving

Gisteravond ben je naar een feest gegaan met een Engelssprekende vriend(in). Daar hebben jullie wat gedronken. Nu heeft één van jullie hoofdpijn en ziet er ook erg bleek uit. De ander heeft altijd aspirientjes op zak. Samen voeren jullie een gesprekje.

Opdracht leerling 1

Jij begint. Probeer erachter te komen wat er precies aan de hand is met je gesprekspartner.

Opdracht leerling 2

Reageer op de vragen van je gesprekspartner. In eerste instantie wil je niet vertellen wat er is. Uiteindelijk zeg je het toch. Probeer bij je gesprekspartner dan ook iets los te krijgen tegen de pijn.

4.3.2 Het nadeel van volledig open opdrachten

Om het vertalen tegen te gaan, wordt ook vaak gewerkt met open opdrachten. Een voorbeeld daarvan staat hieronder.

Schrijfpogave 1, niveau B2

Situatiebeschrijving

Je hebt het afgelopen jaar met je ouders rond de wereld gezeild. De Engelse zeilfederatie promoot dit soort activiteiten en heeft een deel van de website gewijd aan verslagen van een verblijf op zee. Jij wilt ook graag een bijdrage leveren.

Opdracht

Schrijf een verslag van minstens 250 woorden over je ervaringen van het afgelopen jaar.

Het is duidelijk dat een dergelijke open opdracht al snel leidt tot onvergelijkbare leerlingwerken. Onderstaande voorbeelden getuigen daarvan.

Leerling 1

“ ...Every day I woke up at eight o’ clock. Very often we were out of bread, so I usually ate cereals for breakfast. After that I made my homework. My teachers had showed me what to do, because I didn’t want to fail my exams once I would return to Vlaardingen...”

Leerling 2

“... After an exciting morning with lots of whale sightings, we came across three of these mammals who wanted to play with us. For a couple of hours we drifted nearby, trying to keep the prescribed 100 metres away, and these huge, 12 to 15 metre creatures put on an impressive display...”

In een lessituatie is het niet noodzakelijkerwijs een probleem dat een opdracht leidt tot leerlingwerken die inhoudelijk aanzienlijk verschillen, maar in een toetsituatie wel. Omdat leerling 2 ingewikkelder zaken beschrijft dan leerling 1, is de kans op fouten bij leerling 2 veel groter. Bij een beoordeling is men daardoor appels met peren aan het vergelijken. Het feit dat beide werkjes foutloos zijn, wil nog niet zeggen dat beide leerlingen de vormaspecten van de taal even goed beheersen. Want om een foutloos relaas te schrijven vergelijkbaar met dat van leerling 2, is een grotere taalbeheersing nodig dan voor een paar zinnen in de stijl van leerling 1. Aan de andere kant zou het weer onrechtvaardig zijn om leerling 1 af te straffen voor het feit dat zijn stukje inhoudelijk van mindere kwaliteit is. Hij heeft immers gewoon aan de opdracht voldaan. En wie weet is hij ook wel in staat om op het niveau van leerling 2 te schrijven, maar was hij niet zo gemotiveerd of wilde hij fouten vermijden. We zullen dus nooit vast kunnen stellen wat leerling 1 precies in huis heeft, omdat de opdracht hem niet verplicht tot een taalprestatie op een bepaald niveau.

Het is dus aan te bevelen om een open opdracht zoals hierboven te voorzien van een aantal richtlijnen zoals in onderstaand voorbeeld gebeurt. Op deze manier komen teksten van leerlingen zowel qua thema als qua niveau zoveel mogelijk overeen, zonder dat er sprake is van letterlijk vertalen of een opzet waarbij geen enkele creativiteit getoond kan worden.

Schrijfgave 2, niveau B2

Situatiebeschrijving

Je hebt het afgelopen jaar met je ouders rond de wereld gezeild. De Engelse zeilfederatie promoot dit soort activiteiten en heeft een deel van de website gewijd aan verslagen van een verblijf op zee. Jij wilt ook graag een bijdrage leveren, omdat je heel enthousiast bent over wat je hebt meegemaakt.

Opdracht

- Schrijf een verslag van minstens 250 woorden over je ervaringen van het afgelopen jaar.
- Begin met een beschrijving van de voorbereidingen.
- Geef aan met welke tegenslagen jullie onderweg te maken kregen.
- Beschrijf je contacten met de lokale bevolking uit verschillende landen.
- Leg uit hoe je onderweg toch lessen kon blijven volgen.
- Noem een onvergetelijke ervaring van het afgelopen jaar.
- Vertel wat de reis heeft bijgedragen aan de vorming van je karakter.
- Besluit met een opmerking die mensen enthousiast moet maken voor het zeezeilen.

4.3.3 Het uitlokken van een leerlingprestatie op een bepaald ERK-niveau

Om leerlingwerken op een bepaald ERK-niveau uit te lokken is met name de keuze van het onderwerp en de situatiebeschrijving van belang. Over het algemeen geldt dat een hoger ERK-niveau kan worden afgedwongen naarmate de complexiteit van het onderwerp toeneemt. Dat wil niet altijd zeggen dat het bij niveau B2 of C1 dan om heel abstracte zaken moet gaan zoals 'Misdaadpreventie in de detailhandel', of 'De gevolgen van globalisering voor derdewereldeconomieën'. Ook technische onderwerpen zoals uitleg over de werking van een mp3-speler kunnen leerlingen stimuleren om de vreemde taal op een hoog niveau te gebruiken. Voor de situatiebeschrijving geldt dat een formele setting (sollicitatiebrief, vergadering op een kantoor) uitingen van een hoger niveau uitlokt. Datzelfde geldt voor situaties waarin tact en nuance een rol spelen. Denk bijvoorbeeld aan het te woord staan van ontevreden klanten of het geven van feedback aan een goede vriend die plotseling een vervelende eigenschap heeft ontwikkeld.

De niveau-aanduiding boven de opdracht geeft leerlingen ook een idee van de prestatie die van hen wordt verwacht. Daarbij is het van belang om de leerling te informeren over de beoordelingscriteria. Leerlingen presteren namelijk beter als ze weten waarop ze worden afgerekend. Dit betekent overigens niet dat ze op de hoogte moeten worden gebracht van alle implicaties van het correctievoorschrift. Het is voldoende om te laten weten welke criteria een rol spelen bij de beoordeling.

Ten slotte kan de input van de native speaker de leerling een indicatie geven van het taalniveau dat van hem of haar wordt verwacht. Als een leerling een e-mail moet beantwoorden waarin Engels geschreven wordt van een bepaald niveau, dan spreekt vaak voor zich dat de reactie daar qua bewoordingen op moet aansluiten.

4.3.4 Is mijn B1-opgave jouw B1-opgave?

Hierboven noemden we het vergelijken van appels met peren. Dat is een gevaar dat steeds op de loer ligt bij het maken van productieve opdrachten. Zolang bovenstaande richtlijnen gevolgd worden en de hele klas dezelfde opdrachten maakt, treedt het probleem niet op. Maar docenten hebben vaak behoefte aan parallelopdrachten, bijvoorbeeld om te voorkomen dat parallelklassen dezelfde opdrachten krijgen en de inhoud aan elkaar doorvertellen. Daarnaast speelt natuurlijk het probleem van de vergelijkbaarheid over klassen heen. Want hoe garandeer je dat de B1-opdrachten die je als docent hebt gemaakt hetzelfde toetsen als de B1-opdrachten van een docent die 10 kilometer verderop lesgeeft?

Om er zeker van te zijn dat toetsen gelijkwaardig zijn, is het nodig om een groot aantal opgaven af te nemen op een representatieve steekproef van honderden leerlingen. Als we deze afnamegegevens met elkaar in verband brengen, kan worden vastgesteld of de opgaven wel allemaal dezelfde vaardigheid meten. Daarnaast kunnen we afleiden hoe moeilijk de ene opgave is ten opzichte van de andere. Met behulp van al deze gegevens kunnen we vervolgens toetsen samenstellen die volledig gelijkwaardig zijn aan elkaar. Deze werkwijze is voor de individuele docent helaas niet weggelegd. Absolute gelijkwaardigheid van toetsen is voor hem of haar dan ook geen haalbare kaart.

Er zijn echter drie tips waarmee een eerste stap kan worden gezet op weg naar vergelijkbaarheid.

Tip 1: Maak een toets die bestaat uit meerdere opgaven

Bij een toets die uit meerdere opgaven bestaat, is de kans is groter dat opgaven die wat makkelijk uitvallen gecompenseerd worden door opgaven die iets moeilijker zijn dan verwacht. Daarnaast is er nog een andere reden om te kiezen voor een toets die bestaat uit meerdere opgaven. Stel je wilt als docent toetsen of leerlingen een *can do statement* voor het schrijven van correspondentie op niveau B1 beheersen. Deze luidt: 'Kan in brieven of e-mails feitelijke zaken beschrijven en nieuwtjes uitwisselen'. Je zou dan de toets kunnen laten bestaan uit één brief waarin een leerling gevraagd wordt om te vertellen over een denkbeeldige verhuizing naar de andere kant van het land. Nadeel daarvan is dat de inhoud van de toets thematisch beperkt blijft. Leerlingen die goed met dit onderwerp uit de voeten kunnen, zijn daarbij duidelijk in het voordeel en dat is niet de bedoeling. Spreiding van opgaven over meerdere thema's heeft daarom de voorkeur.

Een leerling die de toets goed maakt, heeft dan ook werkelijk aangetoond dat hij of zij de *can do statement* beheerst en dat de goede prestatie niet te danken is aan toevallige kennis van dat ene onderwerp dat nou juist ter sprake kwam. Een derde reden om te kiezen voor een toets die uit meerdere opgaven bestaat, heeft te maken met betrouwbaarheid. Deze neemt toe, naarmate een toets bestaat uit meer metingen of meetmomenten. Dat is eenvoudig te verklaren. Met elk meetmoment komt informatie over het prestatieniveau van de leerling beschikbaar. Helaas bevat elk meetmoment ook altijd een meetfout. Dat houdt in dat men nooit zeker weet of de score die een leerling voor een opgave behaalt ook de score is die een leerling werkelijk verdient. Maar hoe groter het aantal meetmomenten, hoe meer informatie men kan verzamelen over een leerling, hoe betrouwbaarder de meting.

Met name bij schrijfvaardigheid vraagt het wat puzzelwerk om een toets uit meerdere opgaven te laten bestaan. Vooral naarmate een hoger niveau aan de orde is. Een *can do statement* voor de subvaardigheid Verslagen en essays op niveau B2 luidt bijvoorbeeld: 'Kan redelijk gedetailleerde verslagen maken'. Maar het maken van een verslag kost al snel minstens één lesuur. Een toets die bestaat uit een aantal van dit soort opgaven neemt al snel bijna een dag in beslag en dat is om praktische redenen meestal niet haalbaar. In plaats van een verslag van begin tot eind te laten schrijven, kan men leerlingen echter ook vragen om slechts één of twee

alinea's van het totaal te produceren. Op deze manier kan een groot deel van de zaken die een rol spelen bij het maken van een verslag aan de orde komen, zonder dat er al te veel tijd mee gemoeid is. Natuurlijk is het daarbij wel zaak om de leerlingen de tekstfragmenten te tonen waarop ze hun alinea's moeten laten aansluiten, zodat ze de opbouw van het geheel kunnen respecteren en vervolgens de juiste verbindingswoorden kunnen kiezen.

Tip 2: Neem erkende voorbeeldopgaven van een bepaald niveau als richtlijn voor de productie van eigen opgaven

Iedere docent die de *can do statements* van het ERK leest, interpreteert deze op zijn eigen manier. Bijna altijd spelen daarbij de leerlingen waaraan de docent lesgeeft en het gemak waarmee zij een taal leren een rol. Zo zien we regelmatig dat docenten Frans en Spaans al gauw vinden dat er een hoger ERK-niveau vereist is voor het maken van een bepaalde opgave dan docenten Duits of Engels. Daarom is het goed om bij de toetsconstructie een aantal voorbeelden van ERK-opgaven als uitgangspunt te nemen. Die voorbeelden maken concreet wat een bepaalde *can do statement* op een bepaald niveau inhoudt waardoor ze heel geschikt zijn om als richtsnoer te functioneren. Op de website www.erk.nl zijn voor elke taal van dit soort voorbeelden te vinden.

Tip 3: Betrek collega's van andere talen bij het maken van toetsen

Daarnaast is het van belang om bij het maken van opgaven samen te werken met collega's van andere talen. Zij zullen soms een heel ander beeld hebben bij het ERK-niveau dat vereist is voor bepaalde opdrachten dan de maker. Hierdoor ontstaat een vruchtbare discussie over wat er nou wel en niet geëist kan worden op niveau A1, A2, B1, etc. Het resultaat van die discussie is een toets die niet het product is van een niveau-inschatting van één docent, maar van een heel team, waardoor de vergelijkbaarheid van toetsen die op deze wijze worden geproduceerd, toeneemt.

Meer specifieke adviezen en voorbeelden met betrekking tot het toetsen van spreekvaardigheid gespreksvaardigheid zijn te vinden in hoofdstuk 6 van 'Toetsen en beoordelen met het ERK'.

4.4 Het beoordelen van productieve vaardigheden

Bij productieve vaardigheden is het correctievoorschrift even belangrijk als de toets zelf. Samen dienen ze om het ERK-niveau van de leerling te bepalen. Bij volledig open opdrachten ligt de niveaукwestie nog helemaal open en moet tijdens de beoordeling een niveau worden toegekend dat kan variëren van A1 tot C2. In de vorige paragraaf hebben we echter al aangegeven dat deze aanpak niet de voorkeur verdient, omdat leerlingen bij een volledig open opdracht niet altijd gestimuleerd worden om te laten zien wat ze kunnen. Beter is het om een toets voor te leggen die gekoppeld is aan een bepaald ERK-niveau, zodat vervolgens, met behulp van het correctievoorschrift, kan worden vastgesteld of de leerling dit niveau wel of niet beheerst.

4.4.1 Welke beoordelingscriteria moeten worden toegepast?

Bij het maken van een correctievoorschrift moet eerst worden bepaald welke beoordelingscriteria worden opgenomen. Het ERK maakt onderscheid tussen drie aspecten die daarbij als richtlijn kunnen dienen:

- vaardigheden zoals bijvoorbeeld spreken of schrijven;
- strategieën zoals bijvoorbeeld planning of compensatie;
- ondersteunende vaardigheden zoals bijvoorbeeld grammaticale correctheid of coherentie en cohesie.

Gesprekken voeren, spreken en schrijven en de ondersteunende vaardigheden hebben betrekking op het product dat de leerling aflevert. Strategieën zeggen iets over het proces, dat wil zeggen de manier waarop de leerling te werk gaat om tot een bepaald product te komen. Voor het beoordelen hiervan zou men eigenlijk in het hoofd van de leerling moeten kunnen kijken. Zoiets is onder andere mogelijk met een hardop-denken onderzoek. Omdat dit over het algemeen niet tot de praktijk van het talenonderwijs in het voortgezet onderwijs behoort, beperken we ons hier tot de beoordeling van vaardigheden en ondersteunende vaardigheden.

Bij het beoordelen van productieve activiteiten kan eenvoudigweg worden nagegaan in hoeverre de leerling erin is geslaagd om de opdracht (die als het ware een uitwerking is van een *can do statement* op een bepaald niveau) uit te voeren. Het toekennen van scores aan ondersteunende vaardigheden is minder eenvoudig, omdat eerst moet worden bepaald welke ondersteunende vaardigheden in de beoordeling worden meegenomen.

Het ERK onderscheidt er 13 die weer toebedeeld kunnen worden aan drie hoofdcategorieën:

- linguïstische competenties;
- sociolinguïstische competenties;
- pragmatische competenties.

Hieronder worden alle ondersteunende vaardigheden in een tabel weergegeven.

<p>Linguïstische competenties</p> <ul style="list-style-type: none"> • algemeen linguïstisch bereik (taalbeheersing in het algemeen) • bereik van de woordenschat • beheersing van de woordenschat • grammaticale correctheid • fonologische beheersing (uitspraak en intonatie) • orthografische beheersing (spelling, leestekens, lay-out en alinea-indeling)
<p>Sociolinguïstische competenties</p> <ul style="list-style-type: none"> • sociolinguïstische trefzekerheid (beheersing taalregister in verschillende sociale contexten)
<p>Pragmatische competenties</p> <ul style="list-style-type: none"> • flexibiliteit (aanpassen aan situatie en gesprekspartner) • beurten nemen (aan het woord komen en blijven) • thematische ontwikkeling (hoofd- en bijzaken scheiden) • coherentie en cohesie (verbanden en samenhang) • gesproken vloeiendheid • propositionele nauwkeurigheid (nauwkeurigheid van de weergegeven informatie)

4.4.2 Hoe moet gescoord worden?

Als is vastgesteld welke categorieën in het correctievoorschrift worden meegenomen, moet nagedacht worden over de manier van scoren. Het ligt misschien voor de hand om voor elke categorie de hoeveelheid fouten te tellen. Onderzoek heeft echter aangetoond dat deze gedetailleerde aanpak minder betrouwbaar is dan globaal beoordelen (Melse, 1990). Bij globaal beoordelen worden geen fouten geteld, maar wordt op basis van het algemene beeld van een tekst een oordeel gegeven (zie de voorbeelden op de volgende pagina).

Een andere vraag die beantwoord moet worden is hoeveel punten er per categorie maximaal moeten worden toegekend. De ervaring leert dat meer dan zeven schaalpunten in de praktijk

niet goed werkt. Het lukt de beoordelaars dan niet meer om een goed onderscheid te maken tussen een prestatie die bijvoorbeeld zeven punten waard is en een prestatie die acht punten verdient (Kuhlemeier, 2011).

Als per categorie de lengte van de scoreschaal is vastgesteld, moet worden bepaald hoe de schaalpunten worden omschreven. Dat kan op drie verschillende manieren:

- normatief
- descriptief
- met behulp van ankers

Hieronder volgt een schrijfgave op niveau A2 en het werk van een leerling die deze opdracht heeft gemaakt. Vervolgens zullen we laten zien hoe de verschillende schalen worden gedefinieerd en wat dit betekent voor de beoordeling van deze leerling.

Schrijfgave 1, niveau A2

Situatiebeschrijving

Je bent met vrienden op vakantie geweest in Duitsland. Jullie verbleven op 'Campingplatz Kinzigtal', een kampeerterrein naast de boerderij van de heer Tillmanns. De laatste nacht was er sprake van noodweer. Gelukkig mochten jullie in de boerderij slapen. Na terugkomst in Nederland besluit je om de heer Tillmanns een kaartje te sturen.

Opdracht

- Stel jezelf voor.
- Geef aan wanneer je overnacht hebt op het kampeerterrein.
- Herinner de heer Tillmanns aan het noodweer.
- Bedank hem voor zijn hulp en sluit af.

Uitwerking leerling 1

*Sehr geehrter Herr Tillmanns,
Mein Name ist Kiki Burg. In den Sommerferien von dieses Jahr waren ich und meine Freunden auf Ihren Campingplatz Kinzigtal. Das Wetter war ganz schlecht. Wir mochten von Ihnen in ihren hof übernachten. Wir fanden es wirklich sehr nett von Ihnen. Viele Dank für Ihren Hilfe und wir haben ein kleines Geschenk zugefügt. Nochmals viele Dank für Ihren Hilfe.*

*Mit freundlichen Grüßen,
Kiki Burg.*

Normatieve schalen

Bij een normatieve schaal worden omschrijvingen gebruikt als: uitstekend, matig, voldoende, onvoldoende, juist, correct en acceptabel. Het grote voordeel hiervan is de snelle en eenvoudige beoordeling (Kuhlemeier, 2011). Normatieve schalen hebben echter ook een nadeel omdat elke docent zijn eigen interpretatie heeft van wat uitstekend, matig, of zwak is. Dit maakt de beoordeling subjectief. Hieronder staat een voorbeeld van een normatieve schaal voor het aspect 'inhoud' en wordt met behulp van deze schaal een beoordeling gegeven van het werk van Kiki Burg.

Voorbeeld normatieve schaal voor het aspect 'inhoud' bij schrijfopgave 1, niveau A2

0 punten = zwak
1 punt = redelijk
2 punten = goed

Voorbeeldbeoordeling van leerling 1 met bovenstaande schaal

Kiki Burg verdient 2 punten voor inhoud, omdat ze de opdracht goed heeft uitgevoerd.

Descriptieve schalen

Een descriptieve schaal wordt ook wel een rubriek genoemd. In descriptieve schalen wordt bij de schaalpunten een inhoudelijke beschrijving gemaakt. Ook bij deze schaal is het lastig om subjectiviteit te vermijden, omdat het bij taaltoetsing soms lastig is om een eenduidige beschrijving te geven. Voor wat betreft het aspect 'inhoud' valt het vaak nog wel mee, maar het vastleggen van schaalpunten voor bijvoorbeeld 'beheersing van de woordenschat' of 'grammaticale correctheid' is al een stuk lastiger. Als bijvoorbeeld beoordeeld moet worden of een tekst getuigt van een goede beheersing van de woordenschat, dan zullen niet alle docenten dat op precies dezelfde manier doen, omdat niet iedereen hetzelfde beeld heeft bij 'een goede beheersing van de woordenschat'. Hieronder staat een voorbeeld van een descriptieve schaal voor het aspect 'inhoud' en wordt met behulp van deze schaal een beoordeling gegeven van het werk van Kiki Burg.

Voorbeeld descriptieve schaal voor het aspect 'inhoud' bij schrijfopgave 1, niveau A2

0 punten = Twee of meer aspecten van de opdracht ontbreken in de tekst
1 punt = Eén aspect van de opdracht ontbreekt in de tekst
2 punten = Alle aspecten van de opdracht zijn in de tekst terug te vinden

Voorbeeldbeoordeling van leerling 1 met bovenstaande schaal

Kiki Burg verdient 2 punten voor inhoud, omdat alle vier aspecten van de opdracht in haar tekst zijn terug te vinden.

Productschalen

Een derde manier om schaalpunten te omschrijven, is met behulp van ankers. Een anker is een tekst van een leerling die het 'niveau' van het schaalpunt illustreert. Bij spreken of gespreksvaardigheid kunnen ankers uit een audio- of video-opname bestaan. Een schaal met ankers wordt ook wel een productschaal genoemd. Het beoordelen met een productschaal kan men zien als een sorteertaak. De beoordelaar beoordeelt elke tekst door deze te vergelijken met de ankers van de beoordelingsschaal. Productschalen zijn heel geschikt om de betrouwbaarheid objectiever te maken (Kuhlemeier, 2011). Een nadeel is echter dat er voor elke opgave weer een nieuwe productschaal ontworpen moet worden, waardoor deze aanpak erg tijdrovend is. Hieronder staat een voorbeeld van een productschaal voor het aspect 'inhoud' en wordt met behulp van deze schaal een beoordeling gegeven van het werk van Kiki Burg.

Voorbeeld productschaal voor het aspect 'inhoud' bij schrijfpogave 1, niveau A2

- 0 punten = De tekst sluit minder goed aan op de opdracht dan onderstaand anker
1 punt = De tekst sluit even goed aan op de opdracht als onderstaand anker
2 punten = De tekst sluit beter aan op de opdracht dan onderstaand anker

Anker

*Hallo herr Tillmanns,
Kennst du mir noch? Ich bin Debby Koens. Diese Monat habe ich bei du geschlafen weil ich in meine Tent nicht mir schlafen konnte von Angst. Ich ware in Kinzigtal in mei. Meine Freunden und ich wollten Sie danken für deine freundlichkeit und vielleicht bis nächstes Jahr.*

*Mit freundlichen grußen,
Debby Koens*

Voorbeeldbeoordeling van leerling 1 met bovenstaande schaal

Kiki Burg verdient 2 punten voor inhoud, omdat haar tekst inhoudelijk beter aansluit op de opdracht dan het anker (dat wil zeggen, het werk van Debby Koens).

Bij het opstellen van een correctievoorschrift moet ook nagedacht worden over het gewicht van de verschillende beoordelingscriteria. Stel dat een groep docenten de volgende beoordelingscriteria wil laten meewegen voor schrijfvaardigheid:

- communicatieve effectiviteit (slaagt de leerling erin de opdracht uit te voeren);
- beheersing van de woordenschat;
- grammaticale correctheid;
- orthografische beheersing (spelling, leestekens, lay-out en alinea-indeling).

Als vervolgens elk criterium wordt beoordeeld op een schaal van 0 tot en met 3 punten, heeft dat tot gevolg dat er per opgave 12 punten kunnen worden behaald. Men moet zich daarbij wel realiseren dat drie kwart van de punten betrekking heeft op vormaspecten (woordenschat, grammatica en orthografische beheersing) en slechts één kwart van het aantal punten op het overbrengen van de boodschap (communicatieve effectiviteit). Dat is niet voor elk ERK-niveau een wenselijke verhouding. Op niveau A1 bijvoorbeeld mogen nog heel veel taalfouten worden gemaakt. Bij dat niveau past geen correctievoorschrift dat de nadruk legt op vormaspecten. Dat geldt in zekere mate ook voor niveau A2. Over het algemeen kunnen we stellen dat naarmate het ERK-niveau stijgt een toename van aandacht voor vormaspecten geoorloofd is. In hoofdstuk 7 van de publicatie 'Toetsen en beoordelen met het ERK' wordt aangegeven hoe het aandeel van de verschillende beoordelingscriteria in het eindoordeel beïnvloed kan worden.

4.5 Het maken van opdrachten voor receptieve vaardigheden

Voordat men als constructeur toekomt aan het stellen van vragen bij receptieve vaardigheden, moeten er eerst (luister)teksten worden gezocht op het ERK-niveau waarvoor de toets is bedoeld. Omdat er gespreid moet worden over vier verschillende domeinen en *can do statements*, is daarbij enig speurwerk vereist. Het resultaat is echter een heel gevarieerd aanbod aan teksten, representatief voor het tekstaanbod waarmee leerlingen over de grens worden geconfronteerd.

4.5.1 Het constructieproces

Bij productieve vaardigheden is het mogelijk om een constructieteam aan te stellen dat ERK-

opdrachten maakt die in te zetten zijn voor de verschillende talen. Zo is de schrijfpdracht uit hoofdstuk 6 van voornoemde publicatie bruikbaar voor elke taal. Als meerdere secties samenwerken, zorgen de verschillende invalshoeken ten aanzien van het ERK voor een vruchtbare discussie over wat de ERK-niveaus nou precies inhouden. Natuurlijk moet vermeden worden dat leerlingen met steeds dezelfde opdrachten worden geconfronteerd. Een kleine overlap hoeft echter geen probleem te zijn, omdat A2 opdrachten voor Engels of Duits in lagere klassen zullen worden ingezet dan voor Frans of Spaans.

Voor receptieve vaardigheden behoort een constructieteam dat bestaat uit docenten voor verschillende talen helaas niet tot de mogelijkheden. Dit zou immers inhouden dat de uitgangstekst van de ene taal naar de andere vertaald moet worden. Een operatie die al snel veel tijd gaat kosten en die niet altijd de gewenste resultaten oplevert, omdat teksten vaak landgebonden zijn. Wat wel tot de mogelijkheden behoort, is het screenen van opgaven van andere secties. In de eerste plaats is het altijd van belang om opgaven door een ander te laten commentariëren, omdat elke constructeur wel eens wat over het hoofd ziet. In de tweede plaats leidt het screenen en bespreken van opgaven vaak ook tot discussies over wat nou eigenlijk de kenmerken zijn van een opdracht uit bijvoorbeeld het A2-domein. Daarmee komen de verschillende secties meer op één lijn te zitten en wordt voorkomen dat bepaalde interpretaties van het ERK een eigen leven gaan leiden.

Daarnaast is het aan te bevelen om erkende voorbeelden van ERK-opgaven voor een bepaald niveau als richtlijn te nemen voor constructie. Op de site www.erk.nl staat een aantal van deze voorbeelden voor Duits, Engels, Frans en Spaans. Daarnaast zijn op de site van Cito, www.cito.nl > Voortgezet onderwijs > Taaltoetsen bij het ERK, voor lees- en luistervaardigheid opgaven te vinden per ERK-niveau voor Frans, Duits en Engels.

4.5.2 De vraagstelling

Bij de receptieve vaardigheden wordt het ERK-niveau van de opgave bepaald door een combinatie van (luister)tekst en vraag. Meestal verkrijgt men de beste resultaten als het niveau van de vraag aansluit bij dat van de tekst. Natuurlijk is het mogelijk om bij een begrijpelijk stukje tekst een vraag te stellen over dat ene moeilijke woord dat erin staat. De vraag is echter of deze manier van bevragen aansluit bij het ERK. De *can do statements* bij lees- en luistervaardigheid hebben immers betrekking op het begrijpen van teksten in een bepaalde context. Dat pleit voor het stellen van functionele vragen, dat wil zeggen: vragen over het hoe, wat, waarom, etc. Daarvoor is het nodig om het niveau van de vraag te laten aansluiten op dat van de tekst.

Met name bij leestoetsen komt het vaak voor dat vragen worden gesteld over verbanden, zoals bijvoorbeeld:

- Naar welke personen verwijst 'they' in regel 36?
- Hoe verhoudt de eerste alinea zich tot de tweede?
- Welk voegwoord past op de lege plek in de derde alinea?

Dit soort vragen past minder goed bij het ERK, omdat er een tekstanalyse voor gepleegd moet worden en het naar behoren kunnen uitvoeren van een tekstanalyse is nu eenmaal geen *can do statement* in het ERK. Daarom heeft het de voorkeur om vooral functionele vragen te stellen in een ERK-toets. Bijvoorbeeld:

- Welke bijlagen moet je meesturen met het aanmeldingsformulier voor de cursus?
- Wat is het nadeel van het veelvuldig gebruik van sociale media volgens de tekst?
- Hoe zorgt men ervoor dat fruit precies rijp is op het moment dat het in de winkel ligt?

De lengte van de toets

Over het algemeen geldt: hoe meer meetmomenten (scorepunten), hoe betrouwbaarder de

meting. Maar vanuit praktische overwegingen kan de toets maar een beperkte lengte hebben. De vraag die dan bij iedereen op komt luidt vervolgens: hoeveel scorepunten moet een toets dan wel niet hebben om te voldoen aan de minimumeisen van betrouwbaarheid. Om daar antwoord op te geven moet eerst duidelijk zijn wat het doel is van de toets. Voor voortgangscntrole gelden namelijk minder strenge eisen dan voor bijvoorbeeld het eindexamen. Laten we uitgaan van voortgangscntrole, omdat daar in het voortgezet onderwijs meestal sprake van is. Volgens richtlijnen van de COTAN, de Commissie Testaangelegenheden Nederland van het NIP (Nederlands Instituut voor Psychologen), zie www.psynip.nl > Wat doet het NIP, wordt daarvoor een minimale betrouwbaarheidscoëfficiënt geëist van 0,70 (Evers et al., 2009). De ervaring leert dat er bij lees- en luistervaardigheid voor de moderne vreemde talen al gauw 40 scorepunten nodig zijn om een toets te maken die voldoende betrouwbaar is.

Literatuur over het maken van open en gesloten vragen

Over het maken van open en gesloten vragen is al veel geschreven wat ook voor het maken van lees- en luistertoetsen van toepassing is. Voor meer informatie over het maken van open en gesloten vragen zie hoofdstuk 6 en 7 van het algemene deel van Toetsen op School.

4.6 Het normeren van ERK-toetsen

Als een leerling een ERK-toets heeft gemaakt, behaalt deze leerling een score. Maar een score zegt op zich niets over hoe een leerling presteert ten aanzien van het ERK. Daarvoor is het nodig om de toets te normeren. Voor ERK-toetsen betekent dit dat er voor elke toets bepaald moet worden vanaf welke score leerlingen een bepaalde standaard realiseren, in dit geval een bepaald ERK-niveau. In deze paragraaf presenteren we een methode die het voor de individuele docent mogelijk maakt om een ERK-toets te normeren.

Hiervoor hebben we aangegeven hoe je als docent productieve en receptieve opgaven kunt maken volgens de richtlijnen van het ERK. Voor een A2-toets betekent dat bijvoorbeeld dat er opgaven zijn gemaakt die passen bij het domein A2. Binnen dat domein zijn echter niet alle opgaven even moeilijk. Sommige opgaven zullen op beheersingsniveau zijn geformuleerd. Daarbij ga je er als constructeur vanuit dat 80 procent van de leerlingen die zich op niveau A2 bevindt de opgaven goed maakt. Andere A2-opgaven zullen passen bij het instructieniveau. Dat wil zeggen dat leerlingen op niveau A2 tussen de 50 en de 80 procent kans hebben om deze opgaven goed te maken. De leerling beheerst het niveau immers nog niet, maar zit nog in de fase waarin hij of zij aan het leren is. Vandaar de term 'instructieniveau'. Ten slotte kan het zijn dat bepaalde opgaven op 'frustratieniveau' blijken te zitten. Bij een A2-opgave op frustratieniveau hebben A2-leerlingen minder dan 50 procent kans om de opgaven goed te maken. Het blijkt een lastige kwestie om vooraf in te schatten hoe moeilijk een opgave precies zal zijn voor de doelgroep. Zelfs ervaren toetsconstructeurs kunnen zich hier nog wel eens bij vergissen. Opgaven waarvan van tevoren wordt gedacht dat ze voor leerlingen geen enkel probleem zullen opleveren, blijken soms heel ingewikkeld en andersom.

Als de toets eenmaal is samengesteld en afgenomen, is er voor elke leerling een score beschikbaar. Bij het normeren van een ERK-toets krijgen deze scores betekenis. Dat komt omdat er standaarden (of cesuren) worden vastgelegd die bepalen welke scores corresponderen met welk ERK-niveau. Gaat het om een toets op niveau B1, dan bepaalt de normering welke leerlingen zich op, onder of boven dat niveau bevinden. Gaat het bijvoorbeeld om een toets met behulp waarvan een uitspraak kan worden gedaan op meerdere ERK-niveaus (denk aan een zomercursus waarbij leerlingen van over de hele wereld arriveren om Engels te leren) dan geeft de normering aan welk scorebereik past bij welk ERK-niveau.

In hoofdstuk 9 van het algemene deel Toetsen op School worden drie methoden onderscheiden voor het bepalen van standaarden. Deze methoden zijn voor de individuele docent echter niet goed bruikbaar, omdat ze te kostbaar en tijdrovend zijn (methode van Angoff of de methode van de contrasterende groepen) of erg subjectief (intuïtieve methode). Daarom presenteren we hieronder een methode die lijkt op de methode van Angoff maar waarmee de individuele docent wel uit de voeten kan. Deze methode biedt docenten toch de mogelijkheid om verantwoording af te leggen over de keuze voor een bepaalde norm.

De methode houdt in dat je als docent uitgaat van een leerling die zich n t op een bepaald ERK-niveau bevindt. Laten we als voorbeeld een toets nemen die bedoeld is om vast te stellen of leerlingen niveau B1 hebben. Om deze toets te normeren moet een docent zich een voorstelling maken van een (fictieve) leerling die precies niveau B1 beheerst en niet al halverwege B2 is. Het is dus niet de bedoeling om uit te gaan van leerlingen uit een bepaalde klas en te denken: H4c zit ongeveer op niveau B1, dus ik neem deze klas als uitgangspunt. Als men zich eenmaal een beeld heeft gevormd van de minimale B1-leerling, dan moet de docent inschatten hoeveel scorepunten deze leerling zal halen op de toets. Bijvoorbeeld 20 van de 30 scorepunten. Op deze plek ligt dan vervolgens de cesuur. Dat wil zeggen dat leerlingen die minstens 20 scorepunten hebben gehaald, volgens de toets presteren op niveau B1. Hoe de leerlingen in de werkelijkheid presteren op de toets, is niet van invloed op de normering. Als blijkt dat meer dan de helft van de leerlingen van de docent uit bovenstaand voorbeeld onder de 20 punten scoort, dan is het niet de bedoeling dat de normering wordt versoepeld. Dat kan natuurlijk wel, maar dan geeft de cesuur niet langer aan of een leerling presteert op niveau B1 of niet. Dat komt omdat de door ons gekozen manier van normeren absoluut is. Absoluut normeren wil zeggen dat de toetsscores van leerlingen niet worden vergeleken met die van andere leerlingen maar met een toetsscore die als standaard is aangeduid.

4.6.1 ERK-toetsen en cijfers

Als de cesuur (grens tussen voldoende en onvoldoende) eenmaal is vastgelegd, kan een score-cijfertransformatie worden berekend. Met behulp daarvan kan elke score gekoppeld worden aan een cijfer. Voor ERK-toetsing is dit overigens geen absolute must. Soms is het voldoende om vast te stellen of leerlingen wel of niet op een bepaald ERK-niveau zitten. Heb je als docent eenmaal een standaard gezet en is het ERK-niveau van leerlingen alles wat je wilt weten, dan kan je daarmee volstaan. Leerlingen die hoger scoren dan de standaard zitten boven een bepaald ERK-niveau, leerlingen die even hoog scoren, presteren precies op dat ERK-niveau en leerlingen die lager scoren, zitten onder het getoetste ERK-niveau.

De praktijk in het voortgezet onderwijs is echter dat er voor toetsen cijfers worden gegeven. Docenten die een standaard hebben gezet en daarna de scores van leerlingen willen koppelen aan een cijfer kunnen het computerprogramma uit hoofdstuk 9 van het algemene deel van Toetsen op School gebruiken. In het programma moet de totale lengte van de toets worden ingevoerd en de cesuur (de toetsscore die leidt tot de laagste voldoende).

Daarnaast moet een keuze worden gemaakt voor het soort omzetting. Als het om ERK-toetsen gaat, raden we de optie 'lineair met knik' aan. Dit is een transparante omzetting die ervoor zorgt dat bij 0 scorepunten het cijfer 1 wordt toegekend en dat bij elk volgend scorepunt een iets hoger cijfer wordt gegeven. Bij de optie 'lineair' daarentegen wordt een vaste transformatie toegepast (de afstanden tussen de cijfers zijn onder en boven de cesuur steeds even groot) waardoor leerlingen soms bij 10 scorepunten nog een 1 krijgen. Dit is niet altijd makkelijk uit te leggen aan de klas.

5 Het meten van non-cognitieve onderwijsdoelen

5 Het meten van non-cognitieve onderwijsdoelen

Hans Kuhlemeier

Het onderwijs heeft naast de cognitieve ontwikkeling van leerlingen de opdracht de persoonlijke, sociale en morele ontwikkeling te bevorderen. Docenten bereiden leerlingen niet alleen voor op het vervolgonderwijs of toekomstige arbeid, maar ook op deelname aan de maatschappij. Sinds het einde van de vorige eeuw zien we daarbij een hernieuwde aandacht voor de sociale competenties van de leerlingen. De aanleiding is volgens deskundigen de toename van het aantal leerlingen met leer- en opvoedingsproblemen, de toenemende individualisering van de samenleving en het verdwijnen van de traditionele normen en waarden. Behalve voor de leerling als individu wordt sociale competentie ook van belang geacht voor het herstellen van de sociale cohesie in de samenleving en de voorbereiding op kritisch burgerschap. Deze pedagogische opdracht van het voortgezet onderwijs is onder meer vastgelegd in de Wet op het Voortgezet Onderwijs (WVO). Kort gezegd moet de school bijdragen aan het bereiken van de onderwijsdoelen op het gebied van sociale competentie en de sociale integratie en het actieve burgerschap van de leerlingen bevorderen. De mate waarin scholen daarin slagen, is tegenwoordig een belangrijke graadmeter voor de kwaliteit van het onderwijs in het sociale domein.

In het voortgezet onderwijs voeren cognitieve doelen de boventoon. Het onderwijs is vooral gericht op de leerprestaties van de leerlingen. De sterke nadruk op kennis, inzicht en vaardigheden is begrijpelijk. De kerndoelen voor de onderbouw en de examenprogramma's voor de bovenbouw hebben immers vrijwel uitsluitend betrekking op wat leerlingen moeten kennen en kunnen (zie ook hoofdstuk 2). In de kerndoelen zijn identiteit, houdingen, sociaal gedrag en waarden veel minder sterk vertegenwoordigd. Waarom zouden docenten daar dan in hun lessen aandacht aan besteden? Dit hoofdstuk beginnen we met het geven van een antwoord op deze vraag. Daarna bespreken we de belangrijkste non-cognitieve doelstellingen voor het voortgezet onderwijs. We maken daarbij een onderscheid in sociale competentie, burgerschap en morele ontwikkeling. Vervolgens gaan we in op de manier waarop docenten non-cognitieve onderwijsdoelen kunnen meten. Daarna bespreken we bestaande instrumenten voor het vaststellen van sociale competentie, burgerschap en morele ontwikkeling. Tot slot geven we enkele vuistregels en tips voor het zelf ontwikkelen van instrumenten voor het meten van non-cognitieve onderwijsdoelen.

5.1 Waarom aandacht besteden aan non-cognitieve doelen?

Waarom zouden docenten in hun lessen aandacht besteden aan de persoonlijke, sociale en morele ontwikkeling van leerlingen? Zijn de leerprestaties en de examenresultaten niet veel belangrijker? Hieronder staan vier antwoorden op de vraag waarom aandacht voor non-cognitieve onderwijsdoelen belangrijk kan zijn:

- Sommige non-cognitieve onderwijsdoelen zijn wettelijk verankerd, zodat de school daar aandacht aan móet besteden. Het zijn tevens aandachtspunten waar de Inspectie in het toezicht op de kwaliteit van het onderwijs op let. Het betreft onder meer sociale competentie, burgerschap en sociale integratie en het realiseren van een veilig schoolklimaat.

- De meeste vakken kennen ook specifieke non-cognitieve doelen, ook al zijn deze niet in de wet, de kerndoelen of de eindtermen genoemd. Denk bijvoorbeeld aan het bevorderen van het plezier in lezen. Wat als leerlingen de moeilijkste teksten kunnen lezen, maar na hun schooltijd geen boek meer willen aanraken? Andere voorbeelden zijn het wegnemen van angst voor wiskunde of het opwekken van de interesse voor techniek. Deze doelen zijn vaak vakoverstijgend, zoals het bevorderen van een democratische houding of het bevorderen van milieuvriendelijk gedrag.
- In de dagelijkse lespraktijk hebben docenten in toenemende mate te maken met probleemgedrag van leerlingen. Dat kan het lesgeven bemoeilijken of in sommige gevallen zelfs onmogelijk maken. Het voorkomen, verminderen of wegnemen van probleemgedrag wordt dan noodgedwongen een belangrijk onderwijsdoel.
- Aandacht voor non-cognitieve doelen in de lessen kan het leren van leerlingen ten goede komen. Leerlingen voelen zich beter en behalen hogere prestaties als de sfeer in de klas en op school goed is. Onderzoek heeft laten zien dat het welbevinden van leerlingen op school een belangrijke voorspeller is van succes op school, het leven van alledag en de latere beroepspraktijk. Bedenk echter dat een positieve houding of een sterke interesse voor het vak geen noodzakelijke voorwaarden zijn voor het behalen van hoge leerprestaties. Er zijn altijd leerlingen die hun uiterste best doen en hoge leerprestaties behalen zonder dat zij de docent, de lessen, het vak of de school bijzonder waarderen.

De onderwijsdoelen zijn vastgelegd in de kerndoelen. De volgende kerndoelen hebben betrekking op het sociale domein:

- De leerlingen leren zich te gedragen vanuit respect voor algemeen aanvaarde waarden en normen (kerndoel 37).
- De leerlingen leren respectvol om te gaan met verschillen in opvattingen van mensen (kerndoel 38).
- De leerlingen leren zorg te dragen voor de psychische gezondheid van henzelf en anderen (kerndoel 34).
- De leerlingen leren zich redzaam te gedragen in sociaal opzicht (kerndoel 35).

Deze kerndoelen zijn geformuleerd in globale termen en bieden nog weinig concrete handvaten voor docenten. Gelukkig bevat de preambule bij de kerndoelen een nadere toelichting. De preambule noemt zeven onderwijsdoelen in het sociale domein die voor alle leergebieden van belang zijn:

- goede werkhouding;
- reflectie op eigen handelen en leren;
- uitdrukken van eigen gedachten en gevoelens;
- respectvol luisteren en kritiseren van anderen;
- ontwikkelen van zelfvertrouwen;
- verwerven en verwerken van informatie;
- respectvol en verantwoordelijk omgaan met elkaar.

In de kerndoelen voor de onderbouw van het voortgezet onderwijs en de eindtermen voor de bovenbouw van havo/vwo zijn non-cognitieve doelen nagenoeg afwezig. De preambule die aan de eindtermen voor het vmbo voorafgaat, vormt in zekere zin een uitzondering op deze regel. Het bevat de volgende verwijzingen naar non-cognitieve onderwijsdoelen:

Werken aan vakoverstijgende thema's

1.1 het kennen van en omgaan met eigen en andermans normen en waarden

Leren communiceren

4.1 elementaire sociale conventies in acht nemen

4.2 overleggen en samenwerken in teamverband

4.4 verschillen in meningen en opvattingen benoemen en hanteren

4.6 omgaan met formele en informele afspraken, regels en procedures

Ook de examenprogramma's voor de praktijkvakken van het vmbo bevatten verwijzingen naar non-cognitieve onderwijsdoelen. Hieronder staan enkele voorbeelden van professionele, sociale en communicatieve houdingen en gedragingen uit het examenprogramma voor het vak verzorging.

VZ/K/5 Grondhouding

De kandidaat kan:

- respectvol omgaan met anderen
- omgaan met regels en zich houden aan afspraken
- klantvriendelijk handelen

ZWG/K/7 Sociaal pedagogisch werk

De kandidaat kan:

- met eigen emoties en die van anderen omgaan
- begrip tonen voor de gevoelens en wensen van de zorgvrager rond zorgverlening

VZ/K/3 Sociale en communicatieve vaardigheden

De kandidaat kan:

- in een praktijksimulatie sociale vaardigheden hanteren

VZ/K/11 Assisteren en verzorgen van gehandicapten

De kandidaat kan:

- mensen met een functiebeperking aanmoedigen tot deelname aan een activiteit

5.2 Sociale competentie, burgerschap en morele ontwikkeling

In deze paragraaf gaan we in op de drie belangrijkste domeinen van de non-cognitieve opbrengsten van het onderwijs, te weten sociale competentie, burgerschap en morele ontwikkeling. We merken op dat de onderscheiden domeinen niet wederzijds exclusief zijn en elkaar deels overlappen. Het zijn eerder verschillende invalshoeken dan scherp afgegrensde domeinen. Toch gebruiken we deze indeling omdat deze nauw aansluit bij de wetenschappelijke literatuur, de wettelijke vereisten en het toezicht van de Inspectie op de kwaliteit van het onderwijs in het sociale domein.

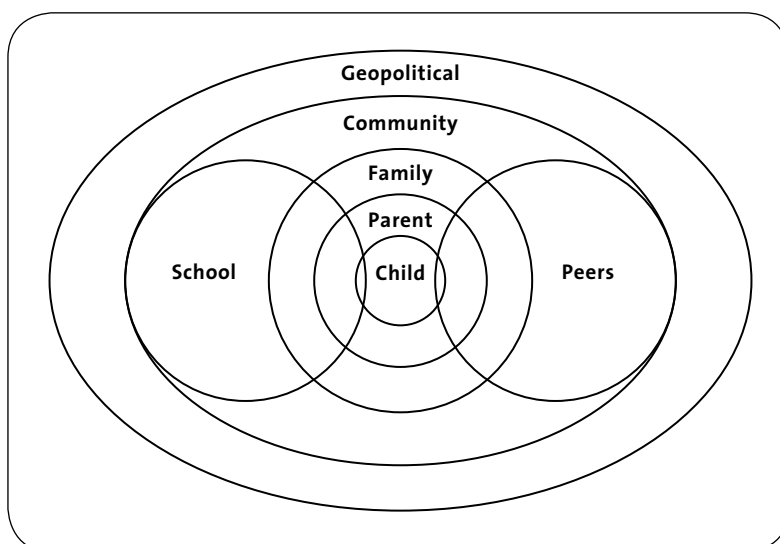
5.2.1 Sociale competentie

Sociale competentie is een containerbegrip dat zeer verschillend wordt ingevuld. Verschillende auteurs onderscheiden verschillende domeinen, ook wel onderdelen of componenten genoemd. Bijgevolg zijn er vele onderwijsdoelen die scholen onder de noemer van sociale competentie kunnen nastreven. In deze publicatie sluiten wij aan bij de veel geciteerde definitie van Ten Dam en Volman. Onder sociale competentie verstaan zij 'het geheel van kennis, vaardigheden en houdingen met betrekking tot het eigen sociale en sociaal-emotionele functioneren, de interacties tussen mensen en de invloed daarop van de inrichting van de samenleving en de waarden die daarin een rol spelen' (Ten Dam & Volman, 1999, pag. 23).

Deze definitie benadrukt dat voor sociaal competent handelen zowel kennis, vaardigheden als houdingen nodig zijn. Leerlingen moeten leren deze drie aspecten in onderlinge samenhang toe te passen in de omgang met zichzelf en anderen. Verder illustreert deze definitie dat sociale competentie zowel de persoonlijke, sociale als morele ontwikkeling van de leerling betreft. Sociale competentie wordt hiermee zowel vanuit het perspectief van de leerling zelf als dat van anderen benaderd. Het omvat naast een intrapersonlijke (de leerling zelf) ook een interpersoonlijke dimensie (de anderen). De intrapersonlijke dimensie heeft betrekking op de ontwikkeling van de eigen identiteit waarbij men onder meer kan denken aan zelfvertrouwen, zelfbeeld en zelfcontrole. De interpersoonlijke dimensie betreft in onderwijssituaties vooral het leren omgaan met medeleerlingen en de docent, door sommigen ook wel schoolburgerschap genoemd.

De hiervoor gegeven definitie van sociale competentie maakt ook duidelijk dat er nog een derde dimensie is, namelijk het maatschappelijke domein van de samenleving als geheel. Hier gaat het erom dat leerlingen leren een bijdrage te leveren aan de maatschappij, democratisch te handelen, maatschappelijk verantwoord te handelen, om te gaan met conflicten en om te gaan met verschillen tussen mensen.

Tot slot wijzen wij erop dat het vermogen tot competent handelen situatiegebonden is. Een onderscheid wordt daarbij gemaakt in sociale situaties zoals het gezin, de school, de buurt, de gemeente, de natuur, vrije tijd, Nederland, Europa en de wereld. Natuurlijk zijn de ouders als eerste verantwoordelijk voor de sociale vorming van hun kinderen. De school heeft een aanvullende verantwoordelijkheid waarbij docenten het goede voorbeeld kunnen geven door gewenst gedrag voor te doen. Het model van Bronfenbrenner, zie figuur 5.1, visualiseert hoe gezin, leeftijdsgenoten, school en samenleving samen de ontwikkeling van sociale competentie beïnvloeden (Sameroff, 2010).



Figuur 5.1: Het sociaal-ecologisch contextmodel van Bronfenbrenner

Om te kunnen bepalen hoe sociaal competent een leerling zich gedraagt, is een criterium nodig. Een veel genoemd criterium is dat het gedrag moet passen bij de ontwikkelingsfase waarin de leerling verkeert. Tot de ontwikkelingstaken van kinderen rond de twaalf jaar behoren onder meer het opbouwen en onderhouden van sociale contacten en vriendschappen (o.a. met leeftijdsgenoten), het verwerven van kennis en vaardigheden op school, het verwerken van sociale informatie (o.a. de bedoelingen en belangen van anderen kunnen inschatten) en het omgaan met autoriteiten en instanties. Sociaal incompetent kinderen zijn niet in staat om de eisen die deze ontwikkelingstaken aan hen stellen het hoofd te bieden. Uitgangspunt van het sociale competentiemodel is dat probleemgedrag een uiting is van onvermogen. Probleemgedrag ontstaat als leerlingen te weinig sociale competenties bezitten om de (ontwikkelings)taken waarvoor zij zich gesteld zien op een sociaal aanvaardbare en effectieve manier uit te voeren.

In het Toezichtkader van de Inspectie, zie ook hoofdstuk 3, is voor het bevorderen van de sociale competentie van de leerlingen een belangrijke plaats ingeruimd. De Inspectie toetst onder meer of de school een specifiek aanbod heeft om de sociale competenties van leerlingen te ontwikkelen (indicator 2.4). Anders dan in het primair onderwijs toetst de Inspectie nog niet of de sociale competenties van de leerlingen op een niveau liggen dat mag worden verwacht. Onder sociale competenties verstaat de Inspectie:

- sociaal-emotionele ontwikkeling (zelfbeeld, zelfstandigheid, regulering van emoties, etc.);
- sociale vaardigheden (samenwerken, zelfredzaamheid, omgaan met conflicten, etc.);
- houdingen/vaardigheden om in uiteenlopende situaties succesvol te functioneren (zelfsturing, omgaan met verschillen, democratisch handelen, sociale en maatschappelijke verantwoordelijkheid, moreel kunnen oordelen, etc.).

5.2.2 Burgerschap en sociale integratie

De laatste jaren heeft burgerschap een prominente plaats gekregen in het overheidsbeleid en de wet- en regelgeving. In de wettelijke regeling 'Bevordering actief burgerschap en sociale integratie' wordt burgerschap omschreven als 'de bereidheid en het vermogen deel uit te maken van de gemeenschap en daar een actieve bijdrage aan te leveren'. De opdracht aan het onderwijs is daarbij omschreven als 'Het onderwijs (a) gaat er mede van uit dat leerlingen opgroeien in een pluriforme samenleving, (b) is mede gericht op het bevorderen van actief burgerschap en sociale integratie, en (c) is er mede op gericht dat leerlingen kennis hebben van en kennismaken met verschillende achtergronden en culturen van leeftijdsgenoten'.

Deze opdracht is ingegeven door de afgenomen betrokkenheid tussen burgers onderling en tussen burgers en overheid. Daarbij zijn, zoals de toelichting bij de wettelijke regeling vermeldt, de plichten en rechten die horen bij burgerschap op de achtergrond geraakt. Dat wordt nog versterkt door de onbekendheid van veel ouders en kinderen met een allochtone achtergrond met de burgerschapstradities en -gebruiken van de Nederlandse samenleving. Door de ontwikkeling van burgerschap op te nemen in de opdracht aan scholen, kan 'een gemeenschappelijk en gedeeld perspectief van jonge mensen op de bijdrage die zij als burgers aan de samenleving kunnen leveren' worden bevorderd.

De opdracht aan de scholen om burgerschap en sociale integratie te bevorderen, heeft naast cognitieve aspecten ook sociale en morele aspecten. De school wordt geacht een leer- en werkomgeving te bieden waarin burgerschap en integratie zichtbaar zijn. Het onderwijsaanbod van de scholen moet erop gericht zijn dat leerlingen de basiswaarden van de democratische rechtsstaat respecteren en in praktijk brengen. Het gaat daarbij onder meer om: vrijheid van meningsuiting, gelijkwaardigheid, begrip voor anderen, verdraagzaamheid (ook wel tolerantie genoemd), autonomie, afwijzen van onverdraagzaamheid en afwijzen van discriminatie. Al deze aspecten hebben ook te maken met moraliteit of morele ontwikkeling. Onderstaande toelichting op de basiswaarden van de democratische rechtsstaat is ontleend aan de Inspectienota 'Toezicht op burgerschap en integratie'.

Vrijheid van meningsuiting. Dit betekent dat je mag zeggen of schrijven wat je denkt, of tegen de opvatting van anderen in mag gaan. Iedereen mag dus ook zijn of haar geloof uitdragen, of zijn of haar mening aan anderen voorhouden. Daarbij moet je je wel houden aan de wet.

Gelijkwaardigheid. Dit betekent dat mensen van gelijke waarde zijn. Daarbij maakt het niet uit wat hun denkbeelden zijn of wat ze geloven. Je hoeft niet te vinden dat die denkbeelden of gebruiken zelf waardevol zijn, maar wel dat mensen met andere denkbeelden en gebruiken niet minder waard zijn dan jij, of dan jouw groep.

Begrip voor anderen. Dit betekent dat je probeert te begrijpen waarom mensen of groepen bepaalde denkbeelden of gebruiken hebben: wat is de achtergrond daarvan en waarom is dat belangrijk voor een ander.

Verdraagzaamheid (ook wel tolerantie genoemd). Dit betekent dat je de mening of het gedrag van een ander accepteert, ook al ben je het er helemaal niet mee eens. En het betekent ook dat je ieder de ruimte wilt geven om zo'n mening of zulk gedrag te hebben. Natuurlijk moet iedereen zich daarbij wel aan de wet houden.

Autonomie. Dit betekent dat iedereen zelf kan bepalen wie hij/zij wil zijn en hoe hij/zij zijn/haar leven wil leiden. Ieder is dus bijvoorbeeld vrij om zelf te bepalen welke denkbeelden of welk geloof voor hem/haar belangrijk is. Daarbij moet je je wel houden aan de wet.

Afwijzen van onverdraagzaamheid. Onverdraagzaamheid (ook wel intolerantie genoemd) is het tegenovergestelde van tolerantie. Het betekent dat je vindt dat andere mensen of groepen, dingen waar jij het niet mee eens bent, niet zouden mogen denken of doen; en dat je het niet nodig vindt dat ieder de ruimte krijgt om zo'n mening of zulk gedrag te hebben.

Afwijzen van discriminatie. Discriminatie betekent dat mensen of groepen bij anderen achtergesteld worden, of dat je vindt dat er voor mensen met andere denkbeelden of gebruiken niet zoveel ruimte hoeft te zijn, of dat die denkbeelden of gebruiken misschien zelfs verboden moeten worden.

In het Toezichtkader vraagt de Inspectie van het Onderwijs van scholen dat zij aandacht schenken aan vier dimensies van burgerschap:

- het ontwikkelen van sociale vaardigheden bij de leerlingen;
- voorbereiding op deelname aan de Nederlandse samenleving;
- bevordering van de basiswaarden van de democratische rechtsstaat;
- het in praktijk brengen van burgerschap door de school zelf.

De scholen zijn vrij een eigen invulling aan burgerschap te geven en een eigen aanpak te ontwikkelen. Zo mag de school zelf weten in welke mate zij de leerlingen aanzet tot aanpassing, gehoorzaamheid en fatsoenlijk gedrag of juist tot mondigheid en een kritische houding.

De Inspectie beoordeelt de kwaliteit van aanbod op het terrein van burgerschap aan de hand van de volgende vier criteria:

- de school heeft een visie op de wijze waarop ze burgerschap en integratie wil bevorderen en geeft daar planmatig invulling aan;
- de school verantwoordt deze visie en de wijze waarop ze daar invulling aan geeft;
- de school evalueert of de doelen op het gebied van burgerschap en integratie worden gerealiseerd;
- de school stemt het aanbod mede af op risico's en ongewenste opvattingen, houdingen en gedragingen van leerlingen rond burgerschap en integratie.

De Inspectie intensificeert het toezicht als het onderwijs strijdig is met de basiswaarden (bijv. discriminatie) of als het burgerschap onder jongeren risico's loopt (bijv. vanwege extremistische ideeën).

De Inspectie constateert dat er op het terrein van sociale competentie, burgerschap en sociale integratie nog geen sprake is van een gericht vormingsaanbod. Wel geven scholen aan het sociaal-emotionele functioneren van leerlingen te bevorderen via de organisatie van het onderwijs en het pedagogische klimaat. Wat verstaat de Inspectie nu onder een gunstig schoolklimaat? In het Toezichtkader benadrukt de Inspectie dat het schoolklimaat gekenmerkt moet zijn door veiligheid en respectvolle omgangsvormen (kwaliteitsaspect 4).

De Inspectie toetst de kwaliteit van het schoolklimaat aan de hand van de volgende indicatoren:

- 4.1 De ouders zijn betrokken bij de school door de activiteiten die de school daartoe onderneemt.
- 4.2 De leerlingen voelen zich aantoonbaar veilig op school.
- 4.3 Het personeel voelt zich aantoonbaar veilig op school.
- 4.4 De school heeft inzicht in de veiligheidsbeleving van leerlingen en personeel en in de incidenten die zich op het gebied van sociale veiligheid op de school voordoen.
- 4.5 De school heeft een veiligheidsbeleid gericht op het voorkomen van incidenten in en om de school.
- 4.6 De school heeft een veiligheidsbeleid gericht op de afhandeling van incidenten in en om de school.
- 4.7 Het personeel van de school zorgt ervoor dat de leerlingen op een respectvolle manier met elkaar en anderen omgaan.
- 4.8 De mentoren stimuleren het welbevinden en de motivatie van de leerlingen.

Andere kwaliteitseisen die vooral met het klimaat in de klas te maken hebben, zijn het duidelijk uitleggen van de leerstof (indicator 5.1), het realiseren van een taakgerichte werksfeer (indicator 5.2) en het realiseren van actieve betrokkenheid van leerlingen bij de onderwijsactiviteiten (indicator 5.3).

5.2.3 Morele ontwikkeling

Moreel handelen heeft betrekking op hoe leerlingen zich daadwerkelijk gedragen in sociale situaties waarin zij een morele beslissing moeten nemen. Morele ontwikkeling betreft de ontwikkeling van het geweten ofwel het besef van goed en kwaad. De veronderstelling is dat leerlingen met een hogere morele ontwikkeling meer tot moreel handelen geneigd zijn dan leerlingen waarvan de morele ontwikkeling is achtergebleven. Onderwijs in sociale competentie wordt vaak ingeperkt tot het aanleren van sociale vaardigheden en het afleren van niet-sociale vaardigheden. De nadruk op sociale vaardigheden blijkt onder meer uit de populariteit van programma's voor het trainen van sociaal vaardig gedrag en het massale gebruik van tests waarmee leerlingen met gedragsproblemen opgespoord kunnen worden. Gebrekkige sociale vaardigheden zijn echter niet de enige oorzaak van antisociaal of agressief gedrag. De oorzaak kan ook gelegen zijn in een achterblijvende morele ontwikkeling of een gebrekkig inlevingsvermogen, ook wel empathie of 'perspectief nemen' genoemd. De laatste jaren is er ook meer aandacht voor de rol die afwijkende opvattingen spelen in het ontstaan van antisociaal gedrag. Het gaat daarbij om egoïstische denkwijzen en de neiging anderen de schuld te geven, het effect van het eigen agressieve gedrag op anderen te bagatelliseren of steeds maar uit te gaan van het ergste. Hieronder geven we een nadere toelichting op drie centrale begrippen van morele ontwikkeling, te weten moreel redeneren, inlevingsvermogen en afwijkende opvattingen van leerlingen.

Moreel redeneren

Het cognitieve aspect van morele ontwikkeling betreft het kunnen redeneren over morele vraagstukken in termen van rechtvaardigheid en het welzijn van anderen en het kunnen toepassen van morele criteria om gedrag te evalueren. Het gaat dan bijvoorbeeld om hoe leerlingen denken over de waarheid vertellen, stelen, pesten, beloftes nakomen en zich aan de wet houden. In de cognitieve benadering is een gebrekkige morele ontwikkeling vooral te wijten aan een gebrek aan kennis van normen en waarden en/of het onvermogen om deze kennis toe te passen in situaties die om een morele beslissing vragen. De belangrijkste vertegenwoordiger van de cognitieve benadering is Lawrence Kohlberg. In zijn fasenmodel onderscheidt hij drie niveaus van morele ontwikkeling met in totaal zes ontwikkelingsfasen. In de egocentrische fasen 1 en 2 heeft het kind de aangeleerde waarden en normen nog niet verinnerlijkt, in de fasen 3 en 4 wordt het belang van de familie en de samenleving in het moreel oordeel meegewogen en in de altruïstische fasen 5 en 6 wegen het persoonlijke geweten en de eigen principes het zwaarst. Mensen zouden deze fasen alle in dezelfde volgorde doorlopen, maar niet iedereen komt even ver. Hieronder geven we een korte beschrijving van de niveaus en fasen die voor leerlingen in het basis- en voortgezet onderwijs van belang zijn.

Niveau 1: Preconventioneel

Op het eerste ofwel preconventionele niveau van morele ontwikkeling oriënteert het jonge kind zich vooral op beloning en straf.

- Fase 1. Het kind is geheel en al gericht op externe autoriteit. Iets is goed omdat een autoriteit (moeder, vader, docent) dat zegt. Slecht is alles wat volgens de voorschriften niet mag en waar straf op staat. De maxime van de eerste fase is 'Wie de macht heeft, heeft ook het recht'.
- Fase 2. Het kind redeneert vooral vanuit eigen belangen en behoeften en hanteert ten opzichte van anderen een weegschaalmodel ('Ik doe iets voor jou omdat jij iets voor mij doet'). Het kind begint te onderhandelen waarbij vooral het behalen van persoonlijk voordeel telt. Slecht is alles wat het eigen belang niet dient. De maxime van de tweede fase is 'De ene hand wast de andere' of 'Voor wat hoort wat'. Maar ook 'Oog om oog, tand om tand'.

Niveau 2: Conventioneel

Op het tweede ofwel conventionele niveau staat het voldoen aan de regels, conventies en verwachtingen van de samenleving en autoriteiten centraal.

- Fase 3. Het kind redeneert vanuit dat gedachte dat het goed is om aardig voor anderen te zijn, zich om anderen te bekommeren, zich loyaal en betrouwbaar jegens anderen te gedragen, aan de verwachtingen van anderen te voldoen en zich aan de regels te houden. Goed is wat de anderen goed vinden, omdat iedereen zo is of doet en het kind niet anders wil zijn dan de anderen, ongeacht eventueel negatieve consequenties. De maxime van de derde fase is 'Wat gij niet wilt dat u geschiedt, doe dat ook een ander niet'.
- Fase 4. Het kind redeneert vanuit de gedachte dat het goed is om plichten tegenover de samenleving na te komen, de sociale orde in stand te houden en bij te dragen aan de maatschappij. Goed is wat de eigen groep en de meerderheid vindt of doet. Het leven wordt verwarrend en moeilijk als de gangbare groepsnormen worden overtreden. De maxime van de vierde fase is 'Rust is de eerste burgerplicht'.

Niveau 3: Postconventioneel

Op het derde of postconventionele niveau worden morele keuzes gemaakt op basis van zelf geaccepteerde morele gronden en universele morele principes.

Volgens Kohlberg is het preconventionele niveau kenmerkend voor het moreel redeneren van de meeste kinderen onder de negen jaar (en daarnaast voor vele criminelen). Het conventionele niveau is karakteristiek voor de meeste adolescenten en volwassenen. Het postconventionele niveau wordt slechts door weinig volwassenen bereikt, en dan meestal pas na de leeftijd van twintig jaar. Colby, Kohlberg, Gibbs en Lieberman (1983) volgden de morele ontwikkeling van 58 jongens van 10, 13 en 16 jaar gedurende een periode van twintig jaar. Op elk van zes meetmomenten legden zij hen negen morele dilemma's voor. Uit de interviews bleek dat ongeveer een kwart van de tienjarigen zich in de allereerste fase van morele ontwikkeling bevond. Bij tien- tot veertienjarigen was fase 2 de meest voorkomende manier van moreel oordelen. Als volwassenen kwamen de meesten niet verder dan het tweede niveau (fase 3 en 4) en het derde niveau (fase 5 en 6) was maar voor zeer weinigen weggelegd. Op twaalfjarige leeftijd bevond zich ongeveer 10% van de jongens in fase 1, 50% in fase 2, 28% in fase 3 en 2% in fase 4.

Inlevingsvermogen

Martin Hoffman heeft de aandacht gevestigd op de belangrijke rol van emoties voor de morele ontwikkeling. De emotionele dimensie van morele ontwikkeling betreft het kunnen ervaren van emoties zoals medelijden, sympathie, compassie, verontwaardiging, schuld, spijt en schaamte. Moreel gedrag veronderstelt dat mensen zich kunnen inleven in de gedachten, emoties en bedoelingen van anderen en daarmee rekening kunnen houden. Dit wordt ook wel inlevingsvermogen, empathisch vermogen of 'perspectief nemen' genoemd. Leerlingen maken zich normen en waarden eigen doordat zij leren zich emotioneel in anderen te verplaatsen. Leerlingen zijn normaliter verontwaardigd en verdrietig als zij of een ander het slachtoffer worden van bijvoorbeeld relationele, materiële of fysieke agressie. Als zij anderen iets aandoen, voelen zij zich doorgaans schuldig en schamen zich voor hun daden. In de emotionele benadering is een gebrekkige morele ontwikkeling vooral te wijten aan het onvermogen tot het ervaren van emoties die passen bij de aard van de (dreigende) morele overtreding.

Afwijkende opvattingen

Van recente datum is de aandacht voor de rol die afwijkende opvattingen spelen in het ontstaan van normoverschrijdend gedrag. Een achterblijvende morele of emotionele ontwikkeling hoeft niet per se tot ernstige gedragsproblemen te leiden. Daarvoor is meer nodig. John Gibbs (2003) toonde aan dat daarvoor ook zelfbeschermende vertekeningen nodig zijn, ook wel zelfbeschermende denkfouten genoemd. Verkeerde interpretaties van sociaal gedrag kunnen verstrekkende gevolgen hebben. Op een vol speelplein loopt een kind bijvoorbeeld per ongeluk tegen een ander kind aan. Het andere kind vermoedt ten onrechte boze opzet en reageert onmiddellijk met een flinke dreun. Onderzoek heeft laten zien dat onjuiste interpretaties van sociaal gedrag een belangrijke rol spelen bij het ontstaan en in stand houden van probleemgedrag. Voorbeelden van onjuiste interpretaties zijn 'Als ik mensen uitlach, is het hun eigen schuld', 'Sommige kinderen vragen om een pak slaag' en 'Schelden is niet erg, het doet toch niemand pijn'. Deze verkeerde interpretaties worden zelfbeschermende denkfouten genoemd omdat ze het eigen belang dienen. Ze maken het probleemgedrag voor de dader acceptabel en zorgen ervoor dat het positieve zelfbeeld geen schade ondervindt. Leerlingen die agressief gedrag vertonen, weten vaak heel goed dat hun gedrag moreel onjuist is. Denkfouten bieden dan een rechtvaardiging voor het gedrag. Zelfbeschermende denkfouten kunnen worden ingedeeld in twee hoofdtypen: primaire en secundaire denkfouten.

Primaire denkfouten zijn egocentrische en egoïstische overtuigingen en houdingen die de drijfveer vormen van antisociaal gedrag. Een leerling met een egoïstisch denkperspectief kan zich niet inleven in de ander en houdt geen rekening met andermans legitieme gevoelens, wensen, behoeften en rechten.

Secundaire denkfouten zijn rationalisaties van antisociaal gedrag die het eigen belang dienen. Ze dragen ertoe bij dat het positieve zelfbeeld intact blijft als dat aangetast dreigt te worden door gevoelens van schaamte of schuld. Een leerling die de schuld altijd bij anderen zoekt of de ernst van zijn of haar antisociaal gedrag ontkent of bagatelliseert, zal zich minder schamen en zich minder schuldig en verantwoordelijk voelen.

Verondersteld wordt dat iedereen vatbaar is voor zelfbeschermende denkfouten. Leerlingen die veel van deze denkfouten maken, neigen echter meer tot probleemgedrag en lopen een groter risico op latere delinquentie.

Preventie- en interventieprogramma's

Voor kinderen en jongeren met ernstig probleemgedrag zijn programma's ontwikkeld die zich richten op moreel redeneren, inlevingsvermogen en zelfbeschermende denkfouten. Bekende interventieprogramma's zijn EQUIP en Agression Replacement Training. De achterliggende gedachte is dat de leerling zich door de training bewust wordt van het problematische karakter van het gedrag, daardoor een voorkeur voor prosociaal gedrag ontwikkelt en uiteindelijk minder probleemgedrag zal vertonen. Kennis van moreel redeneren, inlevingsvermogen en zelfbeschermende denkfouten is van belang voor docenten die in hun klas regelmatig te maken hebben met agressief gedrag of dat willen voorkomen. Door onvolwassen morele redeneringen en onjuiste interpretaties van sociaal gedrag te bespreken, kan de docent een bijdrage leveren aan het verminderen van probleemgedrag. Sommige methoden voor sociaal-emotionele ontwikkeling besteden hier overigens al ruimschoots aandacht aan.

Kanttekening

Bij het onderwijzen en evalueren van houdingen en waarden in onderwijssituaties past een kanttekening. Wetenschappers wijzen erop dat houdingen een vaardighedencomponent en een waardencomponent bezitten. Het bevorderen van een coöperatieve houding kan bijvoorbeeld in twee richtingen worden uitgewerkt: leren hoe men samenwerkt (de vaardigheid) en leren dat men moet samenwerken (de waarde dat samenwerken goed is). Het aanleren van samenwerkingsvaardigheden zien velen als een legitieme onderwijsdoelstelling, terwijl het leren dat er moet worden samengewerkt als indoctrinatie wordt opgevat. Sommigen vinden

dan ook dat waarden niet als doelstellingen onderwezen mogen worden, laat staan ze bij leerlingen te toetsen. Een uitzondering wordt wel gemaakt voor waarden die universeel van aard zijn (in plaats van persoonlijkheidsafhankelijk). Voorbeelden zijn de eerder besproken basiswaarden van de democratische rechtsstaat uit het Inspectietoezicht.

5.3 Probleemgedrag als onderwijsdoelstelling?

In de dagelijkse lespraktijk hebben docenten in meer of mindere mate te maken met probleemgedrag van leerlingen. Dat kan het lesgeven bemoeilijken en in sommige gevallen zelfs onmogelijk maken. Vaak worden deze gedragsproblemen toegeschreven aan een gebrekkige sociale competentie. Leerlingen met sociaal-emotionele problemen zijn vaak druk, agressief en oppositioneel of juist teruggetrokken, angstig en depressief. Velen hinderen niet alleen anderen, maar ondervinden ook zelf hinder van hun gedrag in de zin dat zij door andere leerlingen worden gepest, gemeden of afgewezen. Docenten besteden een niet onaanzienlijk deel van hun tijd en energie aan de begeleiding van 'lastige' of anderszins problematische leerlingen. In de WVO, de kerndoelen en de examenprogramma's komen we onderwijsdoelen met betrekking tot probleemgedrag begrijpelijkerwijs niet tegen. Wel kun je zeggen dat de afwezigheid van probleemgedrag een belangrijk kenmerk is van een gezonde sociaal-emotionele ontwikkeling en een schoolklimaat dat gunstig is voor leren. Het voorkomen en verminderen van probleemgedrag wordt tegenwoordig dan ook als een belangrijke doelstelling van het onderwijs beschouwd. Veel scholen organiseren daartoe sociale vaardigheidstrainingen.

Tip

Een voorbeeld van een sociale vaardigheidstraining is 'Mag ik meedoen?' De training bestaat uit twaalf lessen, waarin de leerlingen effectief gedrag oefenen en succes ervaren. Dat geeft hen het zelfvertrouwen om makkelijker met hun leeftijdsgenoten om te gaan. De training wordt op meer dan 150 scholen gegeven door getrainde docenten van de eigen school.

Met betrekking tot de preventie en behandeling van probleemgedrag kan een onderscheid worden gemaakt in drie soorten sociaal-emotionele problemen:

- externaliserende gedragsproblematiek;
- internaliserende gedragsproblematiek;
- specifieke gedragsproblematiek zoals ADHD en autisme.

Externaliserende gedragsproblematiek

Onder externaliserende problemen wordt negatief, naar buiten gericht gedrag verstaan, zoals ongehoorzaam, dwars en tegendraads gedrag, fysieke, materiële en verbale agressie en delinquentie. De twee belangrijkste stoornissen zijn ODD (Oppositional Defiant Disorder ofwel oppositioneel-opstandige gedragsstoornis) en CD (Conduct Disorder ofwel antisociale gedragsstoornis). ODD kan worden gekarakteriseerd als een terugkerend patroon van negativistisch, opstandig, ongehoorzaam en vijandig gedrag tegenover autoriteitsfiguren dat disfunctioneren tot gevolg heeft. CD is een zich herhalend en duurzaam gedragspatroon waarbij de grondrechten van anderen en belangrijke, bij de leeftijd horende, sociale normen en regels worden overtreden dat zich uit in ernstig agressief en antisociaal gedrag.

Met betrekking tot sociale gedragsproblematiek kan een onderscheid worden gemaakt in oppositioneel-opstandig, agressief en antisociaal gedrag. Oppositioneel-opstandige leerlingen verzetten zich op een hinderlijke en ergerlijke manier tegen anderen, meestal volwassenen.

Ze zijn prikkelbaar, opvliegend, snel beledigd en zoeken de schuld bij voorkeur bij anderen. Agressief gedrag is erop gericht anderen psychische of lichamelijke schade te berokkenen. Het kan gaan om bedreigen, pesten, vechten, slaan en schoppen, maar ook om het opzettelijk vernielen van andermans eigendommen. De agressie kan direct dan wel indirect zijn en openlijk dan wel in bedekte termen voorkomen. Antisociaal gedrag omvat zogeheten predelinquent gedrag waarbij regels en afspraken worden geschonden, zoals liegen en stelen. Op school kan het tot uiting komen in het negeren van de regels en spijbelen.

Internaliserende gedragsproblematiek

Internaliserende problemen zijn meer naar binnen gericht, zoals teruggetrokken gedrag, angst en depressie. Hoewel vaak niet storend voor anderen, kunnen internaliserende problemen het functioneren van betrokkenen ernstig belemmeren. Een onderscheid kan worden gemaakt in algemene, onberedeneerde angst, angstig-teruggetrokken gedrag (o.a. verlegen, teruggetrokken, vermijdingsgedrag) en angstig-depressief gedrag (o.a. verdrietig, neerslachtig, somber, gevoel van waardeloosheid).

Specifieke gedragsproblematiek

Specifieke gedragsproblematiek omvat onder meer ADHD (Attention Deficit Hyperactivity Disorder ofwel aandachtstekort in combinatie met hyperactiviteit) en autisme. Met betrekking tot ADHD kan een onderscheid worden gemaakt in aandachtstekort, hyperactiviteit en impulsiviteit. Leerlingen met aandachtstekort hebben grote moeite hun aandacht bij taken, spel of sociale situaties te houden. Ze kunnen zich slecht concentreren, zijn snel afgeleid, hun aanpak is weinig doordacht en hun werk is slordig. Hyperactiviteit uit zich onder meer in hoge mate van motorische onrust (o.a. onrustig, lawaaierig, bewegingsoverdaad). Impulsieve leerlingen reageren direct op omgevingsprikkels (o.a. ongeduldig, prematuur handelen, moeite met instructies opvolgen, zichzelf in gevaar brengen). Autistische gedragsstoornissen laten we vanwege de lage prevalentie in vmbo, havo en vwo buiten beschouwing.

Verschillende gradaties van gedragsproblemen

Gedragsproblemen kunnen in verschillende gradaties voorkomen. De ernst van het probleem hangt af van hoe vaak het gedrag voorkomt, hoe lang het duurt, hoe intensief het is en in hoeveel situaties het optreedt. In zeer milde vorm gaat het om een betrekkelijk normaal verschijnsel dat deel uitmaakt van het normale functioneren. Komen de verschijnselen in milde mate voor, dan spreken we van problemen (het subklinische gebied). Bij ernstige, langer durende problemen is er sprake van een stoornis (het klinische gebied). Anders dan problemen zijn stoornissen zonder deskundige hulp niet oplosbaar. De docent kan hier beter pas op de plaats maken en de behandeling aan deskundigen overlaten. Het aantal leerlingen in het voorgezet onderwijs dat professionele hulp nodig heeft, is overigens niet gering. Volgens onderzoekers hebben meer dan tweehonderdduizend kinderen in het basisonderwijs gedragsproblemen die meer dan een beetje aandacht vereisen. Tabel 5.1 vermeldt de prevalentie van verschillende soorten gedragsproblemen bij twaalf- tot achttienjarige jongens en meisjes (volgens opgave van hun docenten).

Tabel 5.1: Prevalentie van gedragsproblemen bij jongeren in de leeftijd van twaalf tot achttien jaar (bron: Scholte & Van der Ploeg, 2004)

	Jongens	Meisjes
ADHD	15%	3%
Oppositieel-opstandig gedrag (ODD)	10%	7%
Agressief en antisociaal gedrag (CD)	12%	4%
Angstig en stemmingsverstoord gedrag	17%	15%

5.4 Methoden voor het meten van non-cognitieve doelen

Voor de meting van non-cognitieve onderwijsdoelen komen verschillende onderzoeksmethoden in aanmerking. Een eerste onderscheid is dat tussen instrumenten voor het meten van typisch en maximaal gedrag. Typisch gedrag is het gedrag dat een persoon normaliter in een situatie vertoont. Het wordt meestal op indirecte wijze gemeten via zelfrapportage of rapportage door anderen, bijvoorbeeld met een vragenlijst of een gedragsbeoordelingslijst. Maximaal gedrag is het gedrag dat een persoon vertoont als deze zijn of haar uiterste best doet om een zo goed mogelijke prestatie neer te zetten. In het cognitieve domein is het centraal schriftelijk examen wellicht het beste voorbeeld van het meten van maximaal gedrag. In het sociale domein wordt maximaal gedrag meestal op directe wijze gemeten door de leerlingen onder gecontroleerde omstandigheden een taak te laten uitvoeren, bijvoorbeeld een vignet of een rollenspel. Een tweede onderscheid is dat naar wie de gegevens verstrekt: de leerling, de docent, de ouders of verzorgers, of een externe observator, interviewer of proefleider. Hierna bespreken we enkele onderzoeksmethoden en hun voor- en nadelen.

Zelfrapportage met leerlingenvragenlijst (LV)

Voor de persoonlijke aspecten van non-cognitieve onderwijsdoelen is een vragenlijst vaak de beste keuze. De leerling is immers zelf het beste op de hoogte van zijn of haar gevoelens, motieven, voorkeuren, interesses en waarden. Het grote voordeel van een leerlingenvragenlijst is dat de afname weinig tijd kost. Als de docent een beeldschermtoets met meerkeuzevragen gebruikt, kan de computer de antwoorden van de leerlingen automatisch scoren. Wel mogen de vragen niet al te ingewikkeld zijn. Zelfrapportage of -beoordeling is niet geschikt voor zeer jonge kinderen, per definitie subjectief en gevoelig voor sociale wenselijkheid.

Docent gedragsbeoordelingslijst (DG)

Beoordeling door de docent is waarschijnlijk de meest geschikte methode als het gaat om sociale vaardigheden (bijvoorbeeld hoe de leerling met zijn medeleerlingen of de docent omgaat). De meting vindt meestal plaats met een gedragsbeoordelingslijst. Soms wordt de gedragsbeoordelingslijst gecombineerd met systematische observatie. De docent moet de leerling dan eerst gedurende een bepaalde periode observeren alvorens hem of haar te beoordelen. Gedragsbeoordeling door de docent heeft vele voordelen. De docent is als geen ander op de hoogte van het gedrag van de leerling in de klas. Het oordeel is gebaseerd op vele observaties over een lange periode. Een objectieve meting is mogelijk, al moeten de gedragscategorieën dan wel in observeerbare termen geformuleerd zijn. Het is ook mogelijk om de frequentie van weinig voorkomende maar toch belangrijke gedragingen vast te stellen. De methode levert ook informatie over leerlingen die te weinig taalvaardig zijn om zelf over hun

ontwikkeling te kunnen rapporteren of die dat simpelweg niet willen. Een punt van zorg is echter de gevoeligheid voor storende beoordelaarsfouten, zie hoofdstuk 7 van het algemene deel van Toetsen op School. Een aantal beoordelaarsfouten kan overigens worden bestreden door het gedrag van de leerling te formuleren in termen van enkelvoudige en direct waarneembare gedragingen. Een ander punt van zorg is dat kwaliteit van de meting afneemt naarmate de docent meer leerlingen in een korter tijdbestek moet beoordelen.

Vignet (VIG)

Een vignet kan worden gebruikt voor het meten van sociale competentie en morele ontwikkeling. Het bestaat uit een beschrijving van een hypothetische situatie of probleem (vaak scenario genoemd). Na zich in het scenario ingeleefd te hebben, beantwoordt de leerling een aantal vragen, bijvoorbeeld wat er in deze situatie gedaan zou moeten worden of wat hij of zij daadwerkelijk gedaan zou hebben. De antwoorden van de leerling worden vervolgens gecodeerd, bijvoorbeeld naar de effectiviteit of het morele gehalte van het sociale gedrag. Omdat de vignetmethode nog vrij onbekend is, geven we hieronder een nadere toelichting aan de hand van een voorbeeld. Het voorbeeld is afkomstig uit het Interview Sociale Informatie Verwerking van de hoogleraar ontwikkelingspsychologie Bram Orobio de Castro. Het instrument is gebaseerd op de theorie van sociale informatieverwerking (SIV). Deze theorie probeert te verklaren wat er misgaat bij leerlingen die problemen hebben met de interactie met volwassenen en leeftijdsgenoten. De theorie beschrijft verschillende stappen van informatieverwerking wanneer een leerling reageert op een als onaangenaam ervaren actie van een ander en niet duidelijk is of de ander die actie ook zo bedoelde. Leerlingen met antisociale gedragsproblemen blijken meer te letten op die aspecten van de situatie die een rechtvaardiging voor hun gedrag vormen, zoals een fout die de ander maakt. Zij interpreteren andermans gedrag vaker als een provocatie of aanval, hebben veel meer agressieve gedragingen in hun gedragsrepertoire, kiezen vaker voor antisociaal gedrag en zijn vaardiger in het uitvoeren van dat gedrag. En als hen - al dan niet vermeend - iets wordt aangedaan, worden zij veel bozer dan andere leerlingen en die boosheid sterkt hen in hun agressieve gedrag.

Met het Interview Sociale Informatieverwerking kan worden vastgesteld hoe leerlingen sociale informatie verwerken. Het bevat vijf vignetten, dit wil zeggen verhaaltjes die alledaagse sociale situaties beschrijven waarin een leerling door een leeftijdgenoot gestoord of benadeeld wordt. Daarbij is niet duidelijk wat de emoties en intenties van deze leeftijdgenoot zijn. Een van deze vignetten is hieronder integraal weergegeven.

De limonade

Stel je voor: het is feest op school. Er is een markt op het schoolplein. Iedereen krijgt een bon voor gratis drinken. Het is natuurlijk hartstikke druk: iedereen loopt door elkaar heen. Het is warm, en je krijgt vreselijke dorst. Dus ga je met je bon iets te drinken halen. Je moet lang in de rij staan, maar eindelijk ben je aan de beurt. Je krijgt een grote beker met limonade. Maar als je net je eerste slok wilt nemen, stoot van achter een jongen/meisje tegen je aan. Alle limonade klotst uit de beker in je gezicht!

Nadat de interviewer het verhaaltje heeft voorgelezen, beantwoordt de leerling vragen over de emoties van het vermeende slachtoffer en de emoties en bedoelingen van de vermeende dader. Vervolgens wordt de leerling gevraagd hoe hij of zij zelf gereageerd zou hebben. Tot slot legt de interviewer de leerling een voorbeeld van een agressieve en een prosociale reactie voor, met het verzoek aan te geven in hoeverre zij dat goede reacties vinden.

Een van de interviewvragen gaat over de manier waarop de leerling de bedoeling van de andere leerling interpreteert. Het antwoordmodel aan de hand waarvan de antwoorden gecodeerd worden, is hierna onverkort weergegeven.

Het interpreteren van de bedoelingen van de ander

0 = Weet niet of irrelevant antwoord.

- De leerling geeft ook na doorvragen aan de bedoeling van de ander niet te kunnen inschatten, zonder een afweging te maken als bij antwoordmogelijkheid 3.

1 = Welwillend.

- Het gedrag van de ander is bedoeld om de leerling te helpen, of om voor beiden positief te zijn.

2 = Per ongeluk of stom.

- De ander kon er niets aan doen, er was sprake van overmacht, domheid of onhandigheid.

3 = Ambigu.

- De leerling overweegt zowel de mogelijkheid van opzet als van onopzettelijkheid, maar er wordt ook na doorvragen geen keuze gemaakt.

4 = Opzet of vijandig.

- De ander wil de leerling kwetsen of hinderen, uit woede of jaloezie, of om ten koste van de leerling een eigen doel te bereiken. Hieronder valt ook de keuze voor opzet na het overwegen van mogelijke goede intenties.

Een andere interviewvraag is hoe de leerling zelf op de actie van de vermeende dader gereageerd zou hebben. De antwoorden op deze vraag worden gecodeerd volgens onderstaand antwoordmodel.

Reactie op de vermeende dader

0 = Weet niet of irrelevant

1 = Geen agressie, met als voorbeelden:

- Niets doen. De leerling is passief, negeert of vermijdt het probleem.
- Oplossen. De leerling suggereert het (begin van) een adequate oplossing voor het probleem.
- Vragen/overleggen. De leerling informeert op een niet beschuldigende manier naar de intentie of emotie van de ander, om een betere inschatting van de situatie te kunnen maken.
- Verbale agressie, zoals schelden, vloeken, roddelen, kleineren of treiteren van de ander. Onder deze categorie vallen alleen direct tegen de ander gerichte uitspraken, dus wel tegen de ander gerichte vloeken of scheldwoorden, maar geen vloeken of schelden in het algemeen of tegen kapotte voorwerpen.

2 = Milde agressie, met als voorbeelden:

- Directief: commanderen, dreigen of de ander voor een voldongen feit plaatsen.
- Verbale agressie, zoals schelden, vloeken, roddelen, kleineren of treiteren van de ander. Onder deze categorie vallen alleen direct tegen de ander gerichte uitspraken, dus wel tegen de ander gerichte vloeken of scheldwoorden, maar geen vloeken of schelden in het algemeen of tegen kapotte voorwerpen.

3 = Ernstige agressie, met als voorbeelden:

- Fysieke agressie zoals slaan, duwen, (haren) trekken, wegtrekken of knijpen.
- Materiële agressie zoals het kapot maken, vies maken of afpakken van spullen van de ander (iets kapot maken, limonade op jas gooien).

Door de scores voor de antwoorden over de vijf vignetten bij elkaar op te tellen, worden totaal-scores verkregen op basis waarvan onder meer uitspraken gedaan kunnen worden over:

- het kunnen herkennen en benoemen van eigen emoties;
- het kunnen omgaan met eigen gevoelens van boosheid en verdriet (emotieregulering);
- het kunnen interpreteren van emoties van anderen;
- het kunnen interpreteren van de bedoelingen van anderen (intentietoekenning);
- het kiezen van een reactie op het gedrag van anderen;
- het waarderen en evalueren van reacties op het gedrag van anderen.

Rollenspel (RS)

In een schriftelijk vignet zijn de mogelijkheden tot interactie met de leerling beperkt. Een rollenspel onder leiding van de docent biedt meer mogelijkheden om een levensechte sociale situatie te benaderen (waarbij naast de volwassene ook een leerling als ‘medeproefleider’ kan fungeren). De leerling moet het sociale gedrag in een rollenspel immers echt laten zien. De docent moet dan wel eerst leren hoe hij of zij het rollenspel moet leiden. De voorbereiding en afname van een rollenspel is tijdrovend. Verder is het nog maar de vraag of leerlingen het gedrag dat zij tijdens het rollenspel vertonen ook in de ‘echte’ situatie laten zien. Een punt van zorg is dat deelnemers zich tijdens rollenspelen doorgaans niet bijzonder op hun gemak voelen en van mening zijn dat hun gedrag tijdens het rollenspel niet representatief is voor hun werkelijke sociale gedrag.

Leerlingen beoordelen elkaar

Bij sociometrische technieken beoordelen leerlingen elkaar door aan te geven met wie zij al dan niet graag spelen of samenwerken. Technieken waarbij leerlingen elkaar beoordelen, leveren vaak weinig specifieke informatie op over de sociale competenties van de leerlingen. De oordelen van leeftijdsgenoten zijn meestal weinig stabiel en de overeenstemming tussen de oordelen van verschillende leerlingen over dezelfde leerling is vaak laag. Daarnaast worden de oordelen sterk gekleurd door de sekse, leeftijd, etniciteit, uiterlijke kenmerken en minderheidsstatus van de leerling binnen de klas. Verder wordt het niet ethisch gevonden als leerlingen bijvoorbeeld moeten aangeven wie zij niet leuk of niet aardig vinden. Tot slot kan het moeilijk zijn om van de ouders toestemming voor het onderzoek te verkrijgen.

Objectieve gegevens

In sommige gevallen kunnen non-cognitieve doelstellingen ook door middel van objectieve gegevens worden onderzocht. Zo kan het welbevinden van leerlingen behalve via vragenlijsten ook gemeten worden aan de hand van ‘objectieve’ indicatoren zoals het aantal keren dat een leerling spijbelt of betrokken is bij geregistreerde incidenten.

Welke methoden voor welke doelstellingen?

Niet alle methoden zijn voor alle non-cognitieve onderwijsdoelen even geschikt. Als voorbeeld hebben wij in tabel 5.2 aangegeven naar welke onderzoeksmethoden onze persoonlijke voorkeur uitgaat. Voor enkele non-cognitieve onderwijsdoelen is aangegeven wat onze eerste en tweede keus zou zijn als wij de leerlingen zelf zouden willen testen.

Tabel 5.2: Methoden voor het meten van non-cognitieve onderwijsdoelen

	Eerste keuze	Tweede keuze
Zelfbeeld en zelfvertrouwen	LV	DG
Werkhouding/nauwkeurigheid	DG en LV	DV
Interesse en motivatie voor school en leren	LV	DG
Schoolklimaat (o.a. relaties met medeleerlingen en docent)	LV en DV	?
Morele houdingen, waarden en gedragingen (o.a. de houding t.o.v. de basiswaarden van de democratische rechtsstaat)	VIG en LV	DV
Probleemgedrag (internaliserend, externaliserend en specifieke problemen zoals ADHD)	DG en LV	DV

LV = Leerlingenvragenlijst, DV = Docentvragenlijst, DG = Docentgedragsbeoordelingslijst, VIG = Vignet, RS = Rollenspel

5.5 Instrumenten voor het meten van non-cognitieve doelen

Het zelf ontwikkelen van non-cognitieve instrumenten is veel werk en kost veel tijd. Gelukkig zijn er bruikbare instrumenten in de handel verkrijgbaar die de school zelf kan afnemen. Meestal gaat het om vragenlijsten en gedragsbeoordelingslijsten voor de leerlingen, de docent en/of de ouders. Het materiaal bevat altijd een handleiding met aanwijzingen voor de afname, de verwerking van de gegevens en de interpretatie van de uitslag. Vaak zijn de instrumenten landelijk genormeerd, dit wil zeggen dat de docent de resultaten van zijn of haar leerlingen kan vergelijken met die van een representatieve steekproef van leerlingen van dezelfde leeftijd, opleidingstype, leerjaar en/of geslacht. De instrumenten voor het vaststellen van probleemgedrag bieden doorgaans aanwijzingen voor de interpretatie van de ernst van het gedrag en suggesties voor een eventueel vervoltraject (bijvoorbeeld het inschakelen van professionele hulp van buiten de school). Vele instrumenten bieden ook een rapportageservice maar daar zijn vaak behoorlijke kosten mee gemoeid. De school moet dan een bedrag per leerling aan de uitgeverij betalen, waardoor de totale kosten relatief hoog uitvallen. Een ander nadeel is dat het aantal vragen of aandachtspunten vaak erg groot is en de afname dus veel tijd kost.

Een in de handel verkrijgbaar instrument is de beste keuze als een betrouwbare en valide meting echt noodzakelijk is en de uitslag van de test voor de leerling, de docent of de school belangrijke consequenties heeft. Dat is onder meer het geval bij het diagnosticeren van sociaal-emotionele stoornissen, het plaatsen van een leerling in een andere onderwijssoort of het verwijzen van leerlingen naar het speciaal voortgezet onderwijs. Wil de docent 'slechts' een globaal beeld krijgen van het welbevinden van de leerlingen in zijn of haar klas, dan voldoet een korte zelfgemaakte vragenlijst (zie voor tips en vuistregels, paragraaf 5.6).

Docenten kunnen in de handel verkrijgbare instrumenten (of delen daarvan) voor drie doelen gebruiken:

- Het opsporen en zo nodig begeleiden van individuele leerlingen met sociaal-emotionele problemen.
- Het evalueren en zo nodig verbeteren van het onderwijsleerproces of het klimaat in de klas of school.
- Het inwinnen van een advies ten behoeve van beslissingen met betrekking tot de schoolloopbaan van de leerling (bijvoorbeeld: plaatsing in een andere onderwijssoort of verwijzing naar het voortgezet speciaal onderwijs).

Hieronder beschrijven we kort enkele instrumenten die bedoeld zijn voor gebruik in het voortgezet onderwijs. Als geheel geven deze beschrijvingen een goed overzicht van de onderwijsdoelen die in het voortgezet onderwijs nagestreefd en gemeten kunnen worden. Gebruikers moeten op de kwaliteit van instrumenten kunnen vertrouwen. Daarom vermelden we onder ieder instrument het kwaliteitsoordeel van de Commissie Testaangelegenheden Nederland (COTAN). De COTAN beoordeelt de kwaliteit van tests en toetsen op basis van zeven criteria:

- 1 Uitgangspunten van de testconstructie
- 2 Kwaliteit van het testmateriaal
- 3 Kwaliteit van de handleiding
- 4 Normen
- 5 Betrouwbaarheid
- 6 Begripsvaliditeit
- 7 Criteriumvaliditeit

In hoofdstuk 10 van het algemene deel van Toetsen op School is het COTAN beoordelings-systeem voor de kwaliteit van tests en toetsen besproken. De beoordelingsprocedure is te vinden op www.psynip.nl > Wat doet het NIP.

Sociaal-emotionele vragenlijst (SEV) van Scholte en Van der Ploeg

De SEV meet een breed scala aan sociaal-emotionele problemen. Om de SEV te kunnen invullen moet de docent de leerling en zijn of haar gedrag goed kennen. Er zijn vier basisschalen waarvan er drie zijn opgesplitst in subschalen:

- Aandachtstekort met hyperactiviteit met de subschalen Aandachtstekort, Hyperactiviteit en Impulsiviteit.
- Sociale gedragsproblematiek met de subschalen Oppositioneel-opstandig gedrag, Agressief gedrag en Antisociaal gedrag.
- Angstig en stemmingsverstoord met de subschalen Angstig gedrag in het algemeen, Sociaal-angstig gedrag en Angstig-depressief gedrag.
- Autistisch gedrag.

De COTAN beoordeelde de kwaliteit van de SEV met twee voldoende's en gaf vier keer de kwalificatie goed.

ADHD Vragenlijst (AVL) van Scholte en Van der Ploeg

De AVL meet of, en in welke mate, er bij leerlingen sprake is van gedragsproblemen van ADHD. ADHD verwijst naar drie basisvormen van probleemgedrag, te weten: aandachtstekort (snel afgeleid en weinig taakgericht), hyperactiviteit (overbeweeglijkheid en motorische onrust) en impulsiviteit (ongecontroleerd en ongeremd gedrag) of een combinatie van deze problemen. De AVL kan worden ingevuld door beoordelaars die goed op de hoogte zijn van het gedrag van het kind, zoals ouders, docenten, groepsleiders, psychologen, orthopedagogen of andere professionele hulpverleners. Voor de interpretatie van de uitslag is psycho-diagnostische

scholing vereist (kwalificatieniveau 2). De AVL kan gebruikt worden als screening binnen het onderwijs om ADHD-gedrag vroegtijdig op te sporen.

De AVL kreeg van de COTAN twee voldoende's en vier keer de qualificatie goed.

Vragenlijst Sterke Kanten en Moeilijkheden (SDQ)

De Vragenlijst Sterke Kanten en Moeilijkheden is een vertaling van de Strength and Difficulties Questionnaire (SDQ). De SDQ is gebaseerd op de DSM-classificatie (Diagnostic and Statistical Manual of mental disorders) van psychosociale problemen bij jongeren. Het instrument beoogt tevens problemen met leeftijdsgenoten te meten als indicatie voor psychische problemen.

De vragenlijst is bedoeld om kinderen met een hoog risico op psychosociale problemen te signaleren. De SDQ meet de aanwezigheid van deze problemen en wat de sterke kanten van het kind genoemd worden. De vragenlijst bevat in totaal 25 uitspraken over de aspecten:

- Emotionele problemen.
- Gedragsproblemen.
- Hyperactiviteit en aandachtstekort.
- Problemen met leeftijdsgenoten.
- Pro-sociaal gedrag.

De COTAN heeft de kwaliteit van de handleiding (te weinig informatie) en de normen (niet representatief of representativiteit niet te beoordelen) als onvoldoende beoordeeld. Voor het overige was het kwaliteitsoordeel drie keer voldoende, één keer goed en één keer onvoldoende voor de criteriumvaliditeit vanwege te weinig onderzoek.

Hoe Ik Denk Vragenlijst (HID) van Brugman en collega's

De HID meet zogeheten zelfbeschermende denkfouten of vertekeningen die een rol spelen bij het ontstaan van externaliserend probleemgedrag. Deze verkeerde interpretaties worden zelfbeschermende denkfouten genoemd omdat ze het probleemgedrag voor de leerling acceptabel maken en ertoe bijdragen dat het zelfbeeld geen schade ondervindt. Denkfouten bieden een rechtvaardiging voor externaliserend probleemgedrag waardoor de leerling zich minder schaamt en zich minder schuldig voelt. Voorbeelden van onjuiste interpretaties zijn 'Als ik mensen uitlach, is het hun eigen schuld', 'Sommige mensen vragen om een pak slaag' en 'Schelden is niet erg, het doet toch niemand pijn'. De HID meet vier soorten denkfouten:

- Zelfcentrerend (egoïstisch en egocentrisch). Zodanig belang toekennen aan eigen behoeften, rechten, gevoelens, ideeën en verwachtingen, dat gelegitimeerde behoeften, rechten enzovoort van anderen en ook het eigenbelang op langere termijn, niet of nauwelijks in acht worden genomen.
- Anderen de schuld geven. Schuld afschuiven op externe oorzaken, vooral op een ander persoon, een groep of een tijdelijke toestand (bijvoorbeeld onder invloed van alcohol of drugs) of het onjuist toeschrijven van het eigen slachtofferschap of ongelukkig lot aan onschuldige personen.
- Goedpraten/verkeerd benoemen. Antisociaal gedrag afschilderen als gedrag dat niet erg is, anderen geen schade toebrengt of zelfs acceptabel of wenselijk is; en/of in kleinerende of dehumaniserende termen over andere personen praten ('Hij stelde toch niets voor').
- Uitgaan van het ergste. Het toekennen van vijandige bedoelingen aan andere personen ('Ze deed het met opzet'); het ergste verwachten van een sociale situatie, alsof het onvermijdelijk is; verbeteringen of veranderingen in het eigen gedrag of het gedrag van anderen als onmogelijk beschouwen.

DE HID is nog niet door de COTAN beoordeeld.

Competentiebelevingsschaal voor adolescenten (CBSA) van Treffers en collega's

De CBSA is een leerlingenvragenlijst die de beleving van de eigen competentie en het gevoel van eigenwaarde meet. Het instrument geeft een indicatie van de globale houding van de leerling ten aanzien van de school en zichzelf voor zover van belang voor het onderwijsleerproces.

De CBSA kan worden gebruikt als signaleringsinstrument om leerlingen met een relatief negatieve houding op te sporen en zo nodig te begeleiden. Het kan tevens worden gebruikt als adviseringsinstrument bij beslissingen over de school- en beroepsloopbaan. Tot slot kan de CBSA worden gebruikt als evaluatie-instrument om het onderwijs- en leerproces te evalueren en zo nodig aan te passen. De CBSA bestaat uit zeven schalen, elk bestaand uit vijf items (waarbij een schaal een verzameling items is die samen een bepaald affectief aspect meten):

- Schoolvaardigheden.
- Sociale acceptatie.
- Sportieve vaardigheden.
- Fysieke verschijning.
- Gedragshouding.
- Hechte vriendschap.
- Gevoel van eigenwaarde.

De CBSA kreeg van de COTAN één voldoende, vijf keer de kwalificatie goed en één onvoldoende voor de criteriumvaliditeit vanwege te weinig onderzoek.

Klimaatschaal van Donkers (2010) met versies voor klas- en schoolklimaat

De Klimaatschaal is een vragenlijst waarbij de leerling (leerlingversie) en de docent (docentversie) kunnen aangegeven welk klimaat er in de klas en op school heerst. Leerlingen kunnen ermee aangeven hoe zij het pedagogisch klimaat in de klas ervaren (sfeer en onderlinge relaties) en anderzijds kunnen zij feedback geven op het onderwijsgedrag (pedagogisch en onderwijskundig) van de leraar. Het instrument bestaat uit de volgende vier clusters van vragen:

- Kwaliteit van de onderlinge leerling-relaties.
- Sfeer in de klas.
- Orde in de klas.
- Interactie docent en de leerlingen.

Het instrument is in te zien op www.klimaatschaal.nl. Een leuk en leerzaam filmpje geeft praktische informatie over hoe men het instrument in de lessen kan inpassen.

De Klimaatschaal is niet door de COTAN beoordeeld.

Studeon van Cito

Studeon is een instrument voor het signaleren van de studiehouding en de sociaal-emotionele ontwikkeling van leerlingen in het eerste en tweede leerjaar. Het is een onderdeel van het Cito Volgstelsel voortgezet onderwijs (voorheen VAS) en bestaat uit vier subschalen:

- Persoonlijkheid. Deze schaal meet Zorgvuldigheid (hoe nauwkeurig en geordend ben je?), Vriendelijkheid (in hoeverre houd je rekening met andere mensen?), Emotionele stabiliteit (hoe goed kun je met problemen omgaan?) en Extraversie (hoe graag heb je andere mensen om je heen?).
- Welbevinden. Deze schaal meet Zelfvertrouwen (hoe zeker ben je van jezelf?), Relatie met docenten (hoe is het contact met je docenten?), Relatie met medeleerlingen (hoe goed kun je met je klasgenoten opschieten?) en Algemeen schoolbeeld (in hoeverre heb je het naar je zin op school?).
- Studiehouding volgens de leerling en Studiehouding volgens de mentor. Deze twee tests meten Cognitieve vaardigheden (hoe gemakkelijk kun je (nieuwe) informatie verwerken?), Metacognitieve vaardigheden (hoe goed kun je je manier van studeren aanpassen aan de

situatie?), Timemanagement (hoe goed kun je zelfstandig werken en plannen?) en Werkhouding (hoe actief zet je je in op school?).

Studeon is niet door de COTAN beoordeeld.

Schoolvragenlijst (SVL) van Vorst

De Schoolvragenlijst (SVL) meet verschillende aspecten van het sociaal-emotioneel functioneren en de betrokkenheid bij school. De versie voor het voortgezet onderwijs is bedoeld voor leerlingen uit klas 1, 2 en 3. De vragenlijst bestaat uit de volgende tien schalen:

- Leertaak Gerichtheid (LG): de mate waarin de leerling zich denkt in te zetten voor schoolwerk.
- Concentratie in de Klas (CK): de mate waarin de leerling zich meent te kunnen concentreren op schooltaken.
- Huiswerk Attitude (HA): de mate waarin de leerling zich denkt in te zetten voor het huiswerk.
- Plezier op School (PS): de mate waarin de leerling tevreden zegt te zijn met de school.
- Sociaal Aanvaard voelen (SA): de mate waarin de leerling zegt zich aanvaard te voelen door medeleerlingen.
- Relatie met Leerkrachten (RL): de mate waarin de leerling zegt een functionele relatie te onderhouden met de leerkrachten.
- UitdrukkingsVaardigheid (UV): de mate waarin de leerling zijn/haar gedachten onder woorden kan brengen.
- Zelfvertrouwen bij Proefwerken (ZP): de mate waarin de leerling zegt zelfvertrouwen te hebben in situaties op school waarin een prestatie geleverd moet worden.
- Sociale Vaardigheid (SV): de mate waarin de leerling moeilijke sociale situaties denkt aan te durven.
- Sociale Wenselijkheid (SW): de mate waarin de leerling geneigd is zichzelf in een gunstig daglicht te stellen bij zelfbeschrijving.

De SVL kreeg van de COTAN twee voldoende's, vier keer de kwalificatie goed en één onvoldoende voor de criteriumvaliditeit vanwege te weinig onderzoek.

Schoolattitude Questionnaire Internet (SAQI) van Vorst

SAQI is een beeldschermtest die is gebaseerd op de schriftelijke SchoolVragenLijst (SVL). Dit begeleidingsinstrument is gericht op drie globale houdingen van leerlingen ten opzichte van de school en zichzelf. De SAQI bestaat uit tien basisschalen met elk zestien items en vier samengestelde schalen:

- Motivatie (MOT): de attitudes jegens het schoolwerk, samengesteld uit de basisschalen Leertaakgerichtheid, Concentratie in de Klas en Huiswerkattitude.
- Satisfactie (SAT): de attitudes jegens het schoolleven, samengesteld uit de basisschalen Plezier op School, Sociaal Aanvaard voelen en Relatie met Leerkrachten.
- Zelfvertrouwen (ZELF): de attitudes jegens de eigen competenties, samengesteld uit de basisschalen UitdrukkingsVaardigheid, Zelfvertrouwen bij Proefwerken en Sociale Vaardigheid.
- Sociale Wenselijkheid is een aparte basisschaal die de mate meet waarin de leerling zichzelf in een gunstig daglicht wil stellen.
- Engagement voor school: deze schaal wordt gevormd door de som van de samengestelde schalen MOT, SAT en ZELF.

De COTAN beoordeelde de kwaliteit van de SAQI met twee keer voldoende, drie keer goed en één keer onvoldoende voor de criteriumvaliditeit vanwege te weinig onderzoek.

Belevingsschalen voor wiskunde en Engels van Cito

De Belevingsschaal voor wiskunde (BSW) meet de wijze waarop de klas tegenover het wiskunde-onderwijs staat, zodat de docent het onderwijs daarop kan aanpassen. De leerlingen beantwoorden de vragen anoniem. De BSW bevat vier subschalen:

- Plezier in wiskunde.
- Ervaren angst en moeilijkheid van wiskunde.
- Inzet en interesse voor wiskunde.
- Ervaren nut van wiskunde voor vervolgopleiding en beroep.

De Belevingsschaal voor Engels (BSE) meet de wijze waarop de klas tegenover het onderwijs in het Engels staat. De BSE bestaat uit vijf subschalen:

- Plezier in Engels.
- Integratieve motivatie: interesse voor land, cultuur en volk en taal.
- Instrumentele motivatie: nut voor vervolgopleiding en later beroep.
- Houding ten opzichte van de docent als vakdocent.
- Houding ten opzichte van de docent als persoon.

Tip

Docenten kunnen de BSW en BSE relatief eenvoudig bewerken tot belevingsschalen voor andere vakken. De vragenlijsten zijn opgenomen als bijlage [PDF 5.1](#) en [PDF 5.2](#) in de ToetsSpecial Toetsen op School, <http://www.toetswijzer.nl/berichten/toetsspecial-toetsen-op-school-voortgezet-onderwijs>

De BSW en BSE zijn niet door de COTAN beoordeeld.

Meetinstrument Burgerschap van Ten Dam en collega's

De vragenlijst Burgerschapscompetenties is bedoeld om effecten van onderwijs in burgerschap op leerlingen te meten. De vragenlijst meet het vermogen van jongeren om adequaat als burger te handelen in voor hen betekenisvolle situaties. De inhoud van de vragenlijst is vastgesteld op basis van een uitgebreide internationale literatuurstudie naar 'onderwijs en burgerschap'.

De vragenlijst is opgebouwd rond vier sociale taken die jongeren als burgers in onze samenleving moeten kunnen vervullen:

- Democratisch handelen.
- Maatschappelijk verantwoord handelen.
- Omgaan met conflicten.
- Omgaan met verschillen tussen mensen.

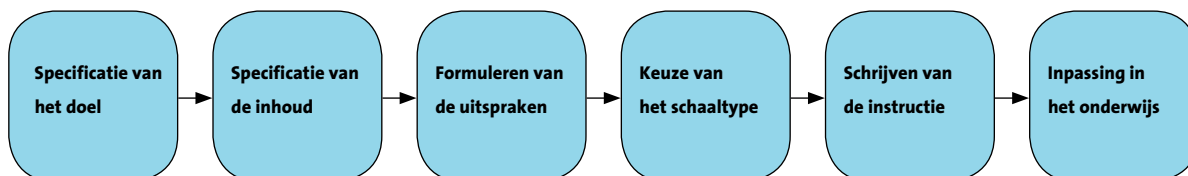
Elk van deze taken is in de vier componenten uiteengelegd: kennis, reflectie, attitude en vaardigheid.

De vragenlijst Burgerschapsgedrag meet het handelen van jongeren als burger. Deze vragenlijst wordt door de docent ingevuld.

Het Meetinstrument Burgerschap is nog niet door de COTAN beoordeeld.

5.6 Het zelf ontwikkelen van non-cognitieve instrumenten

De in de handel verkrijgbare instrumenten zijn vaak erg uitgebreid, waardoor de afname veel tijd kost. Wil de docent alleen een globaal beeld krijgen van het welbevinden van de leerlingen in de klas of de interesse voor het vak, dan voldoet een korte zelfgemaakte vragenlijst. Hieronder beschrijven we de ontwikkeling van een vragenlijst voor gebruik in de 'eigen' klas aan de hand van zes stappen. We beperken ons tot vragenlijsten omdat deze methode weinig afnametijd vergt maar toch veel informatie oplevert en er ter inspiratie goede voorbeelden beschikbaar zijn.



Stap 1 Specificatie van het doel

Ga eerst na waarom u een zelfgemaakte vragenlijst wilt afnemen. Wilt u meer te weten komen over het welbevinden van uw leerlingen of hoe zij tegen uw lessen of uw vak aankijken? Dan is een zelfgemaakte vragenlijst een goede optie. Of wilt u individuele leerlingen met leer-, motivatie- of gedragsproblemen opsporen om deze beter te kunnen begeleiden? In dit laatste geval kunt u beter voor een bestaand instrument kiezen dat daar speciaal voor gemaakt is. Dit laatste geldt ook als u een advies wilt inwinnen over de onderwijssoort die het beste bij de leerling past.

Stap 2 Specificatie van de inhoud

Ga voor uzelf na wat u precies met de vragenlijst wilt vaststellen. Een bekend onderscheid is dat tussen houdingen, interesses, gedragingen en waarden. Een hulpmiddel bij het concretiseren van non-cognitieve doelen zijn de SMART-criteria:

- S = specifiek
- M = meetbaar
- A = acceptabel
- R = realistisch
- T = tijdgebonden

Zie verder: <http://nl.wikipedia.org/wiki/SMART-principe> en <http://nl.wikipedia.org/wiki/RUMBA>.

Stap 3 Formuleren van de uitspraken

De volgende vraag is: hoe kom ik aan bruikbare items of uitspraken over het onderwerp waarover ik de leerlingen wil raadplegen? Het zelf formuleren van uitspraken is arbeidsintensief en tijdrovend. Het voordeel is echter dat de uitspraken optimaal zijn toegesneden op het doel van uw onderzoek en de kenmerken van uw onderwijs en leerlingen. Op het internet zijn met wat zoekwerk vaak bruikbare instrumenten en uitspraken te vinden.

Tip

Sommige leerlingen hebben de neiging om uit gemakzucht bij alle vragen hetzelfde antwoord aan te kruisen (een zogenaamde antwoordtendentie). U kunt dit tegengaan door zowel positief als negatief geformuleerde uitspraken in uw vragenlijst op te nemen. Wel moet u de antwoorden van de leerlingen op de negatief geformuleerde uitspraken bij het berekenen van de score van de leerling even hercoderen. Gebruikt u een vierpuntschaal met de schaalpunten 1, 2, 3 en 4, dan gaat dat als volgt: 1=4, 2=3, 3=2 en 4=1.

Stap 4 Keuze van het schaaltype

Vervolgens moet er een schaaltype worden gekozen. Niet ieder schaaltype is voor elk aspect van de persoonlijke, sociale en morele ontwikkeling van de leerling even geschikt. De twee voor docenten belangrijkste schaaltypen zijn:

- Likert-schaal
- frequentieschaal

Een Likert-schaal is geschikt voor het meten van meningen, interesses, houdingen en waarden. Met een frequentieschaal kunnen docenten het gedrag van leerlingen goed in kaart brengen. Voor een handleiding voor het ontwikkelen van beoordelingsschalen wordt verwezen naar hoofdstuk 8 van het algemene deel van Toetsen op School. Aanvullend op dat hoofdstuk lichten we de beide schaaltypen hierna kort toe.

Likert-schaal

De zogeheten Likert-schaal wordt veel gebruikt voor het meten van houdingen, interesses en waarden. Het is in wezen een samengestelde meerkeuzevraag waarin de leerling gevraagd wordt aan te geven in hoeverre deze het met bepaalde uitspraken of stellingen eens is. Meestal zijn er vier à zes antwoordmogelijkheden. De uiterste antwoordmogelijkheden zijn doorgaans 'helemaal mee eens' en 'helemaal mee oneens', maar andere varianten komen ook voor zoals 'zeer belangrijk' versus 'zeer onbelangrijk'. Onderstaande schaal is een fictief voorbeeld van een korte Likert-schaal voor het meten van de houding ten opzichte van lezen.

In hoeverre ben je het eens met de volgende uitspraken over lezen?
(s.v.p. per regel maar één hokje aankruisen)

	Helemaal mee oneens	Mee oneens	Mee eens	Sterk mee eens
Ik lees alleen als het moet	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Lezen is een van mijn favoriete bezigheden	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Ik vind het leuk om met andere mensen over boeken te praten	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Lezen is voor mij tijdverspilling	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Ik vind het leuk om naar een boekwinkel of een bibliotheek te gaan	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4

De totaalscore van de leerling kan men bepalen door de scores over de vijf uitspraken bij elkaar op te tellen. Een eenvoudige rekensom leert dat de somscore kan variëren van 5 (als de leerling het met alle uitspraken helemaal oneens is) tot 20 (als de leerling het met alle uitspraken sterk eens is). Een andere manier is het uitrekenen van het gemiddelde van de scores van de leerling. Het grote voordeel is dat men de uitslag kan interpreteren in termen van de oorspronkelijke schaal. Een gemiddelde van 3 wijst er dan bijvoorbeeld op dat de leerlingen het over het algemeen met de vijf uitspraken eens is.

Frequentieschaal

De frequentieschaal wordt veelvuldig gebruikt voor het meten van gedragingen, zoals pro sociaal gedrag, probleemgedrag en de gedragscomponent van houdingen. De respondent moet aangeven hoe vaak het gedrag in kwestie voorkomt. Enkele veel gebruikte varianten zijn:

- 1 = (bijna) nooit, 2 = soms, 3 = vaak, 4 = (bijna) altijd
- 1 = (bijna) nooit, 2 = enkele keren per maand, 3 = enkele keren per week, 4 = enkele keren per dag
- 1 = (bijna) nooit, 2 = maandelijks, 3 = wekelijks, 4 = dagelijks
- 1 = nooit, 2 = ongeveer één keer per maand, 3 = ongeveer één keer per week, 4 = (bijna) iedere dag
- 1 = nooit of bijna nooit, 2 = sommige lessen, 3 = de meeste lessen, 4 = elke les

Onderstaande schaal is een fictief voorbeeld van een korte gedragsbeoordelingslijst over ADHD (aandachtstekort, hyperactiviteit en impulsiviteit).

Hoe vaak vertoont de leerling in de klas het volgende gedrag?
(s.v.p. per regel maar één hokje aankruisen)

	(Bijna) nooit	Sommige lessen	Meeste lessen	Elke les
Is afgeleid	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Kan de aandacht niet vasthouden	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Werkt onzorgvuldig	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Gedraagt zich impulsief	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Zit te dagdromen	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Is druk en beweeglijk	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
Is vergeetachtig	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4

Stap 5 Schrijven van de instructie voor de leerlingen

Vergeet niet de vragenlijst anoniem af te laten nemen. Anders krijgt u sociaal wenselijke antwoorden. De leerlingen rapporteren dan niet wat zij werkelijk denken of voelen, maar wat zij denken dat u graag wilt horen. Enkele maatregelen om eerlijke antwoorden te verkrijgen zijn:

- Zeg in de instructie dat er geen goede of foute antwoorden zijn. Elk antwoord is goed zolang de leerling dat maar echt meent.
- Zeg dat de vragenlijst volstrekt anoniem is en dat u niet zult proberen te achterhalen wie wat heeft ingevuld.
- Zeg dat de leerlingen hun naam NIET op de vragenlijst mogen noteren.
- Neem geen open vragen op in de vragenlijst. Het enige wat van de leerlingen verlangd wordt, is het aankruisen of omcirkelen van het antwoord van hun keuze.
- Zeg dat de leerlingen NIET in de vragenlijst mogen schrijven. Dit omdat u anders zonder dat te willen hun identiteit aan de hand van het handschrift zou kunnen achterhalen.

Stap 6 Inpassing in het onderwijs

Docenten nemen een vragenlijst niet alleen af omdat het interessante gegevens oplevert. De leerlingen zullen de vragen serieuzer beantwoorden als de docent aangeeft dat hij of zij samen met hen iets wil verbeteren. Het aantal mogelijke onderwerpen is legio. Enkele voorbeelden zijn de manier van lesgeven, de manier waarop leerlingen in de klas met elkaar omgaan of de interesse en motivatie van de leerlingen voor het vak.

Als u de antwoorden van de leerlingen geanalyseerd heeft, presenteert u de resultaten aan de klas. U kunt dat het beste doen aan de hand van enkele eenvoudige tabellen en/of figuren. Een onderwijsleergesprek met de klas is een goed middel om van de uitkomsten te leren. U kunt het gesprek voeren aan de hand van de volgende vragen:

- Wat gaat goed?
- Wat gaat niet zo goed? En waarom niet?
- Wat zouden we kunnen verbeteren?
- Hoe zouden we dat kunnen doen?
- Hoe bepalen we straks of onze inspanningen succes hebben?

Tip

Bronnen voor verdere studie

- www.leraar24.nl
- www.leraar24.nl/dossier/825
- <http://sociaalemotioneel.slo.nl>
- www.cito.nl > Onderwijs > Voortgezet onderwijs > Alle producten > Studeon
- www.lantaarn.demon.nl > Medemens > Sociaal-emotionele ontwikkeling
- www.leefstijl.nl
- www.edux.nl > Doelgroepen > Voortgezet onderwijs > Sociaalemotionele ontwikkeling
- www.schoolenveiligheid.nl
- www.rijksoverheid.nl/ministeries/ocw > Onderwerpen > Veilig leren en werken in het onderwijs
- www.vo-raad.nl > Thema's > Alle thema's > Veiligheid
- www.venstersvoorverantwoording.nl
- www.pestweb.nl
- www.nji.nl
- www.onderwijsraad.nl > Dossiers > Vorming en burgerschap
- www.burgerschapmeten.nl
- www.onderwijsenburgerschap.nl
- www.leraar24.nl/dossier/612
- www.nji.nl > Kennis > Dossiers > Jeugdparticipatie
- www.kpcgroep.nl > KPC Groep > Diensten ABC > Burgerschap
- www.scholenpanels.nl

Geraadpleegde literatuur

Geraadpleegde literatuur

Alberts, R.V.H., & Erens, B.J.M. (2011). *Verslag van de examencampagne 2011 voortgezet onderwijs*. Arnhem: Cito.

Algra, A. (2004). Eerste en tweede correctie examens: problemen en regels. *Schoolmanagers_VO*, 6, 8-10.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th edition)*. Washington, DC: American Psychiatric Association.

Anderson, L.W., & Bourke, S.F. (2000). *Assessing affective characteristics in the schools (2nd edition)*. Mahwah, NJ: Lawrence Erlbaum.

Appelhof, P., & Walraven, M. (2002). *Sociale competentie ter bevordering van participatie in de samenleving. Een verkenning. Op zoek naar indicatoren*. Utrecht: Oberon.

Béguin, A. (2010). *Concept tijdpad en globale opzet voor de standaardbepaling van de referentieniveaus*. Arnhem: Cito.

Beuving, M., & Waterreus, I. (2010). Aanscherping exameneisen voor basisvakken van vmbo. *Examens*, 1, 17-20.

Branden, N. (1994). *Six pillars of self-esteem*. New York: Bantam.

Bronneman-Helmers, H., & Zeijl, E. (2008). Burgerschapsvorming in het onderwijs. In: P. Schnabel, R. Bijl, & J. de Hart (red.). *Betrekkelijke betrokkenheid. Studies in sociale cohesie. Sociaal en Cultureel Rapport 2008* (p. 170-205). Den Haag: SCP.

Brugman, D., Nas, C.N., Velden, F. van der, Barriga, A.Q., Gibbs, J.C., Granville, B.P., & Liau, A.K. (2011). *Hoe ik Denk Vragenlijst (HID). Handleiding*. Amsterdam: Boom test uitgevers.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development*, 48 (1-2).

Council of Europe (2001). *Common European Framework of References for Languages. Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Dam, G. ten, & Volman, M. (1999). *Scholen voor sociale competentie. Een pedagogisch-didactische benadering*. Lisse: Swets & Zeitlinger Publishers.

Dam, G. ten, & Volman, M. (2000). Sociale competentie: reddingsvest en levenskunst. Over de pedagogische opdracht in de praktijk. *Pedagogiek*, 20, 2, 112-127.

Dam, G. ten, & Volman, M. (2007). Educating for adulthood or for citizenship: social competence as an educational goal. *European Journal of Education*, 42, 281-298.

Dam, G. ten, Volman, M., Westerbeek, K., Wolfram, P., & Ledoux, G., m.m.v. Peschar, J. (2003). *Sociale competentie langs de meetlat*. Den Haag: Transferpunt Onderwijsachterstanden.

- Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2009). *COTAN beoordelingssysteem voor de kwaliteit van tests*. Amsterdam: NIP/COTAN.
- Expertgroep Doorlopende Leerlijnen (2008). *Over de drempels met rekenen. Consolideren, onderhouden, gebruiken en verdiepen. Onderdeel van de eindrapportage van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen*. Enschede.
- Expertgroep Doorlopende Leerlijnen (2008). *Over de drempels met taal. De niveaus voor de taalvaardigheid. Onderdeel van de eindrapportage van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen*. Enschede.
- Expertgroep Doorlopende Leerlijnen (2008). *Over de drempels met taal en rekenen. Hoofdrapport van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen*. Enschede.
- Expertgroep Doorlopende Leerlijnen (2009). *Een nadere beschouwing. Over de drempels met taal en rekenen*. Enschede.
- Gibbs, J.C. (2003). *Moral development and reality. Beyond the theories of Kohlberg and Hoffman*. Thousand Oaks, CA: Sage Publications.
- Gramberg, P., Löhner, S., & Waterreus, I. (2006). Loon naar lesgeven. *Economische Statistische Berichten*, 4497, 566-568.
- Hoffman, M.L. (2000). *Empathy and moral development. Implications for caring and justice*. New York: Cambridge University Press.
- Hoogenkamp, M., Joosten, F., & Voorst van Beest, K. van (2001). *Sociale competentie: een vak apart*. Leuven/Leusden: Acco.
- Inspectie van het Onderwijs (2006). Toezichtkader actief burgerschap en sociale integratie. In: *Staatscourant 5 juli 2006, nr. 128*, p. 13.
- Inspectie van het Onderwijs (2008). Herziening normering Toezichtkader actief burgerschap en sociale integratie. In: *Staatscourant 30 januari 2008, nr. 21*, p. 24.
- Inspectie van het Onderwijs (2009). *De praktijk van de beroepsgerichte examens voor het vmbo. Resultaten van een onderzoek naar de voorbereiding, afname, beoordeling, tweede correctie en herkansing van vier beroepsgerichte examens*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2009). *Het schoolexamen in het voortgezet onderwijs. Verslag van een onderzoek naar de kwaliteit van het schoolexamen bij de vakken Engels, Nederlands, biologie en wiskunde*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2009). *De staat van het onderwijs. Onderwijsverslag 2007/2008*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2010). *Analyse en waardering van opbrengsten primair onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2010). *De staat van het onderwijs. Onderwijsverslag 2009/2010*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs (2011). *Toezichtkader PO/VO 2009. Versie 13 juli 2009. Aanpassing 17 mei 2011*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs (2011). *Opbrengstenoverzicht 2011. Toelichting*. Utrecht: Inspectie van het Onderwijs.

Kerkhoffs, J., Stark E., & Zeelenberg, T. (2006). *Rubrics als beoordelingsinstrument voor vaardigheden*. Enschede: SLO.

Kieviet, Th., Tak, J.A., & Bosch, J.D. (red.) (2002). *Handboek psychodiagnostiek voor de hulpverlening aan kinderen*. Utrecht: De Tijdstroom uitgeverij.

Kleijne, W. (2008). Een gelijkwaardig diploma voor bijzondere kandidaten. Staatsexamens voor voortgezet onderwijs. *Examens*, 5, 5-10.

Kohlberg, L. (1984). *Essays on moral development. The psychology of moral development*. San Francisco: Harper & Row.

Kremers, E., Kuhlemeier, H., & Wiegers, J. (2006). Tweede Fase en prestatieniveau. Niveauvergelijking is geen sinecure. *Examens*, 1, 9-11.

Kuhlemeier, H., Kleintjes, F., & Goldebeld, P. (1998). Hoe verhoog ik de betrouwbaarheid van mijn toetsen? Effect van moeilijkheid en spreiding van itemmoeilijkheden op de betrouwbaarheid van de afsluitingstoetsen basisvorming. *Tijdschrift voor Onderwijsresearch*, 23, 1, 31-45.

Kuhlemeier, H., Steentjes, M., & Kleintjes, F. (2003). *De gelijkwaardigheid van open en meerkeuzevragen bij wiskunde. Effecten van vraagtype en scoringswijze op gemeten vaardigheden, betrouwbaarheid, moeilijkheid en afnametijd*. Arnhem: Cito.

Kuhlemeier, H. (2007). *Internationaal gangbare moderatieprocedures voor het vaststellen en verbeteren van de kwaliteit van schoolexamens*. Paper gepresenteerd op de 34e Onderwijs Research Dagen, 6-8 juni 2007, Groningen.

Kuhlemeier, H. (2011). Het construeren van praktijktoetsen. In: P. Sanders (red.). *Toetsen op School* (p. 125-141). Arnhem: Cito.

Kuhlemeier, H., Weeren, J. van, & Werf, M.P.C. van der (2006). *Scheiding van opleiden en examineren in Nederland en omliggende landen*. Studie verricht in opdracht van de Onderwijsraad. Arnhem: Cito.

Kuhlemeier, H., Gitsels, H., Boom, S., Kerkhof, A. van de, & Sinkeldam, R. (in voorbereiding). *Relaties tussen examenkenmerken en verschillen tussen correctoren in soepelheid bij het CSE geschiedenis, tehatex en Nederlands*. Arnhem: Cito.

Kuhlemeier, H., & Kremers, E. (in voorbereiding). *De praktijk van de eerste en tweede correctie van de centraal schriftelijke examens*. Arnhem: Cito.

Kuhlemeier, H., Rijn, P. van, & Kremers, E. (in voorbereiding). *Eerste correctie, tweede correctie en derde correctie van geannoteerde en blanco examenwerken in de centraal schriftelijke examens. Wat is het verschil?* Arnhem: Cito.

Ledoux, G., Geijssel, F., Dam, G. ten, & Reumerman, R. (2001). Burgerschapscompetenties van jongeren in Nederland. *Pedagogische Studiën*, 88, 1, 3-22.

Liemberg, E., & Meijer, D. (2004). *Taalprofielen*. Enschede: NaB-MVT.

Lieshout, T. van (2009). *Pedagogische adviezen voor speciale kinderen. Een praktisch handboek voor professionele opvoeders, begeleiders en leerkrachten*. Houten: Bohn Stafleu van Loghum.

Melse, L. (1990). *Schrijftoetsen voor de moderne vreemde talen*. Apeldoorn: Van Walraven.

Ministerie van OCW (2009). *Referentiekader taal en rekenen*. Almelo: Lulof druktechniek.

Ministerie van OCW (2011). *Voortgezet Onderwijs 2010-2011. Gids voor ouders, verzorgers en leerlingen*. Den Haag: Ministerie van OCW.

Mönks, F.J., & Knoers, A.M.P. (2009). *Ontwikkelingspsychologie. Inleiding tot de verschillende deelgebieden*. Assen: Van Gorcum.

Nas, C., Brugman, D. & Koops, W. (2005). Effects of the EQUIP programme on the moral judgment, cognitive distortions, and social skills of juvenile delinquents. *Psychology, Crime, and Law*, 11, 421-434.

Nederlandse Taalunie (2008). *Gemeenschappelijk Europees Referentiekader voor Moderne Vreemde Talen. Leren, Onderwijzen, Beoordelen*. Den Haag.

Nitko, A.J., & Brookhart, S.M. (2007). *Educational Assessment of Students*. Pearson Merrill Prentice Hall.

Oenen, S. van (2001). *Sociale competentie en de brede school*. Utrecht: NIZW Uitgeverij.

Onderwijsraad (2002). *Examens in ontwikkeling*. Den Haag: Onderwijsraad.

Onderwijsraad (2003). *Onderwijs en burgerschap*. Den Haag: Onderwijsraad.

Onderwijsraad (2005). *Sociale vorming en sociale netwerken in het onderwijs*. Den Haag: Onderwijsraad.

Onderwijsraad (2006). *Examinering: draagvlak en toegankelijkheid*. Den Haag: Onderwijsraad.

Onderwijsraad (2007). *Onderwijs en sociale samenhang: de stand van zaken*. Den Haag: Onderwijsraad.

Onderwijsraad (2008). *Onderwijs en maatschappelijke verwachtingen. Scholen kiezen zelfbewust positie*. Den Haag: Onderwijsraad.

Onna, M. van (2010). *Vas Cijfertools: toelichting voor externen*. Arnhem: Cito.

Orobio de Castro, B. (2000). Sociale informatieverwerking bij agressieve gedragsproblemen. Van fouten en vertekeningen tot aanpassingen aan aversieve omstandigheden. In: Gerris, J.R.M. (red.). *Jeugdzorg en probleemgedrag. Opvoedingswaarden en vernieuwingen in aanpak* (p. 59-75). Assen: Koninklijke Van Gorcum.

Pennewaard, L. (red.) (2011). *Schoolexamens vmbo. Grip krijgen op kwaliteit*. Enschede: SLO.

Phye, G.D. (1997). *Handbook of classroom assessment. Learning, achievement, and adjustment*. San Diego: Academic Press.

Platform VVVO (2008). *Tweede correctie is gekkenwerk geworden*. Persbericht van 29 mei 2008.

Online beschikbaar via:

www.platformvvvo.nl/brieven-archief/198-persbericht-integrale-tweede-correctie.html

Platform VVVO (2008). *Tijd nodig voor tweede correctie*. Online beschikbaar via:

www.platformvvvo.nl/brieven-archief/233-tijd-nodig-voor-integrale-tweede-correctie.html

Popham, J.W. (2004). *Classroom assessment: what teachers need to know (5th edition)*. Needham Heights, MA: Allyn & Bacon.

Rijn, P.W. van (2009a). Mogelijke effecten van verschillende uitslagregels. *Examens*, 6, 3, 5-8.

Rijn, P.W. van, Béguin, A., & Verstralen, H. (2009b). Zakken of slagen? De nauwkeurigheid van examenuitslagen in het voortgezet onderwijs. *Pedagogische Studiën*, 86, 185-195.

Roelofs, E., & Visser, J. (2011). De inhoud van toetsen. In: P. Sanders (red.). *Toetsen op School* (p. 21-44). Arnhem: Cito.

Sameroff, A. (2010). A unified theory of development: a dialectic integration of nature and nurture. *Child Development*, 81, 1, 6-22.

Sanders, P. (1983). Objectieve beoordeling van open-vragen examens. In: P. Weeda (red.). *Examens in discussie. Een bundel opstellen voor J.W. Solberg* (p. 163-172). Groningen: Wolters-Noordhof.

Sanders, P. (red.) (2011). *Toetsen op School*. Arnhem: Cito. Online beschikbaar via:

www.toetsenopschool.nl

Sanders, P. (2011). De betrouwbaarheid van toetsscores. In: P. Sanders (red.). *Toetsen op School* (p. 45-64). Arnhem: Cito.

Sanders, P., & Verstralen, H. (2011). Het beoordelen van toetsscores. In: P. Sanders (red.). *Toetsen op School* (p.142-155). Arnhem: Cito.

Scholte, E.M., & Ploeg, J.D. van der (2004). *Sociaal-emotionele vragenlijst (SEV). Handleiding*. Houten: Bohn Stafleu van Loghum.

Scholte, E.M., & Ploeg, J.D. van der (2005). *Handleiding ADHD-vragenlijst (AVL)*. Houten: Bohn Stafleu van Loghum.

Staphorsius, G. (red.) (2004). *Handleiding Toelatingstoets Leerwegerondersteund Onderwijs en Praktijkonderwijs*. Arnhem: Cito.

Staphorsius, G. (red.) (2001). *Wetenschappelijke verantwoording Toelatingstoets Leerwegerondersteunend Onderwijs en Praktijkonderwijs*. Arnhem: Cito.

Stiggins, R.J. (2001). *Student-involved classroom assessment (3rd edition)*. Upper Saddle River, NJ: Prentice Hall.

Thijs, A., & Akker, J. van den (2009). *Leerplan in ontwikkeling*. Enschede: SLO.

Til, A. van (2007). Het niveau van de luistervaardigheid in de moderne vreemde talen door de jaren heen. *Levende Talen Magazine*, 94, 3, 5-9.

Til, A. van (2011). *Verantwoording meting taal en rekenen 2010*. Arnhem: Cito.

VO-raad (2011). *Protocol eerste en tweede correctie centrale examens vmbo, havo en vwo*.

Wesseling, J., & Kuhlemeier, H. (2010). Een interactief beeldschermexamen voor het vmbo. *Examens*, 4, 17-20.

Zwitser, R., & Béguin, A. (2011). Het effect van de rekentoets op het percentage misclassificaties. Gaat meetfout bij de rekentoets slachtoffers maken? *Examens*, 4, 23-26.

Index

Index

A

absolute norm 15
absolute normering 16
adaptief toetsen 49
ADHD Vragenlijst (AVL) 119
adviesnormering 15
afname van het centraal examen 51, 59
afsluitingsfunctie 46
afwijkende opvattingen 109, 111
algemene vaardigheden 33–35
ankers 95
antwoordmodel 52, 53, 55–61, 66

B

backwash effecten 74
becijfering kandidaten 66, 69, 72
beeldschermexamens 48–49
beheersingsniveaus 14, 37, 42
Belevingsschaal voor Engels (BSE) 123
Belevingsschaal voor wiskunde (BSW) 123
beoordelaarsfouten 71, 115
beoordelen van scholen 10, 70, 75
beoordelingsaspecten 54, 55, 65, 67–69
beoordelingscriteria 65, 67, 90, 92, 96
beoordelingsmiddelen 65–69
beoordelingsmodel 52, 65, 67, 69
beroepsgerichte examens 50
betrouwbaarheid 49, 67, 71–73, 91, 95, 98, 119
betrouwbaarheid van een examenvak 73
betrouwbaarheid van examens 71–73
betrouwbaarheid van het diploma 73
betrouwbaarheid van het SE 67
bovenbouw 28, 31–35
burgerschap 102, 105–108, 123

C

can do statements 84
CBT 48
CBT-examen (Computer Based Testing) 49
centraal examen 12, 46
Centraal Praktisch Examen (CPE) 48
Centraal Schriftelijk en Praktisch Examen (CSPE) 48–55, 71
Centraal Schriftelijk Examen (CSE) 48, 55–62, 71
centrale rekentoets 13, 41
certificeren 12

cesuur 65, 68, 99
checklisten 51, 52
Checklists van de VO-raad 64
cijfers 20, 46, 48, 61, 62, 64–67, 70, 72, 75, 99
Cito Volgstelsysteem voortgezet onderwijs 10, 21, 121
civiel effect 10, 12, 46
classificeren 10
cohortgegevens 23
College voor Toetsen en Examens (CvTE) 13, 25, 47, 48, 51, 58
commissie-examens 47
Commissie Testaangelegenheden Nederland (COTAN) 71, 98, 119–123
Competentiebelevingsschaal voor adolescenten (CBSA) 121
competenties 23, 74, 93, 102–123
Compex-examens 48, 49
constructieopdracht 47
constructieteam 96
COOL5-18 23
correctievoorschrift 47, 50–56, 59, 80, 90, 92, 93, 96
CPE 48
CSE 48, 55–62, 71
CSPE 48–55, 71

D

descriptieve schaal 95
descriptoren 37
diagnostische toetsen 14, 22, 40
Dienst Uitvoering Onderwijs (DUO) 48
Docent gedragsbeoordelingslijst (DG) 114
doeltaal 87
doorlopende leerlijnen 13–14, 22, 24, 29, 36

E

educatieve domein 82
eerste corrector 58–62
efficiëntie 67
eindcijfers 70
eindexamens 12, 41, 47, 74, 77
eindtermen 28, 31–36, 65–67, 74, 103, 104
einduitslag 70
ERK-indicatie 14
ERK-proof 80, 81

Europees Referentiekader 14, 15, 80–99
Europees Taalportfolio 42
examen 46
Examenbesluit 47, 60, 70
exameneenheden 31–33
examenprogramma 28–34, 41, 50, 62–68,
72, 74, 81, 102, 104, 112
examenreglement 64
examensysteem 46
examenverslagen 72
externaliserende gedragsproblematiek 112

F

fraude 51, 63, 64
frequentieschaal 125, 126
fundamentele niveaus 35, 39

G

gedragscategorieën van een toetsmatrijs 83
geïntegreerd toetsen 81
gespreksopdracht 84
gewicht 64–70, 81, 96
grijze vlekken procedure 51

H

Handreikingen van SLO 64
herkansingenbeleid 50
Hoe Ik Denk Vragenlijst (HID) 120
huiswerkcontrole 20

I

ik kan stellingen 37, 42
inhaaltoetsen 67
inhoudscategorieën van een toetsmatrijs 82
inlevingsvermogen 109–111
instructie voor de examinerator 50–52
integraal nakijken 60
internaliserende gedragsproblematiek 112,
113
Interview Sociale Informatieverwerking 115

K

kerndoelen 28–32, 35, 102–104, 112
kernvaardigheden 10, 34
Kijk- en luistertoetsen 14–16
Klimaatschaal 121
kwaliteitsmonitor schoolexamens 63
kwaliteitszorg 30, 63
kwaliteit van het schoolexamen 63, 64, 69,
70

L

landelijke leerplankaders 28
Likert-schaal 125
lineair met knik 99

M

meerkeuzevragen 47, 72, 114
Meetinstrument Burgerschap 123
meetmomenten 97
methode van Angoff 99
Ministerie van OCW 25, 35
minitoetsen 50
mix van toetsvormen 82
moeilijkheidsgraad van examens 72
moeilijkheid toetsen en opdrachten 65, 66
mondeling examen 46
moreel redeneren 109–111
morele ontwikkeling 102, 105–111, 115, 125

N

Nederlandse Taalunie 80
niveaurapportage 11, 12
niveauspecificaties 65, 66
non-cognitieve doelen 102–104, 114, 117,
118, 124
normatieve schaal 94, 95
normering 13–16, 25, 48, 53, 65, 68, 69, 72,
98, 99
normeringsonderzoek 12
normeringsterm 48
normgroep 12

O

onafhankelijke correctoren 60
onderbouw 28–32
ondersteunende vaardigheden 80
onderwijsnummerbestanden 23
opbrengstenkaart 76

P

pedagogische functie 73
percentielscores 11, 22
persoonlijke domein 82
praktijkopdrachten 50, 51, 53
praktische opdrachten 49, 51, 52, 54, 62
prestatieoverzichten 23
prestatiestandaarden 35
preventie- en interventieprogramma's 111
probleemgedrag 103
producteisen 65, 67
producten voor het schoolexamen 63
productieve vaardigheden 81, 87, 92, 96

productschaal 95, 96
productspecificaties 47
professionele domein 82
programma van toetsing en afsluiting (PTA)
28, 64
protocollen 51, 52, 64
PTA 64
publieke domein 82
publieke verantwoording 30
puntentoekening 68

R

rapportcijfers 46
receptieve vaardigheden 96
referentieniveaus 11, 13, 22, 24, 28, 35–42,
70
rekenlijn 40
Rekentoets VO 13–14, 41
relatief normeren 12, 15
remediëring 21
rollenspel 117
rubriek 95

S

schaalpunten 93–95, 124
Schoolattitude Questionnaire Internet
(SAQI) 122
schoolexamen 12–15, 31–33, 46–76
schoolexamens kijk- en luistervaardigheid
12–15
schoolklimaat 102
Schoolvragenlijst (SVL) 122
schoolwerkplannen 28
scoreschaal 94
scoringsservice 14
simulatie 49
situatiebeschrijving 87
slaag-zakregels 70, 73
SMART-criteria 124
Sociaal-emotionele vragenlijst (SEV) 119
sociale competentie 23, 102, 105, 106, 108,
109, 112, 115, 117
sociale integratie 102, 106, 107, 108
sociometrische technieken 117
soepelheid corrector 52–60, 66, 68, 71, 75
Spearman-Brown formule 73
specifieke gedragsproblematiek 112
staatsexamens 47

standaard 13, 15, 16, 23, 25, 35, 64, 65, 98,
99
standaardenonderzoek 16
streefniveaus 35, 39
strengheid corrector 54–61, 66
Studeon 121
syllabus 31, 32, 50

T

taalbiografie 42
TaalstERK-toetsen 12, 15, 16
talenaspoort 42
toelatingsfunctie 46
toetsen van het leren 19
toetsen voor het leren 19
toetsingskader 64
toetsmatrijs 80, 82–86
toetsplan 80, 82
toetsvorm 19
Toezichtkader van de Onderwijsinspectie
70, 106–108
tweede correctie 51, 52, 54, 59–62, 72
tweede corrector 48, 55, 59–62
Tweede Fase 34

U

uitslagregeling 41

V

vaardigheidsscores 21
validiteit 67, 118–122
Vensters voor Verantwoording 23, 75
vergelijkbaarheid opgaven 65, 67
vernieuwing van het onderwijs 74
Vignet (VIG) 115
volledig open opdrachten 89
voorbeelduitwerkingen 29
voortgangscontrole 98
vraagvormen 47, 72
Vragenlijst Sterke Kanten en Moeilijkheden
(SDQ) 120
vragen stellen 19

Z

Zelfrapportage met leerlingenvragenlijst
(LV) 114
zesjescultuur 70

Cito

Corporate

Toetsen op School

Voortgezet onderwijs

Hans Kuhlemeier en Alma van Til

Piet Sanders (redactie)

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

