

Advies

Toets wijzer

Naar een eigen(tijdse) wijze van toetsen en examineren

Toets wijzer

Naar een eigen(tijdse) wijze van toetsen en examineren

Colofon

De Onderwijsraad is het wettelijk adviesorgaan van de regering en het parlement op het terrein van het onderwijs. De raad adviseert, gevraagd en ongevraagd, over hoofdlijnen van beleid en wetgeving. Hij adviseert over het brede terrein van het onderwijs, van voorschoolse educatie tot postuniversitair onderwijs en een leven lang leren. Gemeenten kunnen in speciale gevallen van lokaal onderwijsbeleid een beroep doen op de Onderwijsraad.

De raad is samengesteld uit personen die zijn benoemd op grond van hun gezaghebbende positie in relevante wetenschappelijke disciplines, in sectoren van het onderwijs, in openbaar bestuur of in maatschappelijke organisaties. De raad gebruikt in zijn advisering verschillende disciplinaire aspecten (zoals onderwijskundige, economische en juridische) en verbindt deze met ontwikkelingen in de praktijk van het onderwijs.

De producten van de raad worden gepubliceerd in de vorm van adviezen, studies en verkenningen. Daarnaast organiseert de raad discussies over onderwerpen die van belang zijn voor het onderwijsbeleid.

Advies *Toets wijzer*, uitgebracht aan de Minister van Onderwijs, Cultuur en Wetenschap en aan de Minister voor Basis- en Voortgezet Onderwijs en Media.

Registratienummer: 20180284/1139, december 2018.

Uitgave van de Onderwijsraad, Den Haag, 2018.

ISBN: 978-946121-066-1

Bestellingen van publicaties:

Onderwijsraad
Prins Willem Alexanderhof 20
2595 BE Den Haag
email: secretariaat@onderwijsraad.nl
telefoon: (070) 310 00 00 of via de website:
www.onderwijsraad.nl

Ontwerp en opmaak:

www.balyon.com

Drukwerk:

Drukkerij Excelsior, Den Haag

© Onderwijsraad, Den Haag.

Alle rechten voorbehouden. All rights reserved.

Aan de Minister van Onderwijs, Cultuur en Wetenschap en
de Minister voor Basis- en Voortgezet Onderwijs en Media
Mevrouw mr. drs. I.K. van Engelshoven
De heer drs. A. Slob
Postbus 16375
2500 BJ Den Haag

Prins Willem Alexanderhof 20
2595 BE Den Haag

Telefoon: 070 310 00 00
Fax: 070 356 14 74
secretariaat@onderwijsraad.nl
www.onderwijsraad.nl

Ons kenmerk
20180284/1139

Contactpersoon

Plaats/datum
Den Haag, 13 december 2018

Uw kenmerk

Doorkiesnummer

Onderwerp
Advies *Toets wijzer*

Mevrouw de Minister, Mijnheer de Minister,

Uw voorgangers hebben de Onderwijsraad advies gevraagd over toetsing en examinering. Met veel genoegen biedt de raad u hierbij het advies *Toets wijzer* aan.

De raad adviseert de rijksoverheid en de onderwijsinstellingen om toe te werken naar een evenwichtiger toets- en examenpraktijk. Binnen de huidige toets- en examenpraktijk is te weinig ruimte voor formatieve toetsing. Daarnaast wordt overwegend gebruikgemaakt van kwantitatieve toetsing. Ten slotte ligt er veel nadruk op centraal ontwikkelde toetsing en examinering.

De rijksoverheid dient vanuit haar stelselverantwoordelijkheid duidelijke kaders te bieden aan onderwijsinstellingen. Deze overheidskaders vormen het tegenwicht voor de eigen toetsing van onderwijsinstellingen en bepalen daarmee hun ruimte. De raad beschouwt om te beginnen een consistent toets- en examenbeleid als duidelijk kader. De overheid dient rolvast en standvastig te zijn in haar maatregelen. Met betrekking tot toetsing en examinering ziet de raad voor alle sectoren mogelijkheden tot verbetering van het toezicht. De raad adviseert de rijksoverheid om voor primair en voortgezet onderwijs een toets- en/of examencommissie te verplichten.

Voor onderwijsinstellingen is het volgens de raad van belang dat toetsing in lijn wordt gebracht met onderwijsdoelen, onderwijsinhoud en onderwijsmiddelen. De raad adviseert hierover uitgebreider in het briefadvies *Curriculumvernieuwing*.

Bovengenoemde verbeteringen in de toets- en examenpraktijk kunnen alleen gestalte krijgen als de toetsdeskundigheid in het onderwijs groter wordt en wordt gedeeld met leerlingen en studenten en hun ouders.

Met beleefde groet,



Prof. dr. H. Maassen van den Brink
Voorzitter



Drs. C. van 't Veen
Secretaris ad interim

Inhoud

Samenvatting	7
1 Aanleiding: toenemend debat en discussie rondom toetsing	9
1.1 Er zijn hoge verwachtingen van toetsing	9
1.2 Verwachtingen leiden tot breed gevoerd debat over kwaliteit van toetsing	10
1.3 Adviesvraag: hoe kan het gebruik van toetsing en examinering worden verbeterd?	12
1.4 Totstandkoming van dit advies	12
2 Hoofdboodschap: werk toe naar een evenwichtiger toetspraktijk	13
2.1 Een evenwichtige toetspraktijk draagt bij aan onderwijskwaliteit	13
2.2 Huidige toetspraktijk is niet evenwichtig	16
Voorbeelden van toetspraktijk uit evenwicht	19
Primair onderwijs	19
Voortgezet onderwijs	20
Middelbaar beroepsonderwijs	22
Hoger onderwijs	24
3 Overheid en instellingen: baseer de toetspraktijk op een heldere visie	27
3.1 Overheid: bied duidelijke kaders aan de onderwijsinstellingen	27
3.2 Onderwijsinstelling: neem een standpunt in op belangrijke toetsvraagstukken	35
4 Vergroot de toetsdeskundigheid in het onderwijs	40
4.1 Actualiseer professionele standaarden op het gebied van toetsing	40
4.2 Geef meer aandacht aan toetsing in professionaliseringstrajecten	43
4.3 Deel informatie over toetsing met andere belanghebbenden	45
Afkortingen	47
Literatuur	48
Geraadpleegde deskundigen	55

Samenvatting

Het belang van toetsing en examinering voor de onderwijspraktijk wordt algemeen erkend. Dit leidt tot hoge verwachtingen: toetsing moet bijvoorbeeld het onderwijsleerproces ondersteunen, onbetwistbare ijkpunten bieden binnen het onderwijsstelsel, en meebewegen met veranderingen in de onderwijspraktijk. Ondanks toenemende aandacht voor het verbeteren van de toets- en examenkwaliteit blijft er discussie en debat rondom het thema. Dit heeft geleid tot de volgende adviesvraag van het ministerie van OCW (Onderwijs, Cultuur en Wetenschap) aan de Onderwijsraad: hoe kan het gebruik van toetsing en examinering in het onderwijs worden verbeterd?

Verbeteren toetspraktijk vraagt om meer evenwicht binnen drie dimensies

De raad adviseert de rijksoverheid en de onderwijsinstellingen om toe te werken naar een evenwichtiger toets- en examenpraktijk. De raad heeft de huidige toets- en examenpraktijk in alle onderwijssectoren bestudeerd en concludeert dat deze binnen drie dimensies uit balans is. Om te beginnen is er te weinig ruimte voor formatieve toetsing door grote nadruk op beslissende toetsing. Daarnaast maakt het onderwijs overwegend gebruik van kwantitatieve toetsing, waardoor kwalitatieve toetsmethoden onvoldoende aan bod komen. Dit geldt vooral voor het primair en voortgezet onderwijs, maar ook het middelbaar beroepsonderwijs en hoger onderwijs neigen naar meer kwantitatieve toetsing. Ten slotte ligt er veel nadruk op centraal ontwikkelde toetsing en examinering, waardoor er decentraal (in de onderwijsinstellingen zelf) te weinig eigenaarschap en deskundigheid aanwezig is.

Evenwicht aanbrengen begint met een heldere visie op toetsing- en examinering

Een evenwichtige toets- en examenpraktijk vraagt zowel van de rijksoverheid als van onderwijsinstellingen een heldere visie. Vanuit deze visie en de daaruit voortvloeiende verantwoordelijkheden kunnen zowel overheid als instellingen duidelijke kaders bieden, zodat alle betrokkenen de bestaande ruimte voor toetsing en examinering beter kunnen benutten.

Overheid: bied duidelijke kaders aan de onderwijsinstellingen

De rijksoverheid dient vanuit haar stelselverantwoordelijkheid duidelijke kaders te bieden aan onderwijsinstellingen. Deze kaders vormen het tegenwicht voor de eigen toetsing van de onderwijsinstellingen en bepalen daarmee hun ruimte. De raad beschouwt om te beginnen een consistent toets- en examenbeleid als duidelijk kader: de overheid dient rolvast en standvastig te zijn in haar maatregelen. Daarnaast vindt de raad standaardisering van de eindtoetsing voor alle sectoren van belang. Gestandaardiseerde toetsing van de ijkpunten biedt basis- en middelbare scholen ruimte om hun schooleigen toetsing en examinering vorm te geven zoals dat past bij hun eigen onderwijsvisie. Maar ook voor het hoger onderwijs vindt de raad enige mate van standaardisering van eindtoetsing, bijvoorbeeld van bachelorexamens van gelijke opleidingen, het overwegen waard. Een derde kader voor de onderwijsinstellingen wordt gevormd door het toezicht, dat duidelijk maakt aan welke doelen moet worden gewerkt. Met betrekking tot toetsing en examinering ziet de raad voor alle sectoren mogelijkheden tot verbetering van het toezicht.

Als laatste kader adviseert de raad de rijksoverheid om voor primair en voortgezet onderwijs, net als voor middelbaar beroepsonderwijs en hoger onderwijs, toets- en/of examencommissies te verplichten. Daarmee kan de borging van de kwaliteit van toetsing en examinering in deze sectoren verbeterd worden en het eigenaarschap worden vergroot.

Onderwijsinstelling: neem een standpunt in op belangrijke toetsvraagstukken

Voor onderwijsinstellingen is het van belang dat toetsing in lijn wordt gebracht met onderwijsdoelen, onderwijsinhoud en onderwijsmiddelen. Deze integrale benadering van curriculum- en toetsontwikkeling ondersteunt het onderwijsleerproces optimaal en gaat versmalling van het curriculum tegen. Bij het vormgeven van deze eigen visie op toetsing en onderwijsdoelen adviseert de raad om meer ruimte te geven aan de formatieve functie van toetsing, om zo het lesgeven van leraren/docenten en het leerproces van leerlingen/studenten beter te ondersteunen. Daarnaast vindt de raad dat beter en meer gebruikgemaakt kan worden van kwalitatieve toetsmethoden. Ook adviseert de raad om de positie van de examencommissies in het hoger onderwijs nog verder te verbeteren.

Voor een evenwichtige toets- en examenpraktijk is meer toetsdeskundigheid noodzakelijk

Bovengenoemde verbeteringen in de toets- en examenpraktijk kunnen alleen gestalte krijgen als de toetsdeskundigheid in het onderwijs groter wordt. Om dit te realiseren adviseert de raad de rijksoverheid en de beroepsverenigingen om toetsing meer en beter aandacht te geven in de beroepsstandaarden van verschillende groepen onderwijsprofessionals. Hiertoe dienen betere bekwaamheidseisen en richtlijnen op het gebied van toetsing te worden geformuleerd voor leraren en docenten. Deze zijn momenteel onvolledig en weinig concreet, behalve voor het hoger beroepsonderwijs. De daar gehanteerde basis- en seniorkwalificatie examinering zouden als voorbeeld kunnen dienen voor andere sectoren. Ook voor andere groepen onderwijsprofessionals, zoals schoolleiders en lerarenopleiders, dienen standaarden met betrekking tot toetsdeskundigheid te worden geformuleerd. Daarnaast beveelt de raad aan om meer, maar ook gericht aandacht te geven aan toetsing in lerarenopleidingen en professionaliseringstrajecten. Hierdoor kunnen veel misverstanden over toetsing en examinering worden weggenomen. Als laatste brengt de raad het delen van kennis onder de aandacht. Ook voor leerlingen en studenten en hun ouders is het nuttig om meer inzicht te hebben in het brede belang van toetsing.

Het belang van toetsing voor de onderwijspraktijk wordt algemeen onderkend. Ondanks toegenomen aandacht voor het verbeteren van de kwaliteit van toetsing en examinering, blijft er veel discussie en debat rondom het thema. Hoe kunnen de verwachtingen van toetsing en examinering worden waargemaakt?

1 **Aanleiding: toenemend debat en discussie rondom toetsing**

Toetsing¹ is in alle onderwijssectoren onlosmakelijk verbonden met de praktijk van alledag. De resultaten van toetsen geven informatie die onontbeerlijk is voor het onderwijsleerproces. Er bestaat geen discussie over het belang van toetsing, maar zodra het gaat over de nadere invulling lopen de meningen uiteen.

1.1 **Er zijn hoge verwachtingen van toetsing**

Onderwijsveld, politiek en maatschappij koesteren hoge verwachtingen van toetsing. Toetsen moeten altijd van onbetwistbare kwaliteit zijn, omdat ze de basis vormen voor belangrijke beslissingen over de toekomst van leerlingen en studenten. Dat toetsing aansluit bij een breed curriculum, waarin niet alleen ruimte is voor kennisoverdracht, maar ook voor doelen op het gebied van socialisatie en persoonsvorming, is voor velen vanzelfsprekend. Verder moet toetsing het onderwijsleerproces ondersteunen, zowel voor leerlingen en studenten als voor leraren en docenten. En ten slotte dient toetsing mee te bewegen met ontwikkelingen op het gebied van personalisering en digitalisering van het onderwijs. Dergelijke verwachtingen leiden tot brede aandacht voor het verbeteren van de kwaliteit van toetsing.

Zo is er in alle sectoren aandacht voor het belang van opbrengstgericht werken, waarbij toetsgegevens een centrale rol innemen. Ook formatief handelen (het ondersteunen van het onderwijsleerproces aan de hand van toetsresultaten) kan rekenen op belangstelling in alle sectoren. In het primair onderwijs zijn daarnaast de ogen gericht op de eindtoets en leerlingvolgtoetsen. In het voortgezet onderwijs geeft de kwaliteit van de centrale examens en van de schoolexamens ieder jaar weer discussie, zowel onder leerlingen als onder vakdocenten en toetsdeskundigen. In het middelbaar beroepsonderwijs wordt vooral gewerkt aan het verbeteren van de kwaliteit van de praktijkexamens. Het hoger onderwijs houdt zich bezig met verbetering van de interne kwaliteitszorg rondom toetsing en examinering.

¹ Omwille van de leesbaarheid wordt in dit advies de term toetsing gebruikt, ook waar zowel toetsing als examinering wordt bedoeld. Waar het specifiek over examinering gaat, wordt dit aangegeven. Zie ook het kader in dit hoofdstuk.

Wat verstaat de Onderwijsraad onder toetsing?

Toetsing is het gericht verzamelen, interpreteren en gebruiken van informatie over kennis, vaardigheden of houding van een leerling of student, doorgaans aan de hand van één of meerdere opdrachten. Wanneer aan het eind van een opleiding prestaties worden vastgesteld, is er sprake van afsluitende toetsing (ook wel eindtoetsing genoemd). In het hoger onderwijs wordt doorgaans gesproken van tentaminering in plaats van toetsing.

Een specifieke vorm van afsluitende toetsing is examinering. Hierbij wordt op basis van toets- of tentamenprestaties een certificaat of diploma toegekend met een civiel effect. Dat laatste wil zeggen dat de leerling of student met het certificaat of diploma in aanmerking komt voor specifieke vormen van vervolgonderwijs of voor functies in de maatschappij.

1.2 Verwachtingen leiden tot breed gevoerd debat over kwaliteit van toetsing

Toetsing moet dus voldoen aan hoge verwachtingen en er is veel aandacht voor de kwaliteit ervan. Het behoeft daarom niet te verbazen dat toetsing onderwerp is van discussie, zowel in het onderwijsveld als in de politiek en breder in de maatschappij. Deze discussies gaan over verschillende aspecten van de toetspraktijk.

Een eerste thema in het debat is de nadruk op gestandaardiseerde toetsing in het onderwijs. Met gestandaardiseerde toetsen is het mogelijk om leerlingen, klassen en scholen met elkaar te vergelijken. Voorstanders van gestandaardiseerde toetsing beargumenteren dat dit een opbrengstgerichte aanpak van het onderwijs stimuleert en daarmee bijdraagt aan kwaliteitsverbetering. In het onderwijs worden gestandaardiseerde toetsen steeds vaker verplicht gesteld. In het primair onderwijs gaat het daarbij om een gestandaardiseerde eindtoets en gestandaardiseerde voortgangstoetsen. In het voortgezet onderwijs zijn de exameneisen aangescherpt. Gestandaardiseerde toetsing wordt vooral, maar niet alleen, gebruikt wanneer toetsing een beslissende functie heeft. Gestandaardiseerde toetsing in het funderend onderwijs geeft volgens de voorstanders duidelijke referentiepunten voor kwaliteitsbeoordeling, doorstroom en civiel effect van diploma's.²

De nadruk op beslissende, gestandaardiseerde en kwantitatieve toetsing leidt tot een tegenbeweging. Tegenstanders stellen dat deze nadruk een strategische aanpak uitlokt, die leidt tot 'versmalling' van de onderwijskwaliteit, aantasting van de pluriformiteit van het onderwijs, en excessieve aandacht voor toetsen.³ Sommige van deze tegenstanders wijzen toetsing volledig af of pleiten voor alleen nog formatief toetsen. Een meer gematigde stellingname is dat toetsing ten dienste moet staan van de brede opbrengsten die het onderwijs nastreeft op de terreinen van kwalificatie, socialisatie en persoonsvorming. Dat betekent dat naast cognitieve, vaak kwantificeerbare opbrengsten ook andersoortige opbrengsten zichtbaar moeten worden gemaakt, zo nodig op kwalitatieve wijze.⁴

Een aan standaardisering gerelateerd thema in het debat is het centraliseringsvraagstuk. Wat moet er centraal worden getoetst en wat mogen onderwijsinstellingen zelf vormgeven qua toetsing en examinering? Dit onderwerp is nauw gerelateerd aan dat van autonomie van onderwijsinstellingen en krijgt veel aandacht in het debat. Welke ijkpunten zijn noodzakelijk om die autonomie te kunnen blijven garanderen? Hoe kan de overheid haar stelselverant-

² Onderwijsraad, 2015a.

³ Berliner, 2011.

⁴ Onderwijsraad, 2016a.

woordelijkheid voor de kwaliteit van het onderwijs centraal waarmaken? Het hoger onderwijs heeft bijvoorbeeld veel ruimte voor eigen invulling. Dit is positief voor het eigenaarschap en de betrokkenheid van docenten, maar brengt uitdagingen met zich mee voor de vergelijkbaarheid van diploma's en mogelijke nadelige gevolgen voor gelijke kansen op de arbeidsmarkt.

Een ander thema in het debat is de toenemende verantwoording die onderwijsinstellingen moeten afleggen over de kwaliteit van de toetsen die zij voor beslissende doeleinden gebruiken. In het hoger onderwijs zijn de afgelopen jaren de eisen aan examencommissies en aan de deskundigheid van hun leden aangescherpt. De verantwoordingsplicht heeft consequenties voor de vorm van de toetsing. Toetsen moeten zowel valide als betrouwbaar zijn en de afname moet objectiveerbaar en kwantificeerbaar zijn. Dat is echter niet altijd mogelijk. Door de verantwoordingsplicht kunnen toetsen waarbij leerlingen bijvoorbeeld houdingsaspecten en vaardigheden in (semi)realistische beroepssituaties demonstreren, onder druk komen te staan. Opleidingen kiezen vanwege de verantwoordingsplicht namelijk eerder voor toetsen waarbij de nadruk ligt op betrouwbaarheid en kwantificeerbaarheid.

Ook is er in het debat aandacht voor de groeiende variatie in toetsen. Dat heeft onder andere te maken met de omslag van een aanbodgestuurde naar een meer vraaggestuurde markt van leermiddelen.⁵ Scholen voor primair onderwijs kunnen naast de centrale eindtoets van het Cito nu bijvoorbeeld ook kiezen voor één van de vijf alternatieve erkende eindtoetsen. Ook voor leerlingvolgtoetsen zijn er verschillende aanbieders. Bovendien hanteert elke methode in het primair en voortgezet onderwijs zijn eigen toetsing. In het middelbaar beroepsonderwijs is sprake van ontwikkeling van toetsing door externe organisaties. In het hoger onderwijs is het mogelijk dat iedere docent zijn of haar eigen toetsen maakt. De markt voor toetsing en examinering is niet altijd transparant en dit kan gevolgen hebben voor de kwaliteit en de doorstroom naar vervolgonderwijs.

Een laatste belangrijk thema in het debat is de digitalisering van toetsen. Digitalisering kan uiteenlopende vormen aannemen en verschillende consequenties hebben voor de toetsing. Leerlingen/studenten en leraren/docenten kunnen eigenaar worden van de toetsing omdat zij zelf de aard, het moment en de locatie ervan kunnen bepalen, alsmede de toetsgegevens kunnen beheer in de vorm van een digitaal portfolio. Daarnaast kunnen digitale toetsen een adaptief karakter hebben, wat betekent dat de moeilijkheid van toetsvragen wordt aangepast aan de gegeven antwoorden.⁶

Naast bovengenoemde inhoudelijke discussiepunten klinken er kritische geluiden over de veelheid aan toetsen die leerlingen en studenten moeten maken en de toetsdruk die hierdoor kan ontstaan, over het belang dat aan toetsing wordt gehecht, en over de mogelijk negatieve gevolgen van toetsing. De gezamenlijke noemer in bovenstaande kritiek- en discussiepunten is dat toetsing een doel op zich lijkt te zijn geworden, waarbij nut en noodzaak buiten beeld verdwijnen.

⁵ Onderwijsraad, 2010b.

⁶ Het blijkt echter niet eenvoudig om digitale adaptieve toetsen vorm te geven. Het College voor Toetsen en Examens kondigde recent aan dat de digitale adaptieve centrale eindtoets in 2019 niet beschikbaar is voor scholen. Dat komt doordat het systeem onvoldoende stabiel is om een goede afname van de toets te garanderen. Zie College voor Toetsen en Examens, 2018c.

1.3 **Adviesvraag: hoe kan het gebruik van toetsing en worden verbeterd?**

De toenemende discussie over verschillende aspecten van toetsing en examinering heeft geleid tot de adviesvraag van het ministerie van OCW (Onderwijs, Cultuur en Wetenschap) over de rol en invulling van toetsing en examinering in het onderwijs.

De adviesvraag luidt: *Hoe kan het gebruik van toetsing en examinering in het onderwijs worden verbeterd?*

De raad gaat in dit advies in op toetsing en examinering in alle sectoren, van het primair onderwijs tot en met het wetenschappelijk onderwijs. Het accent in het advies ligt op de rol en verantwoordelijkheden van betrokkenen uit het onderwijsveld en van de rijksoverheid. De raad geeft daarbij richtlijnen over visie, eigenaarschap, verantwoordelijkheden en deskundigheid die de kwaliteit van toetsing en examinering ten goede komen.

1.4 **Totstandkoming van dit advies**

Voor het tot stand komen van dit advies is literatuuronderzoek verricht en een analyse van relevante beleidsdocumenten gemaakt. Er zijn panelgesprekken gevoerd met diverse groepen betrokkenen uit alle onderwijssectoren, waaronder leraren en docenten, schoolleiders/directeuren en bestuurders. Daarnaast zijn individuele gesprekken gevoerd met diverse andere deskundigen, waaronder vertegenwoordigers van belangen- en sectororganisaties. Er heeft een onderwijsdialoog plaatsgevonden met inleiders, discussianten en deelnemers uit alle sectoren. Ook heeft de Onderwijsraad op zijn website een oproep gedaan om over het advies mee te denken, waarop verschillende reacties zijn verkregen. Een overzicht van gesprekspartners is opgenomen achterin dit advies. Tot slot zijn diverse bijeenkomsten bezocht rondom het thema toetsing en examinering.

Het verbeteren van de huidige toets- en examenpraktijk vraagt om evenwicht. Evenwicht tussen formatieve en beslissende toetsing, tussen centrale en decentrale toetsing, en tussen kwantitatieve en kwalitatieve toetsing.

2 Hoofdboodschap: werk toe naar een evenwichtiger toetspraktijk

De toetspraktijk krijgt vorm door uiteenlopende keuzes die op verschillende niveaus van het onderwijssysteem worden gemaakt. De raad adviseert de rijksoverheid en onderwijsinstellingen om geen eenzijdige keuzes te maken, maar te streven naar een evenwichtige toetspraktijk. Hierin krijgen bijvoorbeeld zowel de borging van het civiel effect van diploma's als het genereren van rijke feedback voor het leerproces een plaats. De raad constateert dat de huidige toetspraktijk in verschillende opzichten uit balans is. Om deze balans aan te brengen is een integrale benadering nodig, met als belangrijkste doel dat toetsing een bijdrage levert aan de onderwijskwaliteit, aan de onderwijsloopbaan van leerlingen en studenten en aan toekomstig functioneren op de arbeidsmarkt en in de maatschappij.

2.1 Een evenwichtige toetspraktijk draagt bij aan onderwijskwaliteit

De raad pleit voor een evenwichtiger toetspraktijk die beter ten dienste staat van de onderwijskwaliteit. Deze toetspraktijk dient verschillende doelen. Voor leerlingen en studenten heeft toetsing als doel om feedback te krijgen over hun leerproces en om inzicht te krijgen in hun vorderingen. Ook speelt toetsing een belangrijke rol bij de overgangen in hun onderwijsloopbaan. Voor leraren en docenten biedt toetsing inzicht in hoe hun onderwijs aansluit bij de leerlingen en studenten en waar verbeterpunten liggen. Voor onderwijsinstellingen geeft toetsing zicht op de kwaliteit van het onderwijs en ondersteunt het beslissingen over bevordering naar een volgend leerjaar en toekenning van een diploma. Daarmee heeft toetsing ook op maatschappelijk niveau een belangrijk doel, namelijk toeleiding naar vervolgopleidingen of de beroepspraktijk.⁷ Bovendien gebruikt de rijksoverheid toetsing om zicht te krijgen op de kwaliteit van het onderwijs, zowel van individuele onderwijsinstellingen als van het stelsel in zijn totaliteit.

Deze doelen van toetsen en examineren staan soms op gespannen voet met elkaar. De raad onderkent bij deze spanningen drie dimensies:

- de mate waarin toetsing een beslissende of formatieve functie heeft;
- de mate waarin toetsing op decentraal niveau of (meer) centraal niveau wordt vormgegeven; en
- de mate waarin wordt gestreefd naar kwantitatieve meting of naar het meer op kwalitatieve wijze zichtbaar maken van onderwijsopbrengsten.

⁷ Onderwijsraad, 2018b.

Wat de juiste balans is binnen de bovengenoemde dimensies verschilt per onderwijssector, maar voor elke sector geldt dat een balans binnen elk van de dimensies de beste garantie is dat toetsing daadwerkelijk een bijdrage leveren aan de onderwijskwaliteit en de onderwijs- en maatschappelijke loopbaan van leerlingen en studenten.

De drie dimensies kunnen niet los van elkaar bekeken worden en vragen om een integrale benadering. Veranderingen binnen de ene dimensie hebben gevolgen voor het evenwicht binnen een andere dimensie. Het is in de praktijk wel vaak zo dat centrale, beslissende en kwantitatieve toetsing samengaan, evenals formatieve, decentrale en kwalitatieve toetsing. Een onderliggende dimensie is dan standaardisering (gelijkschakeling) versus variatie. Toch zijn er ook andere combinaties mogelijk. Zo vindt het centraal schriftelijk en praktisch examen in het vmbo (voorbereidend middelbaar beroepsonderwijs) grotendeels op kwalitatieve wijze plaats, maar wel aan de hand van centraal vormgegeven praktische toetsopdrachten. En gestandaardiseerde kwantitatieve toetsen kunnen ook voor formatieve doeleinden worden gebruikt, zoals gebeurt bij digitale adaptieve toetsen.⁸

Het pleidooi voor een balans binnen deze drie dimensies betekent niet dat de raad voorstelt om de huidige toetspraktijk radicaal anders in te richten. Hij vindt het van belang dat het goede uit de huidige toetspraktijk behouden blijft. Tegelijkertijd dienen er maatregelen genomen te worden die het evenwicht verbeteren.

Formatieve versus beslissende functies van toetsing

Toetsing kan uiteenlopende functies vervullen.⁹ Deze functies kunnen grofweg in twee categorieën worden ingedeeld: beslissende en formatieve functies. De beslissende functies worden in het onderwijsveld ook wel aangeduid als summatieve functies.¹⁰ De raad kiest voor de term beslissende functies in plaats van summatieve functies, omdat er verwarring over de term summatief bestaat. Formatieve toetsing kan namelijk ook gebaseerd worden op summatieve oordelen, de twee sluiten elkaar niet uit.¹¹

Beslissende functies verwijzen naar het gebruik van resultaten van toetsen of examens om beslissingen te nemen. Beslissingen kunnen in de eerste plaats betrekking hebben op leerlingen/studenten. Het gaat dan om zaken als bevordering, toelating, selectie en certificering. Beslissingen op basis van toetsresultaten kunnen daarnaast betrekking hebben op docenten of onderwijsinstellingen. Het gaat dan bijvoorbeeld over het beoordelen van het functioneren van onderwijspersoneel, de toekenning van financiën en het opleggen van maatregelen om de onderwijskwaliteit te verbeteren. In alle gevallen staat er iets op het spel.

Bij de formatieve functies van toetsing gaat het om initiëren, vormgeven en optimaliseren van het onderwijs. Toetsen worden dan gebruikt om specifieke leerbehoeften van leerlingen vast te stellen, feedback te geven en daarmee uiteindelijk het onderwijs te verbeteren. Toetsing is in dit opzicht zowel ondersteunend voor het zelfstandig leren van leerlingen als voor de interne kwaliteitszorg van het onderwijs. Toetsing kan het leer- en onderwijsproces tot op zekere hoogte ook initiëren doordat het richting geeft (er wordt naar de toets of het examen toegewerkt, de toetsen en examens functioneren als een concretisering van de leerdoelen) en moti-

⁸ Onderwijsraad, 2017a.

⁹ Archer, 2017; Newton, 2007; Newton, 2017.

¹⁰ Bloom, Hastings, & Madaus, 1971; Black, 1998.

¹¹ Newton, 2007.

veren (de toets vormt een uitdaging, men spant zich tijdens en voorafgaand aan de toets in om het toetsresultaat te optimaliseren).

Decentrale versus (meer) centrale toetsing

Bij het bepalen van het toetsbeleid moeten keuzes gemaakt worden ten aanzien van vormgeving (het ontwerp, de afname, de beoordeling en het gebruik van toetsresultaten), inhoud (als weerspiegeling van de onderwijsdoelen), kwaliteitsbeleid en organisatie. Deze keuzes kunnen decentraal door de leraar/docent of meer op centraal niveau worden gemaakt, bijvoorbeeld door een groep van instellingen of nog centraler door een instantie van of namens de rijksoverheid.

Zowel centralisering als decentralisering van toetsing heeft voordelen. Decentralisering van de toetsing maakt het mogelijk de toets op maat te snijden van de specifieke groep leerlingen/studenten en het onderwijs dat ze hebben gevolgd. Leraren/docenten blijven daarnaast zeggenschap houden over een belangrijk onderdeel van hun onderwijs. Dat draagt bij aan het gevoel van eigenaarschap en heeft een motiverende werking.¹² Decentrale vormgeving van toetsing versterkt bovendien de mogelijkheden voor het aanbrengen van eigen accenten door onderwijsinstellingen en individuele leraren/docenten. Wanneer een docent zelf zijn of haar toetsen vormgeeft, is er sprake van volledige decentralisering.¹³ Er zijn echter ook tussenvormen, bijvoorbeeld als docenten worden betrokken bij de keuze van toetsen en de manier waarop zij worden ingezet. Dit is bijvoorbeeld het geval in het hoger onderwijs, via examen- of opleidingscommissies. Inspraak bij de vormgeving, de organisatie en de kwaliteitszorg kan ook bijdragen aan eigenaarschap en maakt de toetsing ten dele decentraal.

Een belangrijk voordeel van centralisering van de toetsing is de standaardisering. Gestandaardiseerde toetsing op de ijkpunten, zoals bijvoorbeeld in het funderend onderwijs gangbaar is, maakt het mogelijk om toetsprestaties van instellingen met elkaar te vergelijken. En de individuele leerling/student kan eigen toetsresultaten vergelijken met een grotere populatie. Bij digitale adaptieve toetsvormen zijn dergelijke vergelijkingen in het computerprogramma verwerkt, zodat leerlingen/studenten automatisch een afgestemd aanbod van feedback, instructie en nieuwe toetsopgaven krijgen.¹⁴ Vergelijkbare resultaten maken het ook mogelijk voor onderwijsinstellingen, de rijksoverheid en de maatschappij als geheel om door de tijd heen zicht te houden op de onderwijskwaliteit en kwaliteitsbeleid te voeren. Een ander voordeel van standaardisering is dat dit een objectieve en transparante basis kan bieden bij belangrijke beslissingen, zoals het geven van schooladviezen of toekennen van diploma's. Dit is zowel voor individuele leerlingen/studenten als voor de maatschappij in het algemeen van belang: diploma's hebben op die manier een civiel effect.¹⁵

Kwantitatieve versus kwalitatieve methoden van toetsen

Toetsen kan op kwantitatieve of kwalitatieve wijze. Een toets met gesloten vragen en een vast correctiemodel is kwantitatief. De beoordeling is te herleiden tot (deel)scores en het maakt in principe voor de beoordeling niet uit wie de toets nakijkt. Dat is anders bij kwalitatieve toetsen, zoals een mondeling tentamen, een presentatie of een werkstuk. De beoordelaar moet hierbij beschrijvende informatie of observaties interpreteren. Deze interpretaties vergen inhoudelijke

¹² Onderwijsraad, 2016b.

¹³ Leraren en docenten hebben met betrekking tot deze beoordelingsbevoegdheid een sterke juridische positie. Zie WPO artikel 31a, lid 2; WVO artikel 32e, lid 2; WEB artikel 4.1a.1, lid 2; WHW artikel 7:12c.

¹⁴ Onderwijsraad, 2017a.

¹⁵ Onderwijsraad, 2015a.

expertise en zijn tot op zekere hoogte subjectief. De interpretaties kunnen beschrijvend zijn of uiteindelijk wel weer naar scores of deelscores worden vertaald.¹⁶

Lange tijd is gedacht dat het voor het verkrijgen van betrouwbare en objectieve informatie noodzakelijk was toetsing vergaand te standaardiseren en te kwantificeren en de rol van de beoordelaar te minimaliseren. Een recenter inzicht is dat betrouwbaarheid niet noodzakelijk standaardisatie vereist en vooral het resultaat is van het combineren van verschillende toetsen, bij voorkeur met gevarieerde methoden.¹⁷ Bovendien wordt objectiviteit niet noodzakelijkerwijs verkregen door de rol van de beoordelaar te minimaliseren, maar kan objectiviteit ook worden bereikt door verschillende beoordelaars in te zetten, studenten onderling elkaars werk te laten vergelijken, trainingen te geven, of producten van leerlingen/studenten paarsgewijs te vergelijken.¹⁸

Kwantitatieve en kwalitatieve toetsen kunnen beide bijdragen aan het zichtbaar maken van onderwijsopbrengsten en hebben elk specifieke voordelen. Kwantitatieve toetsen zijn bijvoorbeeld met name geschikt om vergelijkingen te maken tussen grotere groepen leerlingen/studenten. Ze zijn doorgaans ook goedkoper en praktischer, mede doordat ze een kleiner beroep doen op vakdeskundigheid en het nakijken geautomatiseerd kan worden. Kwalitatieve toetsen zijn vooral geschikt om complexere vaardigheden te toetsen. Spreekvaardigheid kan bijvoorbeeld alleen kwalitatief worden getoetst. Kwalitatieve toetsen nemen meestal meer tijd in beslag, bijvoorbeeld omdat ze individueel moeten worden afgenomen. Ze kunnen echter ook een besparing van toetstijd opleveren, omdat vaardigheden geïntegreerd kunnen worden getoetst. Kwalitatieve methoden resulteren in rijke informatie die kan worden ingezet voor feedback of voor het vaststellen van individuele leerbehoeften en maatwerk in gebruik van lesmateriaal.

2.2 Huidige toetspraktijk is niet evenwichtig

De raad constateert dat er binnen de huidige toetspraktijk sprake is van onevenwichtigheden binnen de verschillende dimensies. De nadruk ligt op de beslissende functie van toetsing en op kwantitatieve toetsvormen. Hieronder zijn de algemene bevindingen per *dimensie* van toetsing en examinering weergegeven. Gedetailleerdere beschrijvingen van de disbalans per *sector* zijn te vinden in het katern op pagina 17.

Formatieve toetsing komt nog onvoldoende aan bod

De raad constateert dat toetsing vooral beslissende functies heeft en dat er minder aandacht is voor de formatieve functies. In het primair onderwijs is veel aandacht voor de beslissende functies van de methodegebonden toetsen, de leerlingvolgtoetsen en de eindtoets. Kan een leerling door naar een volgend leerboekje, naar een volgend niveau, naar de volgende groep, naar het voortgezet onderwijs? Uit veldraadplegingen in het funderend onderwijs blijkt dat scholen de formatieve mogelijkheden van dergelijke toetsen, zoals feedback voor leerlingen en bepaling van individuele onderwijsbehoeften, niet altijd goed benutten. In het hoger onderwijs is er door de beslissende toetsing (per vak een tentamen) minder aandacht voor formatieve toetsing dan gewenst. Beslissende toetsing vraagt veel tijd en aandacht, en dit gaat ten koste van de mogelijkheid om via toetsing feedback te verkrijgen die het leren en het onderwijs kan

¹⁶ Van der Vleuten, Schuwirth, Driessen, Govaerts, & Heenemans, 2015.

¹⁷ Van der Vleuten, & Schuwirth, 2005.

¹⁸ Moonen-van Loon, Overeem, Donkers, Van der Vleuten, & Driessen, 2013; Bouwer, Goossens, Mortier, Lesterhuis, & De Maeyer, 2018.

ondersteunen.¹⁹ Naast feedback krijgen ook andere formatieve functies weinig aandacht. Toetsen worden bijvoorbeeld zelden gekozen omdat ze leerlingen/studenten motiveren of het zelfstandig leren bevorderen. Andersom geldt dat de nadruk op beslissende functies op termijn de intrinsieke motivatie van leerlingen en studenten kan verminderen en kan bijdragen aan een passieve leerstijl.²⁰

De raad constateert echter ook dat er soms juist te weinig aandacht uitgaat naar de beslissende functie van toetsing. In elk van de sectoren blijven grote onderdelen van de onderwijsdoelen ongetoetst of wordt de toetsing beperkt meegenomen in beslissingen over bevordering, selectie/schooladvisering en diplomatoekenning. In het funderend onderwijs geldt dit bijvoorbeeld voor productieve taalvaardigheden: schrijven en mondeling presenteren zijn onderbelicht in de eindtoetsing. Hetzelfde geldt voor brede vorming in alle onderwijssectoren: er vindt verhoudingsgewijs weinig toetsing met een beslissende functie op de onderwijsdoelen van de domeinen socialisatie en persoonsvorming plaats. Doordat deze belangrijke doeldomeinen van het onderwijs weinig aan bod komen in toetsing en leraren en docenten geneigd zijn hun onderwijs aan te passen aan dat wat wordt getoetst²¹, treedt er een onwenselijke versmalling van het onderwijs op.²²

Centrale vormgeving van toetsing in verschillende opzichten uit balans

In de huidige toetspraktijk is er nog geen goede balans tussen centrale en decentrale vormgeving van toetsing. In het voortgezet onderwijs begeleidt Cito de ontwikkeling van centrale examens en gebruiken docenten vaak methodegebonden toetsen om leerlingen tussentijds te volgen. Het schoolexamen wordt wel door scholen zelf vormgegeven, maar lijkt in veel gevallen op het centraal schriftelijk examen. In het hoger onderwijs is de vormgeving van de examinering juist nauwelijks gecentraliseerd, terwijl de studiepaden door individualisering en maatwerk steeds verder uiteenlopen. Dit zorgt ervoor dat de borging van de eindkwalificaties en de vergelijkbaarheid van diploma's onder druk komen te staan. In het middelbaar beroepsonderwijs en het primair onderwijs is er sprake van weer een andersoortige disbalans. In deze sectoren is de toetsing voor een groot deel uitbesteed aan toetsaanbieders. De toetsing is daardoor geheel of verregaand gestandaardiseerd. In sommige opzichten draagt dit bij aan de toetskwaliteit. Voor leraren/docenten heeft deze centrale uitbesteding echter een vervreemdend effect omdat zij weinig zeggenschap hebben over een wezenlijk onderdeel van hun onderwijs.²³

Kwalitatieve toetsmethoden krijgen te weinig aandacht

Ook binnen de dimensie kwalitatieve versus kwantitatieve toetsen neemt de raad een disbalans waar. In het primair en voortgezet onderwijs zijn eindtoetsen en examens doorgaans kwantitatief en gericht op afzonderlijke, goed meetbare kennis en deelvaardigheden.²⁴ Kwalitatieve methoden van toetsen, het genereren van rijke feedback voor leerlingen, en het toetsen van brede onderwijsdoelen komen daarmee zoals eerder opgemerkt in het gedrang. In het middelbaar beroepsonderwijs en het hoger onderwijs is er weliswaar een betere balans tussen kwalitatieve en kwantitatieve methoden, maar ook daar is er niet altijd sprake van evenwichtig gebruik van verschillende vormen van toetsing. Geïntegreerde toetsing van kennis en vaardigheden krijgt een belangrijke plaats in proeven van bekwaamheid, scripties en essay-

¹⁹ Van der Vleuten, Schut, & Heeneman, 2018.

²⁰ Lynch, & Hennessy, 2017.

²¹ Popham, 2001.

²² Onderwijsraad, 2016a.

²³ Abrams, Pedulla, & Madaus, 2003.

²⁴ Er zijn uitzonderingen, bijvoorbeeld het praktijkexamen in het middelbaar beroepsonderwijs en het profielwerkstuk in het voortgezet onderwijs.

toetsen. Niettemin bestaat ook in deze sectoren de neiging om meer kwantitatieve methoden te gebruiken. Een voorbeeld is de tendens om uitvoerige checklists te gebruiken en toetsen consequent te voorzien van deelscores en totaalscores.²⁵ Deze vergaande kwantificering is niet altijd wenselijk, zeker niet wanneer toetsing vooral een formatieve functie heeft en de aandacht van de student gericht moet zijn op feedback op haar of zijn werk.

²⁵ Van der Vleuten, Schut, & Heeneman, 2018.

Voorbeelden van onevenwichtigheden in de toetspraktijk

De toetspraktijk is de afgelopen decennia veranderd. In het primair onderwijs zijn gestandaardiseerde eindtoetsen en volgtoetsen ingevoerd. In het voortgezet onderwijs zijn de eisen van het centraal schriftelijk examen aangescherpt. In het middelbaar onderwijs zijn kwalificatiedossiers ingevoerd en verbonden aan de toetsing en wordt toetsontwikkeling uitbesteed aan gecertificeerde organisaties. In het hoger onderwijs hebben modularisering en de invoering van de bachelor-masterstructuur tot tentaminering per vak geleid. Deze ontwikkelingen hebben om goede redenen plaatsgevonden, maar hebben de toetspraktijken in verschillende opzichten uit evenwicht gebracht.

Primair onderwijs

In het primair onderwijs vormen eindtoetsen en volgtoetsen een belangrijk onderdeel van de onderwijspraktijk. De gestandaardiseerde toetsen maken het mogelijk zinvolle vergelijkingen te maken tussen (groepen) leerlingen, scholen en leerjaren. Op basis van deze vergelijkingen zijn de afgelopen jaren steeds meer leraren, scholen en schoolbesturen opbrengstgericht gaan werken. Hierbij worden aan de hand van toetsresultaten doelen geëvalueerd (en bijgesteld) en fungeren de referentieniveaus die aan de toetsen zijn verbonden als ijkpunten waar betrokkenen naar toe werken.²⁶ Het minimumniveau van basisvaardigheden is door deze werkwijze beter geborgd.²⁷ De focus van de eindtoetsen en volgtoetsen op basale reken- en taalvaardigheden zorgt echter ook voor een versmalling van het onderwijs. Bredere opbrengsten, bijvoorbeeld op het gebied van socialisatie en persoonsvorming, waarvan de laatste jaren het belang wordt beklemtoond, krijgen nog weinig aandacht in de toetsing.²⁸

De nadruk die bij volgtoetsen is komen te liggen op schooladvisering, gaat bovendien ten koste van andere functies die toetsen zouden kunnen vervullen, zoals het vaststellen van specifieke onderwijsbehoeften van leerlingen of het bieden van rijke feedback. Doordat dergelijke functies zijn onderbelicht bij het ontwerp en gebruik van de volgtoetsen, bieden zij weinig aangrijpingspunten voor leerling en docent bij het afstemmen van de leerstrategie en de didactisch aanpak. Dat maakt het lastig om toetsresultaten van volgtoetsen daadwerkelijk in te zetten voor opbrengstgericht werken op het niveau van de individuele leerling of de groep.²⁹

²⁶ Inspectie van het Onderwijs, 2010.

²⁷ Onderwijsraad, 2011b.

²⁸ Onderwijsraad, 2016a.

²⁹ Inspectie van het Onderwijs, 2010, 2015a; Faber, Faber, Visscher, & Schut, 2015; Schildkamp, Heitink, Van der Kleij, Hoogland, Dijkstra, e.a., 2014.

Historische ontwikkelingen in de toetspraktijk van het primair onderwijs

Het gebruik van gestandaardiseerde toetsen in het Nederlandse basisonderwijs gaat terug naar de jaren zestig van de vorige eeuw, toen het Cito (Centraal Instituut voor Toetsontwikkeling) werd opgericht. Het Cito werd verantwoordelijk voor de ontwikkeling van een toets met meerkeuzevragen om leerkrachten te ondersteunen bij het geven van een passend advies over de meest geschikte vervolgopleiding van hun leerlingen (de eindtoets). De toets werd onder meer ingevoerd omdat er zorgen waren over de subjectiviteit van beoordelingen door leerkrachten.³⁰ Behalve de eindtoets hebben ook volgoetsen hun intrede gedaan in het onderwijs. Met deze toetsen kan het vaardigheidsniveau gedurende de basisschoolperiode worden gevolgd. Inmiddels zijn scholen verplicht zowel een eindtoets af te nemen als een leerlingvolgsysteem te gebruiken, maar zijn er verschillende varianten en aanbieders waartussen ze kunnen kiezen.

De eindtoetsen zijn geijkt aan de referentieniveaus taal en rekenen, aan de hand waarvan kan worden gekeken of een leerling aan het vereiste basisniveau voldoet.³¹ Daarnaast zijn ze ook zodanig gestandaardiseerd dat zinvolle vergelijkingen kunnen worden gemaakt tussen leerlingen, tussen scholen en tussen leerjaren. Om praktische redenen en om gestandaardiseerde meting beter mogelijk te maken bevatten de toetsen voornamelijk meerkeuzevragen en richten ze zich overwegend op taal- en rekenvaardigheden. Sommige eindtoetsen en volgoetssystemen bevatten ook open vragen of adresseren onderdelen als wereldoriëntatie en socialemotionele ontwikkeling.

Vanaf het schooljaar 2014-2015 is het afnamemoment van de eindtoets naar achteren geschoven, waardoor de toets op een andere manier wordt meegewogen voor het schooladvies. Hiertegenover staat dat veel scholen de volgoetsen, en dan met name de volgoetsen van groep 7, gebruiken als objectieve basis om het schooladvies te onderbouwen. De Inspectie voert inmiddels het beleid om vanuit haar stimulerende rol 'verder te kijken' dan resultaten op de eindtoets, maar dat neemt niet weg dat deze toetsen een belangrijke rol blijven vervullen in het toezichtkader. Lage gemiddelde scores voor taal en rekenen op de eindtoets zijn aanleiding voor de Inspectie om nader onderzoek te doen naar de onderwijskwaliteit. Scholen zijn eraan gewend geraakt om te focussen op deze kwaliteitsindicatoren.

Voortgezet onderwijs

In het voortgezet onderwijs heeft het centraal schriftelijk examen een belangrijke functie. Met de kernvakkenregeling zijn de eisen van het centraal schriftelijk examen aangescherpt en hebben de kernvakken een groter gewicht gekregen bij de beslissing of een kandidaat slaagt of zakt (zie tekstkader op pagina 19). Daarnaast is de Inspectie scherper toezicht gaan houden op slagingspercentages en cijfers op de centrale examens. De normen hiervoor zijn de afgelopen jaren geleidelijk aangepast.³² Ten slotte is de Inspectie er sinds 2009 onder meer op gaan toezien dat op een school de cijfers van de schoolexamens gemiddeld over een periode van drie jaar niet meer dan een halve punt van de cijfers van de centrale examens verschillen. Onderzoek uit 2013 laat zien dat schoolleiders hier scherper op sturen dan is vereist; ze kijken niet alleen naar het gemiddelde over drie jaar, maar ook naar het verschil per jaar en per vak.³³ De verschillende maatregelen die zijn ingevoerd om de examenkwaliteit te borgen, hebben voor een betere examenkwaliteit gezorgd, maar hebben ook ongewenste neveneffecten. Een daarvan is dat op veel scholen het schoolexamen steeds meer op het centraal schriftelijk examen is gaan lijken; het schoolexamen is gaan dienen als een oefening voor het centraal schriftelijk examen. Dit heeft, zoals de raad al eerder constateerde, een versmalling van de toetspraktijk en het onderwijs tot gevolg.

³⁰ De Groot, 1966.

³¹ Onderwijsraad, 2011b.

³² Ministerie van Onderwijs, Cultuur en Wetenschap, 2017b.

³³ Onderwijs Innovatie Groep, 2013.

De toetsing vindt doorgaans klassikaal, schriftelijk en per vak plaats, waarbij vooral de individuele cognitieve prestatie telt en wordt gewaardeerd met een cijfer. Het veelvuldig afnemen van toetsen 'voor een cijfer' dient niet alleen om beter vast te stellen wat het prestatieniveau is, maar is voor sommige docenten ook een stok achter de deur om leerlingen aan het werk te houden. Omdat de leerlingen ook niet anders verwachten, zijn ze vooral bereid zich in te spannen voor die toetsen die mee tellen. Daarbij hebben leerlingen doorgaans een begrijpelijke voorkeur voor kwantitatieve toetsmethoden, omdat de beoordeling daarvan transparant en objectief is.

Deze toetspraktijk loopt echter uit de pas met innovaties in het onderwijs als flexibilisering van leertrajecten, meer aandacht voor vakoverstijgende vaardigheden en interdisciplinaire onderwerpen, zelfstandig leren, werken in groepjes en de inzet van digitale middelen.³⁴ De huidige integrale curriculumherziening brengt hier mogelijk verandering in, maar er is een risico dat deze curriculumherziening stagneert wanneer de toetsing onvoldoende meeontwikkelt.³⁵ Een punt van zorg van de raad is dat toetsvraagstukken vooralsnog onvoldoende zijn meegenomen in de huidige landelijke curriculumherziening.³⁶ De raad adviseert in zijn *Briefadvies curriculumvernieuwing* om toetsvraagstukken op te nemen in curriculumvernieuwing, met een permanente commissie die dit proces monitort en adviseert over herijking als uit periodieke scans blijkt dat dit noodzakelijk is.³⁷

Historische ontwikkelingen in de toets- en examenpraktijk van het voortgezet onderwijs

Met de invoering van de Mammoetwet in 1968 zijn centrale examens in combinatie met schoolexamens de uiteindelijke cijfers op het examen gaan bepalen.³⁸ De centrale examens fungeren daarbij als objectief ijkpunt, terwijl de schoolexamens ruimte aan scholen bieden om de toetsing van hun eigen onderwijsdoelen te realiseren. De Mammoetwet betekende ook een verkleining van het aantal examenvakken (met wel zwaardere eisen) en gaf de leerlingen de mogelijkheid om vakken te kiezen. De keuzevrijheid is geleidelijk aan iets ingeperkt en het aantal examenvakken is toegenomen, wanneer de 'kleine vakken/onderdelen' zonder centraal schriftelijk examen worden meegerekend.³⁹ Sinds de invoering van de tweede fase in 1999 dienen leerlingen in ieder geval examen te doen voor de kernvakken (waaronder Nederlands, Engels en wiskunde⁴⁰), hun profielvakken en een vak naar keuze (vrij deel).

Voorafgaand aan de aanpassing van de zak-slaagregeling in 2011 was het mogelijk dat een leerling met onvoldoendes op het centrale examen voor alle vakken, dankzij hoge schoolexamencijfers toch slaagde voor het eindexamen. De kwaliteit van schoolexamens was niet altijd optimaal en sommige scholen maakten schoolexamens systematisch gemakkelijker dan het centraal schriftelijk examen.⁴¹

Om deze redenen ging de Inspectie er vanaf 2009 op letten dat het verschil tussen schoolexamen en centraal schriftelijk examen op een school niet meer dan een half punt bedraagt. Daarnaast werd in 2011 de zak-slaagregeling aangepast, waarbij het gemiddeld centraal examencijfer een 5,5 of hoger moet zijn om in aanmerking te komen voor het diploma. Ook werd de zogeheten kernvakkenregeling ingevoerd, die voor havo en vwo inhoudt dat leerlingen voor de vakken Nederlands, Engels en wiskunde ten hoogste één 5 mogen staan. Voor het vmbo geldt eenzelfde regeling voor Nederlands.

Sinds 2016 is de indicator van het verschil tussen schoolexamen en centraal schriftelijk examen geschrapt uit het resultatenmodel van de Inspectie, omdat de zak-slaagregeling en de bepalingen over kernvakken

³⁴ Volman, Raban, Heemskerck, Ledoux, & Kuiper, 2018.

³⁵ Een integrale curriculumherziening in de Canadese provincie Quebec, die in 1996 werd ingezet, is bijvoorbeeld ten dele mislukt omdat de toetsing niet mee ontwikkelde; zie Conseil Supérieur de l'Éducation Quebec, 2014.

³⁶ Ministerie van Onderwijs, Cultuur en Wetenschap, 2018b.

³⁷ Onderwijsraad, 2018a.

de indicator overbodig maakten. Bovendien had de indicator het ongewenste effect dat de school-examens meer op de centrale examens gingen lijken.⁴² De Inspectie blijft verschillen tussen schoolexamen en centraal schriftelijk examen wel monitoren om erop toe zien dat het schoolexamen en centraal schriftelijk examen op een school niet langdurig uiteenlopen.⁴³

Middelbaar beroepsonderwijs

In het middelbaar beroepsonderwijs is de ontwikkeling van gedetailleerde kwalificatie-eisen samengegaan met het afzonderlijk toetsen van deze eisen. Dit heeft geresulteerd in een toename van het aantal toetsen met een beslissende functie. Tegelijk zijn de kwaliteitseisen voor toetsing aangescherpt naar aanleiding van langdurige zorgen over de toetskwaliteit in een voornamelijk decentraal georganiseerde toets- en examenpraktijk.⁴⁴ Inmiddels worden kwalificatie-eisen vaker op geïntegreerde wijze getoetst en wordt de examinering steeds vaker uitbesteed aan externe organisaties.⁴⁵ Wat betreft de kwaliteitseisen en de uitbesteding kenmerkt de toetspraktijk zich door tegengestelde bewegingen van centralisering en decentralisering (zie tekstkader). Een constante is echter de verdere afname van het eigenaarschap van docenten over de inhoud van toetsing (en daarmee het onderwijs).

Drie decennia geleden is het middelbaar beroepsonderwijs begonnen om, op basis van beroepsprofielen, de competenties waaraan studenten moeten voldoen op gedetailleerde wijze systematisch te beschrijven.⁴⁶ De elementen van deze kwalificatiedossiers (kerntaken, werkprocessen, kennis en vakvaardigheden) werden aanvankelijk vaak afzonderlijk getoetst, maar inmiddels worden competenties vaker geïntegreerd getoetst, bijvoorbeeld aan de hand van proeven van bekwaamheid en portfolio's.

Om de objectiviteit en betrouwbaarheid van examenbeslissingen te vergroten, is de afgelopen jaren het kwaliteitsbeleid rondom de toetsing aangescherpt. Hoewel het huidige landelijke toezichtkader geen specifieke toetsmethoden voorschrijft en (buiten de centrale examens) de instelling de ruimte laat om zelf de toetsing vorm te geven, ziet de Inspectie erop toe dat de instelling voldoende maatregelen neemt om de toetskwaliteit te borgen. De examencommissies die hier binnen de instelling verantwoordelijk voor zijn, kunnen in principe een risicoschatting maken van wanneer extra borgende maatregelen nodig zijn. Hoewel dit steeds vaker gebeurt, nemen nog steeds veel van hen het zekere voor het onzekere. Ze schakelen voor de toetsing van praktische vaardigheden in het leerbedrijf bijvoorbeeld standaard verschillende beoordelaars in, terwijl de Inspectie aangeeft dat de inzet van verschillende beoordelaars ook steekproefsgewijs kan, zeker wanneer sprake is van meerdere meetmomenten.⁴⁷

38 Van den Bergh, Rohde, & Zwarts, 2003.

39 College voor Toetsen en Examens, 2018b.

40 Andere verplichte vakken zijn: maatschappijleer, culturele en kunstzinnige vorming, lichamelijke opvoeding en rekenen. Voor deze vakken worden geen centrale examens, maar wel schoolexamens afgenomen (voor rekenen gebeurt dit vanaf schooljaar 2019-2020; van pogingen tot centrale examinering van het rekenonderwijs is inmiddels afgezien). Het gemiddelde cijfer voor deze vakken is opgenomen in de zak- slaagregeling en kan bepaalde resultaten compenseren. Dit geldt ook voor keuzeonderdelen, waaronder het profielwerkstuk. Zie College voor Toetsen en Examens, 2018b.

41 Centrale Examencommissie Vaststelling Opgaven, 2006; Rekers-Mombarg, Harms, & Van der Werf, 2009; Lange & Dronkers, 2007.

42 Canton, Aler, Heemskerk, Van der Westen, & Willemsen, 2013.

43 Wanneer het verschil drie jaren achtereenvolgend groter is dan een halve punt, dient de Inspectie dit aan de Minister melden en komt de examenlicentie van de school in het geding. Zie Inspectie van het Onderwijs, 2018e.

44 Inspectie van het Onderwijs, 2014.

45 MBO Raad, 2015.

46 Karsten (2016) noemt als begin van deze ontwikkeling de nota *Beroepsprofiel en leerplanontwikkeling* van het Ministerie van OCW uit 1986.

47 MBO Raad, 2015; Kennispunt MBO Onderwijs en examinering, 2017.

Mbo-opleidingen besteden steeds vaker de vormgeving van toetsen uit aan gespecialiseerde toetsaanbieders. Deze ontwikkeling hangt samen met het aangescherpte kwaliteitsbeleid, dat meer dan voorheen vereist dat de gehanteerde toetsen bewezen betrouwbaar en valide zijn. De toetsen die instellingen afnemen van gecertificeerde toetsaanbieders, voldoen formeel aan die criteria.⁴⁸ De uitbesteding van toetsen zorgt objectief gezien voor een verlichting van de toetsdruk, omdat docenten niet langer zelf hun toetsen ontwikkelen. Docenten ervaren echter vooral een vermindering van eigenaarschap van hun onderwijs, terwijl ze nog wel belast worden met de afname en administratie van de toetsing. Dat geldt tot op zekere hoogte ook voor de landelijke examens voor Nederlands, rekenen en Engels, die de rijksoverheid in 2010 invoerde. Waar reken- en taalvaardigheden onlosmakelijk verbonden zijn met de verschillende vakken en beroepscontexten van de opleidingen, zijn docenten door de centrale examinering meer gedwongen om zich te richten op examens waar deze vaardigheden geïsoleerd en zonder context worden getoetst. Dit zorgt voor een spanning omdat voor mbo-studenten juist de toepassing van deze vaardigheden in de praktijk relevant en betekenisvol is. Het feit dat voor taal naast centrale examens ook instellingsexamens worden afgenomen, neemt deze spanning enigszins weg, maar evenals in het voortgezet onderwijs (zie vorige paragraaf) geldt dat centrale examens al snel de instellingsexamens domineren. Voor rekenen worden nu in plaats van de centrale examens instellingsexamens voorgesteld (zie Kamerbrief 8 november 2018). Met deze maatregel zou het eigenaarschap van de toetsing weer meer bij het veld komen te liggen.

Historische ontwikkelingen in de toets- en examenpraktijk van het middelbaar beroepsonderwijs

De examinering in het middelbaar beroepsonderwijs kenmerkt zich door afwisselende bewegingen van decentralisering en centralisering. Aanvankelijk vond de examinering zowel centraal door de beroepsgroep als decentraal door de instelling plaats. Dat veranderde met de invoering van WEB (Wet educatie en beroepsonderwijs) in 1996. Er ontstonden toen multisectorale roc's (regionale opleidingscentra) die zelf verantwoordelijk werden voor de examinering. In 2010 heeft de rijksoverheid vervolgens landelijke referentieniveaus vastgesteld en op basis hiervan (evenals in het primair en voortgezet onderwijs) landelijke examens voor Nederlands en rekenen in het middelbaar beroepsonderwijs ingevoerd. In 2012 zijn daar de generieke eisen Engels voor mbo 4 aan toegevoegd, die geleid hebben tot centrale examens en instellingsexamens vanaf studiejaar 2017-2018.

Ook het kwaliteitsbeleid rondom de examinering kenmerkt zich door afwisselende bewegingen van decentralisatie en centralisatie. Tegelijkertijd met de invoering van de WEB werd landelijk sectoroverstijgend kwaliteitsbeleid ingevoerd, dat vereiste dat instellingen hun examens moesten laten certificeren. Dat gebeurde aanvankelijk door commerciële bureaus. Omdat die echter te gemakkelijk hun goedkeuring gaven, kwamen er landelijke kwaliteitseisen voor examinering. Een landelijk overheidsorgaan, het KCE (KwaliteitsCentrum Examinering), moest erop toezien dat instellingen aan deze eisen voldeden. Het voorleggen van grote hoeveelheden toetsproducten en processen aan het KCE zorgde voor een grote beheerslast bij instellingen en bij het KCE zelf. Mede om die reden is het KCE weer opgeheven en is de Inspectie gaan toezien op de examenkwaliteit. Het toezicht van de Inspectie verschilt van het toezicht van het voormalig KCE. Het richt zich vooral op de kwaliteitsbewaking door examencommissies. In 2014 meldde de Inspectie echter dat de examenkwaliteit bij veel instellingen niet op orde was.⁴⁹ Hierop heeft de minister besloten om instellingen te verplichten om hun examens in te kopen bij gecertificeerde leveranciers of hun examens te laten valideren. Veel instellingen kiezen voor de eerste optie, omdat dit meer garantie zou bieden dat de Inspectie hun toetsbeleid goedkeurt.⁵⁰

48 MBO Raad, 2018.

49 Inspectie van het Onderwijs, 2014.

50 Karsten, 2016.

Hoger onderwijs

Het hoger onderwijs kent een traditie van decentrale vormgeving in relatief losstaande onderwijseenheden, ook wel vakken of modules genoemd. Hoewel opleidingen ook samenhang aanbrengen tussen deze eenheden, is in vergelijking met andere sectoren het onderwijs in zowel het hoger beroepsonderwijs als het wetenschappelijk onderwijs meer modulair opgebouwd. Het modulaire karakter is de afgelopen decennia nog versterkt met maatregelen die meer flexibele studieprogramma's faciliteren. Hierdoor is een grotere nadruk komen te liggen op beslissende toetsen. Een andere ontwikkeling is dat er, mede als gevolg van de toegenomen studentenaantallen, meer gebruik wordt gemaakt van kwantitatieve toetsmethoden.

Door modularisering, individualisering en maatwerk in het hoger onderwijs is het voor studenten beter mogelijk geworden om zelf een onderwijsprogramma samen te stellen en ook onderdelen bij andere instellingen te volgen, in Nederland of elders in Europa (zie tekstkader). Er is hierdoor echter ook een grotere nadruk komen te liggen op beslissende toetsing. De invoering van de bachelor-masterstructuur heeft tot een verdubbeling van stages en scripties geleid. Voorheen hoefde alleen de doctoraalfase met een stage en een scriptie te worden afgerond, nu geldt dat vaak voor zowel de bachelor- als de masteropleiding. Ook voor de reguliere onderdelen die deel uitmaken van een opleiding is de zak-slaagbeslissing steeds belangrijker geworden, omdat geborgd moet zijn dat de studenten, die steeds vaker uiteenlopende programma's volgen, de gestelde leerdoelen binnen een module hebben behaald.

In reguliere vakken, die doorgaans een omvang hebben van vijf à tien weken, worden minimaal één en vaak meer beslissende toetsen afgenomen, omdat zowel theoretische als toepassingsgerichte elementen van een vak aan bod moeten komen. Daarnaast vindt soms ook nog formatieve toetsing plaats, om studenten voor te bereiden op een beslissende toets. Door de vele toetsen ervaren sommige studenten hun studie als een hordeloop en dit brengt soms stress met zich mee.⁵¹

Ook docenten spreken over toetsdruk, omdat zij de toetsen zelf ontwerpen en beoordelen, aangezien ze op die manier specialistische en actuele kennis kunnen adresseren. Daarbij dienen ze rekening te houden met de aan toetsing gestelde kwaliteitseisen. De afgelopen jaren zijn deze eisen aangescherpt, mede door het kwaliteitsbeleid dat instellingen voeren. Aanleiding voor de aanscherping waren incidenten waarbij aan het licht kwam dat de kwaliteit van de eindwerken niet goed was geborgd.⁵² Daarnaast houden docenten ook rekening met de studenten zelf, die direct of via studentevaluaties, opleidingscommissies of de examencommissie hun mening kunnen geven over de inhoud en kwaliteit van toetsing. Deze verschillende partijen dringen er bij docenten op aan dat zij hun beoordelingen op transparante wijze verantwoorden. Mede hierdoor kiezen docenten sneller voor kwantitatieve toetsmethoden. Een andere reden hiervoor is dat deze methoden efficiënter zijn; dit is van belang gezien de gestegen aantallen studenten in het hoger onderwijs.

Modularisering van opleidingen zet de samenhang van opleidingen onder druk. Tot op zekere hoogte zorgen afsluitende onderdelen als eindwerkstukken, stages en scripties voor samenhang in het toetsprogramma. Deze onderdelen toetsen geïntegreerd de eindtermen. Verantwoordelijke docenten en examencommissies worstelen echter met het toewerken naar deze afsluitende onderdelen. Dit komt doordat de modulaire opzet van een programma en een gevarieerde studentenpopulatie als gevolg van de zijinstroom een goede opbouw van leer-

⁵¹ Rouwhorst, Sinke, Van Kessel, & Zondervan, 2017.

⁵² Commissie Externe Validering Examenkwaliteit Hoger Beroepsonderwijs, 2012.

doelen en toetsing soms in de weg staan. Verder geldt dat de afsluitende onderdelen weliswaar veel, maar zeker niet alle eindtermen van de opleiding adresseren. Het blijft daardoor de vraag of studenten door het accent op kwantitatieve toetsing de vaardigheden en kennis, die in eerdere vakken zijn 'afgetoetst', voldoende hebben geconsolideerd. Bovendien heeft ook hier de kwantitatieve manier van toetsing een versmalling van de curriculuminhoud tot gevolg.

Historische ontwikkelingen in de toets- en examenpraktijk van het hoger onderwijs

In het hoger onderwijs heeft de toetsing belangrijke functies voor het bevorderen en diplomeren van studenten. In de modulaire opbouw die kenmerkend is voor zowel het hoger beroepsonderwijs als het wetenschappelijk onderwijs, worden vakken afgesloten met tentamens. Wanneer studenten een tentamen halen, krijgen ze studiepunten, waarmee zij uiteindelijk een diploma behalen. De afgelopen decennia is in Europees verband gewerkt aan het vergroten van de flexibiliteit van studieprogramma's en het vereenvoudigen van de uitwisseling van studenten. Hiertoe is de bachelor-masterstructuur ingevoerd en hebben studiepunten een vaste omvang gekregen (volgens het European Credit Transfer System). Instellingen zelf zijn daarnaast ook overgegaan tot vaste blokindelingen.⁵³ De meeste opleidingen hanteren een jaar- of semesterindeling van blokken met bijvoorbeeld zes keer twee parallelle blokken van vijf studiepunten, waarbij doorgaans in de laatste week een of meer afsluitende toetsen worden afgenomen.⁵⁴ Van de blokstructuur kan worden afgeweken; aan studieonderdelen als stage of scriptie, die een plaats hebben aan het eind van de opleiding, worden grotere blokken (meer studiepunten) toegekend. Soms worden aan onderdelen van de scriptie, zoals het schrijven van een plan, deelbeoordelingen toegekend.

Toetsing speelt in het hoger onderwijs eveneens een belangrijke rol in de beoordeling van de kwaliteit van het onderwijs. Hoewel er geen door de overheid gecoördineerde examens worden afgenomen, borgt de overheid wel de kwaliteit van de toetsing en examinering via de Inspectie en de NVAO (Nederlands-Vlaamse Accreditatieorganisatie). De NVAO verleent accreditatie aan een opleiding op basis van een visitatie. Zusteropleidingen zitten in een cluster en worden in dezelfde periode gevisiteerd.⁵⁵ Toetsing en examinering spelen een belangrijke rol bij deze visitaties. Nagegaan wordt of de opleiding beschikt over een adequaat systeem van toetsing (standaard 3) en of de gerealiseerde onderwijsopbrengsten die aan de hand van dit systeem worden aangetoond voldoen (standaard 4). Standaard 3 wordt voornamelijk beoordeeld op basis van de kwaliteitszorg rondom examens van de opleiding, in het bijzonder het functioneren van de examencommissie. Standaard 4 wordt steekproefsgewijs beoordeeld aan de hand van afsluitende toetsen, waaronder eindwerkstukken, scripties en portfolio's en soms ook toetsen in eerdere fasen van de opleiding.⁵⁶

Om de kwaliteit van de toetsing beter te borgen, hebben examencommissies in het hoger onderwijs met de invoering van de Wet versterking besturing in 2010 een grotere rol gekregen. De wet regelt dat examencommissies meer dan voorheen verantwoordelijkheden en bevoegdheden hebben ten

53 Sursock, Smidt, & Davies, 2010.

54 Behalve door het idee van flexibele modules is deze toetsing per onderdeel ook ingegeven door beleidsdoelstellingen om studenten regelmatig te laten studeren en de studielast beter te spreiden in vergelijking met bijvoorbeeld een aanpak waarbij examinering op het eind van de studiejaar plaatsvindt (Inspectie van het Onderwijs, 2009).

55 De NVAO verleent ook accreditaties aan instellingen, maar toetsing van leerresultaten komt daarbij meer zijdelings aan de orde (bij standaard 4: evaluatie en monitoring)

56 Nederlands-Vlaamse Accreditatieorganisatie, 2016.

aanzien van het borgen van de kwaliteit van de toetsing en gerealiseerde opbrengsten van afgestudeerden. Naast deze maatregelen zijn er ook andere maatregelen genomen.⁵⁷ Het gaat dan bijvoorbeeld om een andere rol van opleidingscommissies, externe validering van examens, scholing en certificering van docenten op het gebied van toetsing, en in enkele gevallen ook instellingsoverstijgende toetsing.⁵⁸

⁵⁷ Een belangrijke aanleiding hiervoor vormden een aantal incidenten medio 2010 waarbij kwaliteitsgebreken op het gebied van toetsing in het hoger beroepsonderwijs aan het licht kwamen. Zie Commissie Externe Validering Examenkwaliteit Hoger Beroepsonderwijs, 2012.

⁵⁸ Inspectie van het Onderwijs, 2016a.

Een evenwichtige toetspraktijk vraagt zowel van de rijksoverheid als van onderwijsinstellingen een heldere visie op toetsing en examinering. Vanuit deze visie en daaruit voortvloeiende verantwoordelijkheden kunnen zowel overheid als instellingen duidelijke kaders bieden, zodat alle betrokkenen in het veld de bestaande ruimte voor toetsing en examinering beter kunnen benutten.

3 Overheid en instellingen: baseer de toetspraktijk op een heldere visie

Toewerken naar een evenwichtiger toetspraktijk vraagt om een heldere visie. Zowel de rijksoverheid als onderwijsinstellingen dienen een dergelijke visie te formuleren als vertrekpunt voor het vormgeven van het toetsbeleid. De raad adviseert de rijksoverheid om een consistent toetsbeleid te voeren en duidelijke kaders te stellen. De instellingen beveelt hij aan om binnen de gegeven kaders een eigen visie op toetsing te formuleren en daarmee de ruimte te benutten die beschikbaar is. In beide gevallen is rolvastheid voor alle betrokken partijen noodzakelijk. Rolverwarring draagt niet bij aan de kwaliteit van de toets- en examenpraktijk.

Een visie formuleren betekent stellingname. Overheid en instellingen dienen zich in de ogen van de raad, vanuit hun verschillende rollen, uit te spreken over:

- 1 de mate waarin toetsing beslissende en formatieve functies heeft;
- 2 de mate waarin de besluitvorming rond toetsing centraal of decentraal plaatsvindt; en
- 3 de mate waarin onderwijsopbrengsten op kwantitatieve wijze of meer kwalitatieve wijze zichtbaar worden gemaakt.

Zoals aangegeven in hoofdstuk 2 dienen overheid en instellingen te zoeken naar een balans binnen deze dimensies, waarbij uiteenlopende doelen met elkaar worden verenigd en toetsing ten dienste staat van de onderwijskwaliteit, de leerloopbaan van de leerling en student en de te verwerven positie op de arbeidsmarkt en in de maatschappij.

3.1 Overheid: bied duidelijke kaders aan de onderwijsinstellingen

Op welke onderdelen legt de rijksoverheid de toetsing vast? Hoeveel ruimte geeft zij aan de instellingen? Dat verschilt per sector, maar alle onderwijsinstellingen hebben belang bij duidelijke kaders die de beschikbare ruimte afbakenen. Deze kaders worden om te beginnen gevormd door rolvastheid en consistentie in het landelijke toetsbeleid. Voor een evenwichtige toetspraktijk wijst de raad daarnaast op de voordelen van standaardisering van de eindtoetsing, die – in bepaalde mate – voor alle sectoren kan worden overwogen. De kwaliteitszorg rondom de toets- en examenpraktijk in de onderwijsinstelling wordt in de ogen van de raad versterkt door toets- en examencommissies te verplichten voor primair en voortgezet onderwijs. Ook het toezicht op de toetspraktijk in de verschillende sectoren vraagt volgens de raad om aanscherping.

Bewaak de consistentie in het toetsbeleid en wees rolvast

De raad acht het van groot belang dat de rijksoverheid standvastig is in haar toetsbeleid en dat zij niet buiten haar rol treedt. Standvastig beleid biedt onderwijsinstellingen duidelijkheid bij het vormgeven van hun eigen visie op toetsing. Het onderwijsveld moet erop kunnen vertrouwen dat de standpunten van de rijksoverheid ten aanzien van toetsing niet telkens wijzigen. Soms worden in rap tempo maatregelen ingevoerd, die achteraf onvoldoende doordacht blijken te zijn. Dan worden maatregelen weer teruggedraaid. Of de rijksoverheid begeeft zich op het terrein van de onderwijsinstelling met bepaalde maatregelen en is daarmee niet rolvast. Dit scheidt verwarring waar juist behoefte is aan helderheid.

Een voorbeeld van zo'n maatregel is het in de tijd naar achteren verschuiven van het afname-moment van de eindtoets (zie pagina 20). Deze maatregel is meteen breed ingevoerd op alle basisscholen, zonder dat vooraf goed in kaart was gebracht wat invoering zou betekenen voor leerlingen, ouders en scholen, en welke consequenties dit zou hebben voor toegankelijkheid en gelijke kansen. Door de eindtoets minder gewicht toe te kennen, lijkt het advies dat leerlingen krijgen afhankelijker te worden van hun achtergrond.⁵⁹ Dit doet afbreuk aan het principe van gelijke kansen voor iedereen. Inmiddels gaan weer stemmen op voor het terugdraaien van de maatregel.⁶⁰

Het komt vaker voor dat besluiten met betrekking tot toetsing snel weer worden teruggedraaid. Zo werd een streep gezet door de wettelijke plicht van de diagnostische tussentijdse toets in het voortgezet onderwijs.⁶¹ Begin 2018 deelde de minister nog mee dat de rekentoets toch niet meetelt voor het vwo-eindexamen⁶² en inmiddels ligt er een voorstel om rekenen weer wel mee te laten tellen, maar dan via het schoolexamen⁶³ (zie onderstaand kader). En recent werd het voornemen gedeeld om het aantal studiepunten dat behaald moet worden voor het bindend studieadvies in het hoger onderwijs te verlagen.⁶⁴

De rekentoets in het voortgezet onderwijs

De rekenprestaties van alle leerlingen verbeteren was een van de doelen bij het invoeren van de referentieniveaus.⁶⁵ De Tweede Kamer besloot destijds om aan het eind van een opleiding te toetsen of leerlingen het vastgestelde referentieniveau hebben gehaald. Aan het eind van het voortgezet onderwijs werd een rekentoets als verplicht onderdeel van het examen ingevoerd voor alle leerlingen van alle schoolsoorten. De rekentoetsen in het voortgezet onderwijs werden sinds het schooljaar 2013-2014 verplicht, maar telden op dat moment nog niet mee in de zak-slaagbeslissing. Wel werd het cijfer voor de rekentoets vermeld op (een bijlage bij) de cijferlijst.

De rekentoets is in enkele opzichten een novum in het voortgezet onderwijs. Exact dezelfde toets en cesuur gelden voor meerdere opleidingen (2F voor vmbo en 3F voor havo en vwo) en hoewel de leerinhoud wordt afgesloten met een centrale toets, betreft het geen zelfstandig vak. Bovendien worden de rekentoetsen geheel digitaal afgenomen en automatisch gescoord en genormeerd aan de hand van een vaardigheidsschaal.⁶⁶

De invoering van de rekentoets voortgezet onderwijs was van meet af aan onderwerp van discussie. Meningsverschillen waren er onder meer over de methode, over de te toetsen leerinhoud (wel/geen

⁵⁹ Dronkers, Efficers, Karsten, & Van der Velden, 2016; Feron, Schils, & Ter Weel, 2016.

⁶⁰ Huygen, 2016.

⁶¹ Ministerie van Onderwijs, Cultuur en Wetenschap, 2018a.

⁶² Ministerie van Onderwijs, Cultuur en Wetenschap, 2017a.

⁶³ Ministerie van Onderwijs, Cultuur en Wetenschap, 2018c.

⁶⁴ Rijksoverheid, 2018.

⁶⁵ Expertgroep Doorlopende Leerlijnen Taal en Rekenen, 2008.

⁶⁶ College voor Toetsen en Examens, 2018a.

rekenmachine, kale sommen of contextopgaven), over de positie in de zak-slaagregeling, over de voorbereiding door scholen, over de toets zelf (het aantal opgaven, het digitale karakter) en over de snelheid van invoering. Ook vielen de resultaten op de rekentoetsen tegen en bleek 2F niet voor alle leerlingen haalbaar. Dit resulteerde in verschillende debatten, algemene overleggen en rondetafelgesprekken in de Tweede Kamer, het instellen van een tweetal adviescommissies⁶⁷, en het uitvoeren van een aantal onderzoeken en intensiveringstrajecten.⁶⁸ Als antwoord hierop zijn de toetsen en examens op verschillende punten aangepast en zijn er varianten gekomen voor specifieke doelgroepen.

Het laten mee tellen van het cijfer op de rekentoets vo was voorzien vanaf schooljaar 2015-2016, maar is verschillende keren uitgesteld. Uiteindelijk heeft de rekentoets vo alleen voor het vwo één jaar meegeteld voor het diploma. In het regeerakkoord van het kabinet-Rutte III is aangekondigd dat er een alternatief voor de rekentoets komt. Dit alternatief treedt uiterlijk in het schooljaar 2019-2020 in werking en wordt daarmee voor alle leerlingen op alle niveaus een geïntegreerd onderdeel van het examen. Inmiddels ligt er een plan om rekenen mee te laten tellen voor het behalen van een diploma in het vmbo, havo, vwo en mbo, via het school- of instellingsexamen.⁶⁹

Ook gebeurt het dat de rijksoverheid, naar aanleiding van een maatschappelijk of politiek debat over toetsing en examinering, regels vastlegt die de ruimte voor de school en de leraar onnodig inperken. Een voorbeeld daarvan is het recente verbod op het gebruik van kleuter-toetsen in het leerlingvolgsysteem.⁷⁰ De regering kiest hiermee positie in een langlopende discussie over voor- en nadelen van toetsen voor jonge kinderen.⁷¹ De raad vindt dat het niet aan de rijksoverheid is om voor scholen te bepalen of zij wel of niet gebruikmaken van kleutertoetsen. Er zijn leraren die het toetsen van kleuters onverstandig vinden en liever een observatie-instrument gebruiken. Maar er zijn ook leraren die juist vinden dat het toetsen van kleuters zicht biedt op de ontwikkeling en handvatten geeft voor hun onderwijspraktijk.

Standaardiseer eindtoetsing waar nodig en mogelijk

De raad wijst op de voordelen van gestandaardiseerde eindtoetsing bij belangrijke overgangen in het stelsel. Niet alleen bij de overgang van het primair naar het voortgezet onderwijs en van het voortgezet onderwijs naar een vervolgopleiding, maar ook bij andere overgangen zijn er voordelen van standaardisering van de eindtoetsing. Gestandaardiseerde eindtoetsing biedt een duidelijk kader voor alle betrokkenen. Vanuit dat kader ontstaat er ruimte voor instellingen om hun eigen onderwijsdoelen en de toetsing daarvan vorm te geven. Voor het funderend onderwijs en voor het middelbaar beroepsonderwijs (de centrale examens) geldt bovendien dat gestandaardiseerde eindtoetsing garandeert dat de ijkpunten bewaakt worden. Daarmee ontstaat vrijheid voor het vormgeven en uitvoeren van de eigen visie op onderwijs en toetsing van de instelling.

In het **primair onderwijs** maken alle leerlingen een eindtoets. Doordat er tegenwoordig meerdere aanbieders zijn in de markt van eindtoetsen, hebben basisscholen meer keuzevrijheid dan voorheen.⁷² Deze keuzevrijheid heeft een prijs. Het is lastig de verschillende eindtoetsen zo te normeren dat ze onderling vergelijkbaar zijn.⁷³ De Inspectie waarschuwde in 2017 voor de afnemende vergelijkbaarheid van de eindtoets als gevolg van de toelating van nog drie nieuwe ver-

67 Commissie Verantwoorde Invoering Rekentoets, 2014; Bosker, & Van de Vorle, 2014.

68 Zie College voor Toetsen en Examens, 2015; Ministerie van Onderwijs, Cultuur en Wetenschap, 2015.

69 Ministerie van Onderwijs, Cultuur en Wetenschap, 2018c.

70 Ministerie van Onderwijs, Cultuur en Wetenschap, 2018d.

71 *Voor- en tegenstanders over de Cito kleutertoets*, 2018.

72 Voor 2019 kunnen scholen kiezen uit de volgende toetsen: centrale eindtoets po van het College voor Toetsen en Examinering; eindtoets ROUTE 8 van A-VISION; IEP-eindtoets van Bureau ICE; DIA-eindtoets van Diataal; AMN-eindtoets van AMN.

73 Glas, Emons, & Berding-Oldersma, 2016.

sies op de markt.⁷⁴ In 2018 bleek dat zwakkere scholen gebruikmaken van andere eindtoetsen dan betere scholen.⁷⁵ Hiermee staat een van de ijkpunten in het Nederlandse onderwijsstelsel onder druk. Eén eindtoets zou dit ondervangen, maar doet geen recht aan de verschillende visies en voorkeuren met betrekking tot toetsing onder basisscholen. Het belang van de eindtoets als ijkpunt voor het onderwijsstelsel is dusdanig groot, dat de raad vindt dat de toetsen van onbetwistbare kwaliteit moeten zijn. Dat houdt in dat verschillen in de normering van de verschillende eindtoetsen niet acceptabel zijn en dat alles op alles gezet moet worden om deze toetsen door de overheid (College voor Toetsen en Examens) te kalibreren en stelselmatig te evalueren.

De raad heeft diverse malen het belang van gestandaardiseerde eindtoetsen als ijkpunten in het **voortgezet onderwijs** onderstreept.⁷⁶ Om daadwerkelijk als ijkpunten te kunnen fungeren dienen centrale examens van hoge kwaliteit te zijn en dienen alle betrokkenen, waaronder leerlingen, ouders, scholen en vervolgopleidingen, hiervan overtuigd te zijn. Uit onderzoek blijkt dat de centrale schriftelijke eindexamens in ons land van goede kwaliteit zijn en dat ze een bijdrage leveren aan goed onderwijs.⁷⁷ Dat neemt niet weg dat het volgens de raad belangrijk is om scherp te blijven en om incidenten met betrekking tot de inhoud en afname van de centraleschriftelijke eindexamens te voorkomen. Waar zich incidenten voordoen is het van belang er open over te zijn, ervan te leren en ze in perspectief te plaatsen. Voorkomen moet worden dat het vertrouwen in het gehele centrale examen ondermijnd wordt door specifieke incidenten in een bepaalde context.

Om de kwaliteit en het vertrouwen te borgen zijn zorgvuldige en transparante procedures nodig, waarbij er voldoende deskundigheid en 'checks and balances' zijn. De raad constateert met tevredenheid dat de procedures van het College voor Toetsen en Examens op dit punt zijn verbeterd. Zo hebben leraren en vakverenigingen meer mogelijkheden gekregen om mee te denken over de correctievoorschriften.⁷⁸ De raad moedigt het College voor Toetsen en Examens aan zijn nieuwe beleid voort te zetten en verder aan te scherpen. Aan het bredere onderwijsveld doet de raad enerzijds de aanbeveling om gebruik te maken van de mogelijkheden om mee te denken over het centrale examen en anderzijds om realistische verwachtingen te koesteren. Het centrale examen is op zijn best de grootste gemene deler van hoe in het voortgezet onderwijs de eindtoetsing wordt gewenst. Het kan dus per definitie niet tegemoetkomen aan alle persoonlijke wensen en kritiek. Tegelijkertijd is het centraal examen een noodzakelijke tegenhanger van de vrijheid van onderwijs die ons stelsel kenmerkt. De raad wil er ten slotte op wijzen dat een deel van die vrijheid in het voortgezet onderwijs ook ligt bij het schoolexamen en dat die vrijheid beter benut en ingevuld zou kunnen worden dan nu het geval is.

Voor het **middelbaar beroepsonderwijs** adviseert de raad om pas op de plaats te maken met verdergaande wijzigingen in toetsing. De sector is in de afgelopen jaren blootgesteld aan vele wijzigingen in toetsing en heeft tijd nodig om deze wijzigingen te kunnen borgen. Met betrekking tot standaardisering van de eindtoetsing vindt de raad de huidige aanzet voor sectorale examenprofielen voor vergelijkbare kwalificatiedossiers een vorm van standaardisering die het civiel effect van mbo-diploma's ten goede komt. Van belang is wel dat de beroepssec-

74 Inspectie van het Onderwijs, 2017.

75 Inspectie van het Onderwijs, 2018a.

76 Onderwijsraad, 2014, 2015a, 2017b.

77 Van de Werfhorst, Elffers & Karsten, 2015.

78 Het College voor Toetsen en Examens probeert met *Project Ieders Examen* het draagvlak voor examens te vergroten. Zij betreft vakverenigingen meer bij het vaststellen van correctievoorschriften om zo meer consensus over de voorschriften te ontwikkelen. Ook geeft zij, bijvoorbeeld via webinars meer openheid over de procedures voor de ontwikkeling van examens en scoringsvoorschriften.

toren deze examenprofielen nog concreter invulling geven, zodat ze daadwerkelijk handvaten voor examinering bieden en in de onderwijspraktijk 'gaan leven'. De raad is minder gelukkig met de deelcertificering die nu voor sommige examenonderdelen mogelijk is (het gaat daarbij om onderdelen waarvan is aangetoond dat ze afzonderlijke betekenis hebben voor de arbeidsmarkt). Deze certificering is weliswaar begrijpelijk in het licht van gepersonaliseerd onderwijs, maar leidt tot versnippering van kwalificatie-eisen en doet daarmee afbreuk aan het civiel effect van diploma's.⁷⁹ Daarnaast ontnemt de losse certificering studenten een lange-termijnperspectief op hun opleiding, waarbij ze toewerken naar complete beheersing van de eindtermen aan het eind van de opleiding.

De raad plaatst ook vraagtekens bij het op voorhand valide verklaren van examenonderdelen van examenleveranciers die over een keurmerk beschikken. De raad constateert dat op dit punt de examencommissie, die formeel opdrachtgever is aan de toetsaanbieders, al te snel wordt gepasseerd. De examencommissie moet immers een oordeel vellen over de validiteit van de examens en zo nodig aanpassingen in de examinering kunnen afdwingen wanneer daarmee een betere aansluiting ontstaat bij eigen toetsvisie en de plaats van de examens in de opleiding. Op die manier behoudt de opleiding het eigenaarschap over een belangrijk deel van het onderwijs. Daarnaast zou uitbesteding niet de voorkeursoptie moeten zijn (de richtlijn dat uitbesteding de voorkeur heeft, wordt thans afgegeven door de MBO Raad en de Nederlandse Raad voor Training en Opleiding). Opleidingen die de examinering zelf of samen met de externe toetsaanbieders vorm wensen te geven, moeten worden aangemoedigd. Aanmoediging betekent ook dat de hiervoor benodigde deskundigheid bij docenten wordt vergroot (zie hoofdstuk 4).

In het **hoger onderwijs** vindt de eindtoetsing voornamelijk decentraal plaats. De opleiding bepaalt binnen de kaders van de Dublin-descriptoren wat de algemene eisen zijn voor succesvolle afronding van een bachelor- of masteropleiding.⁸⁰ De docenten die het desbetreffende vak verzorgen, zorgen doorgaans voor de vormgeving van de examenonderdelen. Deze decentrale regie op de eindtoetsing heeft als voordeel dat de toetsing eerder aansluit bij de visie van de opleiding en de docenten. Een nadeel kan echter zijn dat de eisen aan studenten van dezelfde opleiding bij verschillende universiteiten uiteen gaan lopen. Opleidingen maken weliswaar afspraken om beroepsprofielen of kwalificatie-eisen af te stemmen (dit gebeurt bijvoorbeeld in het kader van visitaties of akkoorden met de beroepssector in het hoger beroeps-onderwijs), maar deze standaarden hebben een algemeen karakter. Opleidingen en docenten geven verschillende interpretaties aan deze standaarden waar het gaat om examinering. Het resultaat is dat er verschillen ontstaan tussen opleidingen in de feitelijke eisen die aan de student worden gesteld.

Omdat de raad hecht aan een duidelijk eigen profiel van hogeronderwijsopleidingen, juicht hij accentverschillen in profilering tussen zusteropleidingen toe.⁸¹ Wanneer er echter grote niveaueverschillen tussen zusteropleidingen ontstaan, ziet hij wel een probleem. Een diploma van dezelfde studie aan verschillende universiteiten of hogescholen heeft dan op papier wel, maar in de praktijk niet dezelfde waarde en dat betekent dat het civiel effect van het diploma onder druk komt te staan.⁸² De raad constateert bovendien dat het risico op dergelijke niveaueverschillen groter is geworden doordat het al moeilijk is om binnen een opleiding het niveau

79 Onderwijsraad, 2015a.

80 Dublin-descriptoren zijn de Europese eindtermen voor het hoger onderwijs. Zie Bologna Working Group, 2005.

81 Onderwijsraad, 2011b.

82 Onderwijsraad, 2010a.

te borgen, als gevolg van modularisering, de stijgende diversiteit van studentenpopulaties en de flexibilisering van onderwijsprogramma's. Dergelijke ontwikkelingen maken het noodzakelijk om na te denken over meer centralisering van toetsing. De raad denkt hierbij niet aan door de landelijke overheid gecoördineerde toetsing, omdat dat niet past bij de aard van de sector en bij de academische vrijheid.⁸³ Hij denkt aan verdergaande samenwerking tussen zusteropleidingen. Behalve de gezamenlijke formulering van kwalificatie-eisen is het ook zinvol om na te gaan of zusteropleidingen, net zoals in het middelbaar beroepsonderwijs, per sector of discipline gezamenlijke examenprofielen kunnen maken of, zoals bij de universitaire medische opleidingen, op onderdelen hun toetsing gezamenlijk kunnen vormgeven. Ook is zo'n ontwikkeling aan te bevelen om internationale mobiliteit te bevorderen.⁸⁴

Geef ook het toezicht op toetsing en examinering vorm als duidelijk kader voor het veld

Naast consistent beleid en, waar nodig, standaardisering van de eindtoetsing, vormt het toezicht een duidelijk kader voor onderwijsinstellingen met betrekking tot toetsing en examinering. Onderwijsinstellingen weten vanuit de toezichtkaders waar ze aan moeten voldoen en waar ze hun eigen ruimte kunnen gebruiken. De raad is van mening dat de positie van toetsing in het toezicht verbeterd kan worden. Hij licht dit hieronder per sector toe.

Voor het **primair onderwijs** kan het toezicht op de toetsing buiten de eindtoets om zich beperken tot de deugdelijkheid van toetsprocessen. Scholen moeten inzichtelijk kunnen maken hoe zij hun kwaliteitszorg rondom toetsing hebben ingericht, met aandacht voor bijvoorbeeld de deskundigheid van betrokkenen, de mate waarin de gekozen instrumenten de onderwijsdoelen dienen, en de samenhang tussen toetsen. Het vernieuwde toezicht is al meer dan voorheen gericht op processen, maar heeft volgens de raad nog onvoldoende aandacht voor toetsing als apart onderwerp. Het ontwerp van het nieuwe onderwijsresultatenmodel bevat voorstellen voor breder kijken naar resultaten.⁸⁵ De raad is hier voorstander van, maar vindt nadrukkelijk dat gewaakt moet worden voor het aanvullen van verantwoordingskaders met een set met 'af te vinken' aspecten van brede onderwijskwaliteit. Dit brengt namelijk het risico met zich mee dat scholen zich hierdoor laten sturen, met als gevolg dat zij de aandacht gaan richten op zaken die daarvoor in kaart moeten worden gebracht in plaats van de eigen visie te ontwikkelen en in de praktijk te brengen. De raad heeft in zijn advies *De volle breedte van onderwijskwaliteit* suggesties gedaan voor hoe dit wel vorm zou kunnen krijgen.⁸⁶ Het valt volgens de raad te overwegen om het schooleigen gedeelte van de toetsing meer gewicht te geven. Dit kan dan ook een duidelijkere rol spelen in de advisering voor het vervolgonderwijs en zou dan tevens aandacht moeten krijgen in het toezicht. De raad is daarmee nog geen voorstander van een eindexamen primair onderwijs. Het gaat hem om het verder vergroten van de aandacht voor wat scholen doen buiten de onderdelen die getoetst worden in de eindtoets.

De Inspectie gebruikt eindtoetscijfers in het primair onderwijs ook in het toezicht op de kwaliteit van de school. Voor toeleiding naar het vervolgonderwijs zijn de vaardigheden die in de eindtoets aan bod komen redelijk goede voorspellers van succes⁸⁷, maar als peilstok voor de Nederlandse onderwijskwaliteit vormen ze een te smalle basis. Instrumenten zoals de peilingsonderzoeken⁸⁸ bieden een bredere kijk op de onderwijskwaliteit dan taal- en rekenvaardighe-

⁸³ Commissie Externe Validering Examenkwaliteit Hoger Beroepsonderwijs, 2012.

⁸⁴ Verhoeven, Snellen-Balendong, Hay, Boon, Van der Linde, e.a., 2005. Zie ook Organisation for Economic Co-operation and Development, 2015.

⁸⁵ Inspectie van het Onderwijs, 2018b.

⁸⁶ Onderwijsraad, 2016a.

⁸⁷ De Boer, Bosker, & Van der Werf, 2010; Timmermans, Kuyper, & Van der Werff, 2013; Lenhard & Schröppel, 2014.

⁸⁸ Inspectie van het Onderwijs, 2018d.

den in de eindtoets. Eindtoetscijfers zijn niet een op een te relateren aan de kwaliteit van een school, ook niet na controle voor achtergrondkenmerken of een toets in groep 3.⁸⁹ Scholen die soepeler zijn in het accepteren van nieuwkomers of scholen die minder snel verwijzen naar het speciaal onderwijs, zullen minder goed scoren dan verwacht. Scholen die daarentegen veel voor de eindtoets oefenen, zullen beter scoren. De raad heeft dan ook eerder aangegeven dat scholen niet generiek dienen te worden 'afgerekend' op eindtoetsresultaten.⁹⁰ De Inspectie baseert zich in het toezicht tegenwoordig op een veel breder palet aan indicatoren dan eindtoetsen alleen. De raad vindt dat een belangrijk gegeven. Zodra scholen namelijk worden afgerekend op hun (gemiddelde) toetsprestaties, kunnen perverse prikkels ontstaan die de kwaliteit van het onderwijs onder druk zetten.

Voor het **voortgezet onderwijs** acht de raad het wenselijk dat er in de praktijk duidelijker onderscheid wordt aangebracht tussen het schoolexamen en het centraal schriftelijk examen. Het schoolexamen zou een zelfstandig onderdeel van de toetspraktijk in het voortgezet onderwijs moeten blijven, waarin scholen de gelegenheid krijgen om brede onderwijskwaliteit en eigen specialisering in kaart te brengen. Het schoolexamen is niet bedoeld als oefenexamen voor het centraal schriftelijk examen, maar wordt hier wel regelmatig voor ingezet. De raad heeft tijdens veldraadplegingen diverse interessante mogelijkheden gehoord voor het vergroten van de zelfstandige rol van het schoolexamen. Het kan gaan om het omdraaien van het schoolexamen en het centraal schriftelijk eindexamen, om een portfoliomodel waarbij het schoolexamen een soort meesterproef of een serie meesterproeven wordt, of om het beperken van het centraal schriftelijk examen tot de kernvakken. De raad is er voorstander van deze alternatieven te onderzoeken en te evalueren.

Bij een zelfstandige rol van het schoolexamen wordt het des te belangrijker om de kwaliteit hiervan goed te borgen en hier goed op toe te zien. De perikelen op het vmbo Maastricht tijdens de afgelopen examenperiode hebben geleid tot een hernieuwde discussie over het belang van toezicht op het schoolexamen.⁹¹ De raad vindt dit begrijpelijk, omdat de consequenties van de Maastrichtse zaak zeer ingrijpend zijn geweest voor de betrokken leerlingen. Toch zou de raad de rijksoverheid willen oproepen om de zogenoemde risico-regelreflex te vermijden.⁹² Aan scholen komt ruimte toe om het schooldeel van het examen zelf in te vullen, uiteraard binnen de kaders van de eindtermen en met goede borging van de kwaliteit en goed toezicht.

Bij het **middelbaar beroepsonderwijs** kan het toezicht op de instellingsexamens volgens de raad generieker plaatsvinden, zoals ook al in het meest recente onderzoekskader mbo van de Inspectie wordt voorgestaan.⁹³ Door examinering van een opleiding als geheel te bekijken, kan de examendruk omlaag en kan worden toegewerkt naar een kleinere en beter behapbare set toetsen. Voorheen richtte de verantwoording op de dekking van het kwalificatiedossier zich te veel op details, waarbij instellingen zich weinig vrijheden en ruimte voor eigen invulling permitteerden. Opleidingen waren vooral gericht op volledige dekking van het kwalificatiedossier,

89 Onderwijsraad, 2014.

90 Onderwijsraad, 2017c.

91 Inspectie van het Onderwijs, 2018f.

92 "De risico-regelreflex is de neiging om het voortdurend verminderen van risico's, al dan niet naar aanleiding van een incident, als vanzelfsprekende opgave aan te merken. Deze reflex kan leiden tot disproportionele ingrepen: maatregelen waarvan de baten bij nadere beschouwing duidelijk niet opwegen tegen de kosten en de maatschappelijke bijwerkingen, of waarbij van de overheid meer wordt verwacht dan nodig of realiseerbaar is. Maatregelen kunnen uiteenlopen van wet- en regelgeving en normstelling tot toezicht, uitvoeringsmaatregelen en systeemveranderingen." Definitie afkomstig uit Trappenburg, 2011.

93 Inspectie van het Onderwijs, 2018c.

maar gaven aan dat het ondoenlijk was om aan alle gestelde eisen te voldoen bij de afnames en beoordelingen van examens in de praktijk. In het meest recente onderzoekskader van de Inspectie zijn de eisen niet meer zo specifiek geformuleerd, maar instellingen vinden het lastig om de vrijheid die er nu wel is, te benutten.

De raad signaleert wel behoefte aan duidelijke richtlijnen voor toetsing in de praktijk, omdat er nu vaak geen aandacht is voor de plaats van een onderdeel in het totaal en er te weinig wordt gedaan om dubbel toetsen te voorkomen. Uiteraard is het niet de bedoeling dat de sector terugkeert naar de situatie waarvoor het fijnmaziger toezicht toentertijd in het leven werd geroepen (zie pagina 23). De raad ziet de toenmalige wijziging ook als een uiting van de hierboven genoemde risico-regelreflex. Hij ziet de oplossing in meer deskundigheid en betrokkenheid (zie ook hoofdstuk 4) en niet in het voortdurend aanscherpen van de regelgeving.

Waar het gaat om de examencommissies in het middelbaar beroepsonderwijs is het volgens de raad verstandig de komende jaren vooral goed de vinger aan de pols te houden, conform de huidige vormgeving van het toezicht.⁹⁴ De taken en bevoegdheden van de examencommissies zijn in augustus 2017 aangescherpt. De commissies hebben nieuwe taken toegewezen gekregen en ook de verantwoordelijkheden van het bevoegd gezag ten aanzien van de examencommissies zijn veranderd. De kwaliteitsborging van deze commissies heeft in het onderzoekskader van de Inspectie een prominentere plaats gekregen. Zolang de beleidsmatige invoering van de nieuwe taken nog in ontwikkeling is, voert de Inspectie vaker volledig onderzoek uit naar de examencommissies. De raad ziet echter een degelijke regelmatige evaluatie als noodzakelijke onderdeel van de kwaliteitscyclus.

De visitaties in het kader van de NVAO in het **hoger onderwijs** richten zich volgens de raad formeel op papier in voldoende mate op de kwaliteit van de toetsing en examinering. Een adequaat systeem van toetsing en examinering is een van de vier standaarden aan de hand waarvan accreditatie wordt verleend en dit komt indirect ook aan bod bij de standaard van het aantonen van gerealiseerde onderwijsopbrengsten. Dit laatste gebeurt voornamelijk door middel van examinering. Het is vooral zaak om deze standaarden daadwerkelijk goed te evalueren. Dat betekent bijvoorbeeld dat het toezicht niet beperkt blijft tot het bekijken van een aantal eindwerken, maar dat ook de opbouw van het gehele toetsprogramma en de kwaliteit van de toetsen in beschouwing wordt genomen. Het betekent ook dat belangrijke competenties, zoals het uitvoeren van een onderzoek, niet eenmalig maar herhaaldelijk worden getoetst, waarbij de opdrachten steeds complexer worden. Het toezien op dergelijke kwaliteiten vergt toetsdeskundigheid. De raad beveelt aan om de toetsdeskundigheid van visitatiepanels verder te vergroten door scholing of het aantrekken van deskundige externen.⁹⁵ In het hoger beroepsonderwijs is het noodzakelijk dat toekomstige werkgevers niet alleen in voldoende mate betrokken zijn, maar ook inhoudelijk gekwalificeerd zijn om deel te nemen aan de eindbeoordeling. Daartoe is afstemming tussen de eisen en verantwoordelijkheden van toekomstige werkgevers en de beroepsopleidingen zelf van cruciaal belang.

Borg kwaliteit door verplichten toets- of examencommissies in het primair en voortgezet onderwijs
De raad adviseert de rijksoverheid om voor het primair en het voortgezet onderwijs een uit leraren/docenten bestaande toets- of examencommissie bij wet te verplichten. In het voortgezet onderwijs gaat het dan om een toets- en examencommissie, in het primair onderwijs om een toetscommissie, aangezien deze sector geen examens kent. Binnen de dimensie cen-

⁹⁴ Inspectie van het Onderwijs, 2018c.

⁹⁵ Inspectie van het Onderwijs, 2016b.

traal-decentraal leidt een dergelijke commissie, mits goed ingericht, tot het verbeteren van de kwaliteit van de toetsing en examinering op het decentrale niveau. Dergelijke commissies verbeteren de kwaliteitsborging van toetsing en examinering. Daarnaast dragen ze bij aan het vergroten van het eigenaarschap van toetsing en examinering bij onderwijsinstellingen en in het bijzonder bij leerkrachten/docenten. Ook neemt de deskundigheid op dit gebied zo toe. Alle betrokkenen in het onderwijs zouden, zoals de raad bepleit in hoofdstuk 4, over enige mate van toetsdeskundigheid moeten beschikken. Gezien de complexiteit van de materie acht de raad een commissie met specifieke verantwoordelijkheden en expertise op dit gebied echter wenselijk. De raad benadrukt dat hierbij teamverantwoordelijkheid met specifieke toetsdeskundigheid de voorkeur verdient om toename van werkdruk te voorkomen.

Net als in het middelbaar beroepsonderwijs en het hoger onderwijs is een toetscommissie in het primair onderwijs en een toets- en examencommissie in het voortgezet onderwijs primair belast met borging van de kwaliteit van toetsing en examens.⁹⁶ Voor de precieze taken en bevoegdheden kan inspiratie ontleend worden aan de regelingen voor de examencommissies in het middelbaar beroepsonderwijs en het hoger onderwijs. Tegelijk is het zaak de wettelijke regelingen voor de genoemde commissies af te stemmen op specifieke omstandigheden in die sectoren. Bij het opstellen van de wettelijke regeling is het van belang om naast de taken en bevoegdheden van de commissie ook de samenstelling en de positionering ervan goed te regelen. Hoe verhoudt de toetscommissie of de toets- en examencommissie zich tot het bestuur, de schoolleiding, het onderwijsteam, de individuele leraar/docent en de medezeggenschapsorganen?

In het funderend onderwijs kan een toetscommissie of een toets- en examencommissie op verschillende manieren vormgegeven worden. Bij grote scholen of scholengemeenschappen kan een school een eigen commissie krijgen. Het zou echter ook mogelijk moeten zijn om te kiezen voor één commissie voor meerdere scholen onder hetzelfde bevoegd gezag. In het geval van éénpitters met kleine scholen kan gewerkt worden met extra expertise van buitenaf, als binnen de school nog onvoldoende deskundigheid beschikbaar is. De raad vindt het in alle gevallen wel wenselijk om in ieder team kennis van toetsing in ieder geval bij één iemand te beleggen, ook als er sprake zou zijn van een overkoepelende commissie. Dit versterkt het eigenaarschap van toetsing en examinering binnen de school.⁹⁷

3.2 **Onderwijsinstelling: neem een standpunt in op belangrijke toetsvraagstukken**

Onderwijsinstellingen moeten op hun beurt duidelijke kaders bieden aan hun leraren en docenten, net zoals de rijksoverheid duidelijke kaders moet bieden aan de instellingen. De basis voor deze kaders wordt gevormd door een toetsvisie op instellingsniveau. Deze visie geeft aan hoe toetsing wordt afgestemd met onderwijsdoelen, onderwijsinhoud en onderwijsmiddelen. Voor het vormgeven van deze visie dienen instellingen keuzes te maken ten aanzien van belangrijke toetsvraagstukken. Welke ruimte krijgen leraren en docenten voor het zelf vormgeven van toetsen? Hoeveel wordt op hoger niveau belegd? Hoeveel aandacht krijgt formatieve toetsing? Welke gevolgen heeft dat voor de beslissende toetsing? Waar worden kwalitatieve toetsmethoden ingezet en waar wordt vastgehouden aan kwantitatieve toetsing?

⁹⁶ WHW 7:12b, lid 1a.

⁹⁷ Zie ook Onderwijsraad, 2015b, waarin de raad adviseert over examencommissies in het hoger onderwijs.

Vanuit dergelijke keuzes kunnen onderwijsinstellingen de ruimte die hun ter beschikking staat, optimaal benutten en de balans in hun eigen toets- en examenpraktijk aanbrengen.

De keuzes zullen leiden tot verschillen tussen onderwijsinstellingen. De duidelijke kaders van de rijksoverheid uit paragraaf 3.1 bieden ruimte voor deze verschillen en de raad moedigt deze verschillen aan als zij daadwerkelijk het resultaat zijn van visiegestuurde beleidskeuzes ten aanzien van toetsing. Niettemin doet de raad hieronder enkele generieke aanbevelingen. De raad adviseert de formatieve functie van toetsing te versterken en vaker te kiezen voor alternatieve vormen van toetsen en beoordelen. Daarnaast raadt hij het hoger onderwijs aan om de interne borging van de examenkwaliteit verder te versterken.

Breng toetsing in lijn met onderwijsdoelen, onderwijsinhoud en onderwijsmiddelen

Goed toetsbeleid begint in de ogen van de raad met de integratie van toetsing, onderwijsdoelen, onderwijsinhouden en onderwijsmiddelen.⁹⁸ Dit wordt gezien als een kenmerk van succesvolle onderwijssystemen, omdat leerlingen hiermee optimaal kunnen profiteren van het onderwijsaanbod.⁹⁹ Dit vraagt om een op uitkomsten gebaseerde benadering van onderwijs, waarbij van tevoren bepaald wordt wat de beoogde leeropbrengsten zijn. Onderwijsactiviteiten en evaluatiemethodieken worden dan, eveneens van tevoren, afgestemd op deze beoogde leeropbrengsten.¹⁰⁰ Het principe klinkt voor de hand liggend, maar is dit in de huidige onderwijspraktijk geenszins. Het vraagt niet alleen dat leraren hun eigen onderwijspraktijk zorgvuldig onder de loep nemen en een eigen onderwijsvisie ontwikkelen¹⁰¹, maar het vraagt ook op verschillende organisatieniveaus (teamleiders, directie en bestuur) dat onderwijs en toetsing als samenhangend systeem beschouwd worden.¹⁰² In Nederland experimenteert vooral het hoger beroepsonderwijs met een dergelijke aanpak op systeemniveau van verbetering van onderwijskwaliteit.¹⁰³ De raad ziet goede mogelijkheden voor dergelijke benaderingen in andere onderwijssectoren.

Bovengenoemde benadering zal er, als het goed is, toe leiden dat curriculum en toetsing gelijktijdig en in voortdurende onderlinge afstemming worden ontwikkeld en telkens weer worden verbeterd.¹⁰⁴ Toetsing kan worden bekeken op het niveau van losse toetsitems, een losse toets en het programma van toetsen als geheel, zowel per domein of leergebied/vak als voor de hele opleiding. Voor elk van deze niveaus gelden verschillende kwaliteitscriteria.¹⁰⁵ Goed toetsbeleid dat daadwerkelijk bijdraagt aan de onderwijskwaliteit, adresseert al deze niveaus.¹⁰⁶ Voor een criterium als validiteit betekent dit bijvoorbeeld dat het niet volstaat dat een toets op zichzelf aansluit bij de leerdoelen.¹⁰⁷ Ook het programma van alle toetsen samen dient aan te sluiten bij de gestelde doelen. Toetsing en curriculum zijn op die manier niet los te zien van elkaar. Een heldere visie op toetsing dient daarom duidelijk te maken hoe het systeem van toetsing

98 Organisation for Economic Co-operation and Development, 2010.

99 Scheerens, 2017; Scheerens & Exalto, 2017.

100 Biggs, 1996; Biggs, 2014.

101 Schönwetter, Sokal, Friesen, & Taylor, 2002.

102 Schildkamp, Heitink, Van der Kleij, Hoogland, Dijkstra, e.a., 2014.

103 Van Schilt-Mol, Sluijsmans, Peters, De Beer, & Jakobs, 2016; Sluijsmans, Joosten-ten Brinke, & Van Schilt-Mol, 2015.

104 In de software-ontwikkeling wordt gesproken van 'iteratieve en incrementele ontwikkeling'. Deze termen zijn vooral bekend sinds de introductie van 'agile' ontwikkelmethoden in dit vakgebied. Zie bijvoorbeeld Larman & Basili, 2003 en <http://agilemanifesto.org/>.

105 Kwaliteitscriteria op toetsitemniveau zijn: relevantie, objectiviteit, efficiëntie en moeijelijkheid. Kwaliteitscriteria op toetsniveau zijn: validiteit, betrouwbaarheid, objectiviteit, transparantie en normering. Kwaliteitscriteria op toetsprogrammaniveau zijn: passenheid bij doel, vergelijkbaarheid, acceptatie, betekenis, reproduceerbaarheid, eerlijkheid, zelf-toetsing, authenticiteit. Zie Sluijsmans, Peeters, Jakobs, & Weijzen, 2012.

106 Van Schilt-Mol, e.a., 2016.

107 Baartman, & Gulikers, 2014.

zich verhoudt tot het curriculum en hoe beide kunnen worden ontwikkeld zodat de samenhang tussen beide gewaarborgd is.

Een mogelijkheid is hierbij dat toetsing niet het sluitstuk, maar juist een integraal onderdeel vormt van het curriculumontwerp. De raad adviseert hierover in het briefadvies over curriculumherziening. In dat advies onderscheidt de raad drie processen bij curriculumvernieuwing. De overheid gaat over de herijking van kerndoelen en eindtermen. De onderwijsinstellingen zijn betrokken bij curriculumontwikkeling, inclusief toetsing, en bij onderwijsontwikkeling, inclusief toetsing.¹⁰⁸ De raad roept onderwijsinstellingen, net als de rijksoverheid, op om rolvast te zijn in dit proces. De ruimte voor onderwijsinstellingen zit vooral in de onderwijsontwikkeling en de toetsing die daarmee gemoeid is. Die ruimte kunnen zij beter benutten dan nu het geval is.

Versterk de formatieve functie van toetsing voor het onderwijsleerproces

De raad beveelt aan om de formatieve functie van toetsing in het onderwijs te versterken. In de huidige toetspraktijk krijgen toetsen vooral beslissende functies toegekend. Deze zijn, zoals eerder beargumenteerd, nodig en van groot belang, maar dragen er in de huidige toetspraktijk ook aan bij dat formatieve functies nog relatief weinig aandacht krijgen. Om beter evenwicht binnen de dimensie beslissend-formatief aan te brengen, zal soms van beslissende functies moeten worden afgezien, of zal waar mogelijk aan deze functies minder gewicht moeten worden toegekend.

Onder formatieve functies worden in het onderwijsveld uiteenlopende dingen verstaan. De raad hanteert een ruime definitie, waarbij formatieve toetsing kan worden ingezet om onderwijsbeleid te verbeteren, het onderwijs af te stemmen op de behoefte van leerlingen, en het leerproces van leerlingen te initiëren, vorm te geven en te optimaliseren.¹⁰⁹ Voor al deze verschillende formatieve functies geldt dat versterking nodig is. Het primair en het voortgezet onderwijs zetten toetsresultaten bijvoorbeeld beperkt in om onderwijsbeleid te verbeteren en doen te weinig aan feedback op het leerproces.¹¹⁰ Voor de leerling/student wordt vaak onvoldoende aangegeven wat er precies goed ging en wat nog niet, wat zij nu nodig hebben, en wat hen te doen staat.¹¹¹

Maar zelfs wanneer toetsing wel rijke feedback biedt, zoals uitvoerig commentaar op werkstukken, gebeurt het vaak dat leerlingen/studenten dit terzijde schuiven omdat zij geïnteresseerd zijn in het cijfer, en niet in het leren van hun fouten. Om een meer lerende houding bij hen te bewerkstelligen kan het helpen om minder vaak beslissend te toetsen. Een andere mogelijkheid is toetsen minder zwaar mee te laten tellen, zodat leerlingen/studenten het zich beter kunnen permitteren om een keer een lager cijfer te halen.¹¹² Het wegnemen van de druk van beslissende toetsing kan ook op schoolniveau helpen om meer aandacht voor formatieve functies te krijgen. De schoolleiding kan er bijvoorbeeld voor zorgen dat 'opbrengstgericht werken' wordt vormgegeven vanuit intrinsieke motivatie en niet is ingegeven vanuit een afrekencultuur of is opgelegd door de Inspectie.¹¹³

108 Onderwijsraad, 2018a.

109 In dit verband wordt ook gesproken van 'assessment for learning' (Schildkamp e.a., 2014)

110 Inspectie van het Onderwijs, 2018a.

111 'Feed up' informatie beschrijft wat de prestatiedoelen zijn, wat leerlingen moeten leren en welke criteria hiervoor gehanteerd worden. 'Feed back' informatie geeft aan in hoeverre die doelen al zijn bereikt. 'Feed forward' informatie geeft aan welke strategieën kunnen worden ingezet om de gestelde doelen te bereiken. Zie Fisher, & Frey, 2009; Hattie, & Timperley, 2007; William, 2016.

112 Onderwijsraad, 2017b.

113 Schildkamp e.a., 2014.

Het reduceren van het gewicht van beslissende toetsing volstaat echter niet om formatieve toetsing succesvol te implementeren. Effectieve formatieve toetsing vergt uiteenlopende randvoorwaarden die betrekking hebben op de toets, de docent, de leerlingen en de context.¹¹⁴ Een voorbeeld van een randvoorwaarde met betrekking tot de toets is het formuleren van goede feedback. Goede feedback is doelgericht, beschrijvend en voldoende concreet (maar niet te gedetailleerd) en biedt ruimte voor verbetering.¹¹⁵ Randvoorwaarden voor docenten zijn bijvoorbeeld voldoende vakinhoudelijke kennis en het vermogen om toetsing te gebruiken bij de evaluatie van het onderwijsproces. Een voorbeeld van een randvoorwaarde voor leerlingen/studenten is het vermogen tot zelfreflectie. Randvoorwaarden in de omgeving zijn onder andere een coherent toetsbeleid en een lerende cultuur.

De lerende cultuur als voorwaarde van formatieve toetsing vraagt onder meer dat leerkrachten en docenten de resultaten van hun leerlingen zien als het resultaat van hun eigen handelen en niet alleen als het resultaat van de inzet van de leerling.¹¹⁶ Dat vraagt een bereidheid om te kijken naar eigen handelen en kritisch te reflecteren op de gevolgen hiervan: toetsresultaten geven informatie over de mate waarin leerlingen kunnen profiteren van het onderwijs dat docenten verzorgen. Alle sectoren van het onderwijs zijn hier al mee bezig, door middel van feedbacktrainingen, collegiale consultaties, het vormgeven van lerende organisaties en professionele leergemeenschappen, enzovoort. Er valt daar in alle sectoren echter nog veel winst te behalen.¹¹⁷

De lerende cultuur die nodig is voor formatieve toetsing hangt ook samen met het toetsbeleid. Zolang toetsen op zichzelf staan en aan het eind van een onderwijsonderdeel worden gegeven, zullen leerlingen en studenten weinig betekenis toekennen aan de feedback die ze krijgen. De volgende toets is waarschijnlijk toch weer anders. De lerende houding van leerlingen en studenten kan alleen tot stand komen als er samenhang tussen toetsen bestaat en als er sprake is van een langetermijnperspectief. Om deze reden is formatieve toetsing vaak in verband gebracht met programmatisch toetsen en visiegestuurd toetsbeleid.¹¹⁸

Gebruik vaker kwalitatieve en alternatieve toets- en beoordelingsvormen

Evenwichtig toetsen betekent dat naast gestandaardiseerde toetsen ook alternatieve vormen van toetsen worden gebruikt. Dit vraagt om positie innemen binnen de dimensie kwantitatief-kwalitatief. Er kan bijvoorbeeld meer geïntegreerd getoetst worden, zoals dat gebeurt bij de meesterproeven in het middelbaar beroepsonderwijs. De raad acht meesterproeven ook inzetbaar in andere sectoren. Naast geïntegreerde vormen van toetsing kunnen alle sectoren meer gebruikmaken van globale/holistische oordelen in plaats van de meer gebruikelijke analytische toetsen.¹¹⁹ Ook adviseert de raad om de mogelijkheden van digitale toetsing beter te benutten.¹²⁰

Keuzes maken binnen de dimensie kwantitatief-kwalitatief kan ook tot gevolg hebben dat niet alle toetsresultaten uitgedrukt worden in een cijfer. De raad adviseert om, waar mogelijk en waar van toepassing, alternatieve vormen van beoordelen of inhoudelijke feedback te gebruiken. Cijfers liggen voor de hand en zijn handig voor het vergelijken van leerlingen en onder-

114 Schildkamp e.a., 2014.

115 Gulikers & Baartman, 2016; Yeh, 2010.

116 Abrams e.a., 2003.

117 Onderwijsraad, 2016b.

118 Schuwirth & Van der Vleuten, 2011; Dijkstra, Van der Vleuten, & Schuwirth, 2010.

119 Van der Vleuten & Schuwirth, 2005.

120 Zie ook Onderwijsraad, 2017a.

wijnsinstellingen. Cijfers zijn voor sommige functies van toetsing ook noodzakelijk, bijvoorbeeld voor het civiel effect van examens. Cijfers kunnen echter ook bijdragen aan het verminderen van de aangeboren intrinsieke motivatie tot leren.¹²¹ Kinderen gaan door becijfering van hun inspanning leren voor een cijfer.¹²² Ook het langetermijnperspectief is voor leerlingen moeilijk vast te houden door de becijfering van hun inzet, ze leren van toets naar toets zonder samenhang in het curriculum te kunnen zien of voor ogen te kunnen houden waar ze naar toe werken.

Borg kwaliteit door versterken examencommissies in het hoger onderwijs

In het hoger onderwijs ziet de raad, ondanks de aandacht voor de rol en positie van examencommissies in de afgelopen jaren,¹²³ nog ruimte voor verbetering. Vooral de rolvastheid van de verschillende commissies die betrokken zijn bij de kwaliteitszorg verdient aandacht. Formeel zijn examencommissies belast met de borging van de kwaliteit van tentamens en examens.¹²⁴ De opleidingscommissie adviseert over het bevorderen en waarborgen van de kwaliteit van de opleiding.¹²⁵ Examencommissies signaleren als er zaken niet goed verlopen in de toets- en examenpraktijk binnen een opleiding. Een examencommissie heeft een beperkt aantal instrumenten om zelf in te grijpen. Zo kan een commissie de aanwijzing van een docent als examinerator intrekken¹²⁶ of de richtlijnen aanscherpen.¹²⁷ Een belangrijk aandachtspunt hierbij is dat docenten ook op de hoogte zijn van de richtlijnen en geborgd wordt dat richtlijnen daadwerkelijk effect op de onderwijspraktijk hebben. Verder kan een commissie signalen alleen doorspelen naar het management en bestuur van de opleiding. Deze laatste partijen zijn verantwoordelijk voor het verbetertraject. In de huidige toets- en examenpraktijk in het hoger onderwijs komt het voor dat de examencommissie signaleert en de opleidingscommissie de verantwoordelijkheid voor het verbetertraject oppakt, in plaats van het bestuur of het management. Hiermee ontstaat een complexe en onwenselijke dubbelrol voor de opleidingscommissie, die uiteindelijk zijn eigen vlees moet keuren.¹²⁸

De raad wijst er verder op dat het functioneren van een examencommissie als formeel instrument voor kwaliteitsborging moet passen binnen het algehele kwaliteitsbeleid van een instelling en opleiding. In zijn advies *Kwaliteit in het hoger onderwijs* heeft de raad aandachtspunten benoemd voor de afstemming tussen effectieve kwaliteitszorg en een sterke kwaliteitscultuur op opleidingsniveau en wat dat betekent voor de opstelling van bestuurders en organen binnen een instelling.¹²⁹ Een examencommissie zal in elk geval uit moeten gaan van de onderwijsvisie van de opleiding, zodat er geen frictie ontstaat tussen het toetsbeleid en wat docenten en studenten als opleidingsgemeenschap goed onderwijs vinden.¹³⁰ Hetzelfde geldt voor de afstemming tussen het toetsbeleid en wat (nieuwe) docenten meekrijgen in het kader van professionaliseringsactiviteiten, zoals een bko- of bke-traject (respectievelijk basiskwalificatie onderwijs en basiskwalificatie examinering).

121 Carbonneau, Vallerand, & Lafrenière, 2012; White, 1959; Lepper, Greene, & Nisbett, 1973.

122 Butler & Nisan, 1986; Pulfrey, Darnon, & Butera, 2013.

123 Inspectie van het Onderwijs, 2015b; 2016c. Zie ook <https://zoek.officielebekendmakingen.nl/stb-2016-273.html>

124 WHW artikel 7:12b.

125 WHW artikel 9:18.

126 WHW artikel 7:12c. De wet zelf spreekt alleen over aanwijzing van examinatoren, uit jurisprudentie volgt echter dat intrekking ook kan, mits dit goed gemotiveerd is.

127 WHW artikel 17:12b, lid 1, onder b.

128 Van den Berg, 2018; Van de Veen, Pol, & Moerkerke, 2018.

129 Onderwijsraad, 2015b.

130 De examencommissie dient in dit kader te werken binnen het kader van de Onderwijs- en Examenregeling (WHW 7:12b lid1 onder b).

Voor een evenwichtige toetspraktijk is het vergroten van de toetsdeskundigheid in het onderwijsveld noodzakelijk. Dit is te realiseren door het aanpassen van beroepsstandaarden, het verbeteren van het aanbod in professionaliseringstrajecten, en het delen van informatie met alle gebruikers van toetsen.

4 Vergroot de toetsdeskundigheid in het onderwijs

Doordat instellingen niet vertrouwd zijn met formatieve toetsing en kwalitatieve toetsmethoden, wordt toetsing tot nu toe beperkt ingezet voor het geven van feedback. Dit leidt onbedoeld tot een versmalling van de focus van onderwijs. Ook leidt gebrek aan toetsdeskundigheid tot onnodig vaak uitbesteden van de vormgeving van toetsing aan externe partijen. Om de handelingsbekwaamheid en het eigenaarschap ten aanzien van toetsing te versterken is het nodig de toetsdeskundigheid van onderwijsprofessionals (waaronder leraren/docenten en schoolleiders/onderwijsmanagers) te vergroten. De raad adviseert de rijksoverheid en de beroepsverenigingen om de beroepsstandaarden van verschillende groepen onderwijsprofessionals te actualiseren en meer en beter aandacht te besteden aan toetsing in professionaliseringstrajecten. Daarnaast adviseert hij onderwijsprofessionals om hun kennis te delen met een bredere groep van betrokkenen, waaronder bestuurders en ouders. Deze andere betrokkenen hoeven geen toetsen te kunnen ontwikkelen, maar moeten wel toetsresultaten kunnen interpreteren en eventueel mee kunnen denken over het toetsbeleid.

4.1 Actualiseer professionele standaarden op het gebied van toetsing

Bij toetsing zijn verschillende groepen onderwijsprofessionals betrokken, waaronder leraren/docenten, leidinggevendenden in het onderwijs, examenfunctionarissen, lerarenopleiders en toetsaanbieders. Hoewel niet iedere professional over dezelfde toetsdeskundigheid hoeft te beschikken, geldt voor al deze groepen dat het geheel van noodzakelijke kennis, vaardigheden en houdingsaspecten aanzienlijk is en op doelgerichte en samenhangende wijze nagestreefd dient te worden.¹³¹ De raad acht professionele standaarden daarbij zeer behulpzaam, mits deze standaarden voldoende specifiek, samenhangend en compleet zijn. Dat is op dit moment voor de meeste beroepsgroepen en sectoren niet het geval. Daarom adviseert de raad aan de rijksoverheid en de beroepsverenigingen om de professionele standaarden op het gebied van toetsing te actualiseren dan wel te ontwikkelen. Hieronder wordt nader toegelicht dat professionele standaarden zowel voor leraren/docenten als voor andere groepen professionals nodig zijn.

¹³¹ Brookhart, 2011.

Formuleer omvattende bekwaamheidseisen en aanvullende richtlijnen voor leraren/docenten

Ten aanzien van toetsing zijn de huidige bekwaamheidseisen voor leraren/docenten in het primair en voortgezet onderwijs en het middelbaar beroepsonderwijs onvolledig en bovendien weinig concreet. Aanvullende richtlijnen, zoals de kennisbases voor lerarenopleidingen, schieten vaak ook tekort. De raad doet daarom de aanbeveling aan de rijksoverheid om de bekwaamheidseisen omvattend te maken ten aanzien van toetsing. Daarnaast beveelt hij overige relevante landelijke organisaties voor onderwijsgevend personeel aan om specifieke richtlijnen te onderscheiden, waarbij de eisen ten aanzien van toetsing bij elkaar staan en op complete wijze zijn uitgewerkt tot op het niveau dat ze goed toetsbaar zijn. Als voorbeeld kunnen hierbij de basis- en seniorkwalificaties examinering (bke en ske) van de Vereniging Hogescholen worden genomen.

Voor het primair en voortgezet onderwijs en het middelbaar beroepsonderwijs zijn de bekwaamheidseisen voor onderwijspersoneel wettelijk vastgelegd en onlangs herijkt.¹³² Een nadere specificatie van pedagogische, vakinhoudelijke en vakdidactische domeinen heeft geresulteerd in een aanzienlijk aantal verwijzingen naar toetsdeskundigheid. Dit geldt met name voor het didactische domein, waar is uitgegaan van een cyclische aanpak van observeren, toetsen, evalueren, bijstellen en ontwikkelen.¹³³ De raad constateert met tevredenheid dat het uitgangspunt van een cyclische aanpak heeft gezorgd voor een ruime opvatting van toetsing, waarbij ook de formatieve functies een plaats krijgen. Het toetsen en evalueren is bij deze aanpak immers gericht op het bijstellen en (door)ontwikkelen van het onderwijs. Niettemin constateert de raad dat de herijkte bekwaamheidseisen niet compleet zijn. Dat wordt duidelijk als de eisen worden gelegd naast recente indelingen van toetsdeskundigheid die door onderzoekers zijn opgesteld.¹³⁴ Opvallend afwezig zijn de eisen dat de leraar/docent externe toetsresultaten moet kunnen gebruiken voor onderwijsbeslissingen; dat hij resultaten en beslissingen moet kunnen communiceren naar de studenten, de schoolleiding en de ouders; en dat hij leerlingen moeten kunnen helpen om op basis van toetsinformatie de eigen leerstrategie te bepalen. Ook de wettelijke en ethische verantwoordelijkheden die de leraar of docent heeft ten aanzien van toetsing, blijven ongespecificeerd.¹³⁵

De genoemde bekwaamheidseisen bieden onvoldoende houvast voor professionalisering, begeleiding en toezicht op het gebied van toetsdeskundigheid. Het is bijvoorbeeld niet duidelijk hoe de leraar of docent invulling kan geven aan “toetsen...ontwikkelen, toetsresultaten beoordelen, analyseren en interpreteren en de kwaliteit van toetsen en examens...beoordelen”¹³⁶ of hoe kan worden geëvalueerd of de docent aan deze eis voldoet.

Beroepsstandaarden omvatten, naast de wettelijke bekwaamheidseisen, ook richtlijnen, die tot op zekere hoogte een aanvulling of een uitwerking op de eisen bieden of een vertaling naar het algemene aanbod van lerarenopleidingen. In het primair en voortgezet onderwijs zijn dit bijvoorbeeld *Kennisbasis leraar basisonderwijs* en *Generieke kennisbasis voor tweedegraads lerarenopleidingen* van de Vereniging Hogescholen, *Vakspecifieke beroepsstandaarden Talen* van de vereniging van Levende Talen¹³⁷ en de *Geactualiseerde kennisbasis* van de Vereniging van Lerarenopleiders. Voorbeelden van aanvullende beroepsstandaarden in het middelbaar beroepsonderwijs zijn *Kennisbasis Leraar Technisch Beroepsonderwijs* van 10 voor de leraar en *Specifieke*

¹³² Besluit bekwaamheidseisen onderwijspersoneel, 16 maart 2017.

¹³³ Onderwijscoöperatie, 2016.

¹³⁴ Brookhart, 2011; DeLuca & Bellara, 2013; Mandinach, & Gummer, 2016.

¹³⁵ Vergelijk Brookhart, 2011.

¹³⁶ Besluit bekwaamheidseisen onderwijspersoneel, artikel 2.9k.

¹³⁷ De Graaff & Bodde, 2009.

beroepscompetenties van docenten in het middelbaar beroepsonderwijs van de Vereniging Hogescholen. In het hoger onderwijs zijn er basis en seniorkwalificaties onderwijs van de VSNU (vereniging van universiteiten) en de basis- en seniorkwalificatie examinering van de Vereniging Hogescholen.

Deze niet-wettelijke richtlijnen voor docentcompetenties of voor het aanbod van lerarenopleidingen blijven ten aanzien van toetsdeskundigheid vaak summier en incompleet. De *Generieke kennisbasis voor tweedegraads lerarenopleidingen* onderscheidt bijvoorbeeld relevante kernconcepten en kernmethodieken op het gebied van toetsing, maar een onderwerp als communicatie van toetsresultaten ontbreekt.¹³⁸ Het blijft daarnaast onduidelijk wat precies van studenten op het gebied van toetsdeskundigheid wordt verwacht en hoe deze deskundigheid kan worden geëvalueerd. Een belangrijke uitzondering vormen de bke- en ske-eisen die de Vereniging Hogescholen recent heeft geformuleerd.¹³⁹ De bke- en ske-eisen hebben een aantal belangrijke voordelen. Ze zijn verbonden met indicatoren die concreet maken wanneer docenten aan de eis voldoen. Daarnaast omvatten de eisen naast de vormgeving van toetsing bij afzonderlijke vakken ook de vormgeving van toetsing op het niveau van het curriculum. En ten slotte wordt een basisniveau gevolgd door een seniorniveau dat zich richt op de toetsing van grotere delen van het curriculum. Om deze redenen zouden het primair en voortgezet onderwijs en het middelbaar beroepsonderwijs hier een voorbeeld aan kunnen nemen. Dat geldt ook voor het wetenschappelijk onderwijs. Weliswaar maakt in deze sector toetsing onderdeel uit van de certificeringstrajecten bko en sko (basiskwalificatie en seniorkwalificatie onderwijs), maar deze eisen zijn ondanks landelijke afspraken verschillend per instelling¹⁴⁰ en doorgaans minder omvattend en specifiek dan de bke- en ske-eisen in het hoger beroepsonderwijs.

Formuleer ook toetsingsstandaarden voor andere groepen onderwijsprofessionals

Professionele standaarden ten aanzien van toetsing zijn niet alleen nodig voor leraren/docenten, maar ook voor andere professionals. De beroepsprofielen van schoolleiders verwijzen bijvoorbeeld nauwelijks naar toetsdeskundigheid,¹⁴¹ terwijl de vormgeving van het toetsbeleid en het gebruik van toetsing voor de kwaliteitszorg taken van schoolleiders zijn. De raad adviseert beroepsverenigingen in alle onderwijssectoren de eisen op het gebied van toetsdeskundigheid van leidinggevenden nader uit te werken of vorm te geven¹⁴² en daarbij aan te sluiten bij de literatuur die hiervoor specifieke competenties onderscheidt.¹⁴³ De raad acht dergelijke standaarden ook nodig voor lerarenopleiders en nascholers. Hun taken op het gebied van toetsing zijn tweeledig. Zij moeten er niet alleen voor zorgen dat zij de kwaliteiten van (aankomende) leraren en docenten op effectieve wijze beoordelen, maar moeten ook hun studenten/cursisten begeleiden bij het verwerven van toetsdeskundigheid. De landelijke beroepsstandaarden voor lerarenopleiders verwijzen echter slechts summier naar toetsdeskundigheid.¹⁴⁴

¹³⁸ HBO-raad vereniging Hogescholen, 2011.

¹³⁹ Expertgroep BKE/SKE, 2013.

¹⁴⁰ Inspectie van het Onderwijs, 2016b.

¹⁴¹ Krüger & Andersen, 2017; VO Academie, 2014; Basoski & Koot, 2014.

¹⁴² Voor onderwijsmanagers in het hoger onderwijs geldt dat er nog geen landelijke beroepsstandaarden zijn. Weliswaar worden professionaliseringstrajecten als *Onderwijskundig leiderschap in het hoger onderwijs* afgesloten met een certificaat, maar de bekwaamheidseisen van dergelijke trajecten lopen uiteen tussen instellingen.

¹⁴³ McMillan, 2000; Hargreaves, Boyle, & Harris, 2014.

¹⁴⁴ Velon, 2018.

Lerarenopleidingen besteden volgens de raad te weinig aandacht aan toetsdeskundigheid. Dat geldt ook voor andere individuele professionaliseringstrajecten, zoals de basiskwalificatie onderwijs in het wetenschappelijk onderwijs. De raad adviseert om in de verschillende trajecten niet alleen meer, maar ook meer gericht aandacht aan toetsing te geven. Hiertoe is het in de eerste plaats nodig om concrete beroepsstandaarden ten aanzien van toetsdeskundigheid te formuleren (zie paragraaf 4.1). Ten tweede beveelt de raad aan om naast geïntegreerde scholing in toetsing (als onderdeel van algemene didactiek of vakdidactiek) ook afzonderlijk onderwijs over toetsing te geven. Ten derde adviseert hij om de scholing in toetsing beter te borgen. Naast individuele professionaliseringstrajecten ziet de raad ook goede mogelijkheden om de toetsdeskundigheid te versterken via collectieve professionaliseringstrajecten.

Naast geïntegreerde ook afzonderlijke scholing in toetsing

Eerste- en tweedegraads lerarenopleidingen besteden in beperkte mate aandacht aan toetsingsvraagstukken. Nascholingstrajecten richten zich eveneens op het vergroten van toetsdeskundigheid. Deze scholing blijft echter vaak beperkt tot een korte cursus of zijdelingse behandeling van toetsthema's. Niet alleen in Nederland, maar ook in andere landen wordt beperkt aandacht gegeven aan toetsing in lerarenopleidingen en nascholing. Scholing in toetsing in de Verenigde Staten bijvoorbeeld is wisselend, doorgaans kort, eenzijdig en niet passend bij de beleidsdoelstellingen. Terwijl het onderwijsbeleid bijvoorbeeld het belang van opbrengstgericht werken benadrukt, wordt aan formatieve toetsthema's zoals functies van toetsing en communicatie van toetsresultaten weinig aandacht geschonken.¹⁴⁵ Een van de verklaringen voor de beperkte aandacht voor toetsing is dat toetsing vaak geïntegreerd is in algemene of vakdidactische cursussen waarin het ontwerpen van lessen en het lesgeven centraal staat.¹⁴⁶ Het voordeel is dat duidelijk gemaakt kan worden dat toetsing nauw is verweven met vakinhoudelijke kennis en algemene didactische principes. Om een goede toets te maken, moeten leraren/docenten bijvoorbeeld begrijpen hoe het leerproces bij een bepaald vakdomein verloopt en hoe leerdoelen concreet kunnen worden gemaakt in toetstaken.¹⁴⁷ Het risico van het integreren van het toetsingsonderwijs is dat het ondergesneeuwd raakt en dat daarmee de samenhang van theoretische noties van toetsing verloren gaat. Om die reden doet de raad de aanbeveling om toetsing ook als apart onderdeel in de opleiding op te nemen.

Borg scholing in toetsing

De beperkte aandacht voor toetsing in lerarenopleidingen is ook een gevolg van de ondergeschikte plaats die toetsing heeft in de kennisbasis van eerste- en tweedegraads lerarenopleidingen. Duidelijke eisen vormen een eerste aspect in de borging van het vergroten van toetsdeskundigheid. Een tweede facet is het toetsen van de eisen die aan toetsdeskundigheid worden gesteld. Dit gebeurt momenteel zeer beperkt. Zo vormt toetsdeskundigheid geen domein in de landelijke kennistoetsen die een groot deel van de tweedegraadslerarenopleidingen gezamenlijk afneemt. Een derde aandachtspunt voor borging is dat bij opleidingsaccreditaties de toetsdeskundigheid van studenten/cursisten een belangrijk onderdeel dient te vormen. Afgaand op de systeembrede analyses van de visitaties van de verschillende lerarenopleidingen is dat nu niet het geval.¹⁴⁸ De raad is van mening dat de rijksoverheid via de Inspectie en de NVAO bijzondere aandacht zou moeten hebben voor het realiseren van de gestel-

¹⁴⁵ Greenberg, & Walsh, 2012; DeLuca & Bellara, 2013.

¹⁴⁶ DeLuca & Bellara, 2013.

¹⁴⁷ Baartman & Gulikers, 2014; Sluijsmans e.a., 2012.

¹⁴⁸ Nederlands-Vlaamse Accreditatieorganisatie, 2016.

de eisen ten aanzien van toetsdeskundigheid. Een dergelijke focus is gelegitimeerd, omdat toetsdeskundigheid van aankomende leraren een belangrijke waarborg vormt voor de onderwijskwaliteit in de sector waar ze les gaan geven. Het vierde aspect van borging is voldoende toetsdeskundigheid bij lerarenopleiders. Ook hiervoor geldt dat nadere specificatie van de beroepsstandaarden nodig is.

Neem misverstanden over toetsing weg

Traditioneel wordt toetsingsonderwijs gericht op het maken van een toets (het geheel van items en het scoringsvoorschrift) en het hanteren van de kwaliteitscriteria die daarvoor gelden, zoals betrouwbaarheid, validiteit, objectiviteit, normering en transparantie.¹⁴⁹ Wanneer het toetsingsonderwijs hiertoe beperkt blijft, kan dit bijdragen aan de veelvoorkomende misconceptie dat 'een goede toets' volstaat. Zowel de eerste- en tweedegraads lerarenopleidingen als de basis- en seniorkwalificatie onderwijs en diverse nascholingsprogramma's dienen te benadrukken dat toetsing ook het gebruik van toetsresultaten, het toetsprogramma als geheel, en de organisatie rondom toetsing omvat.¹⁵⁰ Op dit meer systemische niveau van toetsing gelden andere kwaliteitscriteria, zoals aansluiting bij bredere onderwijsdoelen, vergelijkbaarheid van toetsresultaten, acceptatie van de toetsen, de betekenis die toetsing heeft voor betrokkenen, de authenticiteit van de toets, de herhaalbaarheid van toetsresultaten en de eerlijkheid van beslissingen die worden gebaseerd op toetsen.¹⁵¹ Uit trainingen die gericht zijn op de kwaliteit van toetsprogramma's blijkt dat veel onderwijsprofessionals uit zichzelf deze bredere toetsingscriteria niet hanteren. Dat toetskwaliteit over meer gaat dan de toets beschouwen ze als een eye-opener.¹⁵² In het verlengde van een systemische benadering van toetsing kunnen andere veelvoorkomende misconcepties aan de orde komen, zoals het idee dat ondersteunende functies en beslissende functies van toetsing onverenigbaar zouden zijn (dit is een misverstand omdat feedback op toetsen met een beslissende functie juist waardevol is)¹⁵³ of dat goede toetsing volledig objectief moet zijn (dit is een misverstand omdat maatregelen kunnen worden getroffen om met beoordelaarseffecten om te gaan).¹⁵⁴

Oefenen van vaardigheden en langdurige professionalisering

Het wegnemen van misverstanden over toetsing en het bijbrengen van specifieke toetsingsconcepten zijn belangrijk, maar toetsdeskundigheid omvat ook vaardigheden. Lerarenopleidingen dienen tijd in te ruimen voor het oefenen, bijvoorbeeld in het gebruik van toetsmatrijzen. Het analyseren van toetsen en toetsitems is een complexe vaardigheid.¹⁵⁵ Dat geldt ook voor het construeren van toetsen en toetsitems die aansluiten bij leerdoelen.¹⁵⁶ Tegelijkertijd zijn deze vaardigheden essentieel om bijvoorbeeld geen reproductieve toetsen te geven wanneer het doel is om inzicht te meten.¹⁵⁷

Verwerving van toetsdeskundigheid is een proces dat zich verder uitstrekt dan een lerarenopleiding of een enkele nascholingscursus.¹⁵⁸ De raad doet de aanbeveling om verschillen-

¹⁴⁹ Baartman, 2008.

¹⁵⁰ Dat gebeurt nu in Nederland nog te weinig. Zie: Joosten-Ten Brinke, 2011; Sluijsmans e.a., 2012; Van Schilt-Mol e.a., 2016. Ook in andere landen, zoals de Verenigde Staten krijgen 'formatieve toetsthema's', zoals functies van toetsing en communicatie van toetsresultaten doorgaans weinig aandacht (Greenberg & Walsh, 2012; DeLuca & Bellara, 2013).

¹⁵¹ Baartman, 2008; Gulikers, Sluijsmans, & Baartman, 2008.

¹⁵² Gulikers, & Baartman, 2016; zie ook DeLuca & Bellara, 2013, die in dit verband spreken van 'een omslag in het denken'.

¹⁵³ Baartman & Gulikers, 2014.

¹⁵⁴ Van der Vleuten, Norman & Graaff, 1991.

¹⁵⁵ Krathwohl, 2002.

¹⁵⁶ Baartman & Gulikers, 2014; Sluijsmans e.a., 2012.

¹⁵⁷ Brookhart, 2011.

¹⁵⁸ Schildkamp e.a., 2014.

de professionaliseringstrajecten op elkaar te laten aansluiten, zodat onderwijsprofessionals gaandeweg hun deskundigheid kunnen verdiepen. Een goed voorbeeld in dit verband zijn de opeenvolgende trajecten van de basiskwalificatie examinering en de seniorkwalificatie examinering in het hoger beroepsonderwijs.¹⁵⁹

Gezamenlijk ontwikkelen van toetsdeskundigheid

De kwaliteit van toetsing wordt niet alleen bepaald door afzonderlijke onderwijsprofessionals, maar ook door de mate waarin zij samenwerken. Gezamenlijke ontwikkeling van het toetsbeleid kan in belangrijke mate bijdragen aan de professionalisering van zowel de instelling als de individuele professionals.¹⁶⁰ Een geschikt startpunt hierbij is de gezamenlijke analyse van de eigen toetspraktijk aan de hand van bewijsstukken, zoals verslagen van de examencommissie, toetsmatrijzen, trainingsmateriaal, resultaat- en ontwikkelingsgesprekken en studentevaluaties.¹⁶¹ Belangrijk bij deze collectieve projecten is dat de verschillende groepen onderwijsprofessionals en belanghebbenden bij toetsing zijn vertegenwoordigd.¹⁶² Om toetsing praktijkgericht te maken kan het verder zinvol zijn om professionals uit het beroepenveld erbij te betrekken.¹⁶³ In alle gevallen is voldoende toetsexpertise van belang om de gezamenlijke projecten te begeleiden. Zo nodig dienen externe toetsexperts te worden ingeschakeld.¹⁶⁴ Collectieve trajecten waarbij professionalisering en innovatie hand in hand gaan, hoeven uiteraard niet gebonden te zijn aan één instelling. Er zijn ook trajecten mogelijk waarbij verschillende instellingen toetsbeleid ontwikkelen, zoals de medische faculteiten doen.¹⁶⁵ Andere vormen zijn instellingsoverstijgende leergemeenschappen of onderzoeksconsortia die professionalisering, innovatie en onderzoek op het gebied van toetsing combineren.¹⁶⁶

4.3 Deel informatie over toetsing met andere belanghebbenden

Naast onderwijsprofessionals hebben ook andere groepen belang bij toetsing – leerlingen, ouders, bestuurders, leden van medezeggenschapsorganen, beleidsmakers, werkgevers, politici en de burger in het algemeen. Toetsresultaten informeren deze groepen over onderwijsopbrengsten en de onderwijskwaliteit. Voor leerlingen/studenten geldt dat toetsing een belangrijke onderdeel van hun dagelijkse praktijk vormt en dat zij toetsresultaten kunnen benutten om hun leerproces te begrijpen en te optimaliseren.¹⁶⁷ Voor al deze groepen is tot op zekere hoogte kennis over toetsing nodig. Kennis maakt het mogelijk om toetsresultaten goed te interpreteren en beredeneerd stelling te nemen in de discussie over de vormgeving van de toetspraktijk.

Er kan niet van alle belanghebbenden bij toetsing worden verwacht dat zij professionele scholing volgen. Het is wel reëel en wenselijk dat onderwijsprofessionals hen informeren over basale principes van toetsing. De raad beveelt onderwijsprofessionals daarom aan om hun toetsdeskundigheid te delen met bredere groepen belanghebbenden. Evenals bij onderwijsprofessionals geldt voor deze groepen dat het wegnemen van misconcepties over toetsing

¹⁵⁹ Expertgroep BKE/SKE, 2013.

¹⁶⁰ Van Schilt-Mol e.a., 2016.

¹⁶¹ Sluijsmans e.a., 2012.

¹⁶² Van Schilt-Mol e.a., 2016.

¹⁶³ Schildkamp e.a., 2014.

¹⁶⁴ Van Schilt-Mol e.a., 2016.

¹⁶⁵ Van der Vleuten e.a., 2018.

¹⁶⁶ Teurlings & Uerz, 2009; Hollenbergh, Keuning, & Meijer, 2016.

¹⁶⁷ Brookhart, 2011.

prioriteit heeft. Het delen van kennis over toetsing helpt onderwijsprofessionals om draagvlak te creëren en steun te organiseren voor het gevoerde toetsbeleid. Voor leerlingen/studenten geldt dat naast basale kennis over toetsen ook vaardigheden moeten worden onderwezen om toetsen te benutten voor het leerproces. Deze vaardigheden zijn nauw verweven met vaardigheden voor zelfstandig leren en vereisen langdurige oefening. Een ultiem doel daarbij is dat leerlingen/studenten gaandeweg beter in staat worden om zichzelf te toetsen¹⁶⁸ of bruikbare peer feedback kunnen geven aan medeleerlingen/studenten.¹⁶⁹

¹⁶⁸ Adesope, Trevisan, & Sundararajan, 2017.

¹⁶⁹ Dochy, Segers, & Sluijsmans, 1999.

Afkortingen

bke	basiskwalificatie examinering
bko	basiskwalificatie onderwijs
Cito	Centraal Instituut voor Toetsontwikkeling
KCE	KwaliteitsCentrum Examinering
mbo	middelbaar beroepsonderwijs
NVAO	Nederlands-Vlaamse Accreditatieorganisatie
OCW	Onderwijs, Cultuur en Wetenschap
roc	regionaal opleidingencentrum
ske	seniorkwalificatie examinering
sko	seniorkwalificatie onderwijs
vmbo	voorbereidend middelbaar beroepsonderwijs
vo	voortgezet onderwijs
VSNU	(landelijke vereniging van Nederlandse universiteiten)
WEB	Wet educatie en beroepsonderwijs
WHW	Wet op het hoger onderwijs en wetenschappelijk onderzoek
WPO	Wet op het primair onderwijs
WVO	Wet op het voortgezet onderwijs

Literatuur

- Abrams, L.M., Pedulla, J.J., & Madaus, G.F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice*, 42(1), 18-29.
- Adesope, O.O., Trevisan, D.A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659-701.
- Archer, E. (2017). The assessment purpose triangle: Balancing the purposes of educational assessment. *Frontiers in Education*, 2(4), 41-48.
- Baartman, L. (2008). *Assessing the assessment: Development and use of quality criteria for competence assessment programmes*. Proefschrift. Utrecht: Universiteit Utrecht.
- Baartman, L.K.J., & Gulikers, J.T.M. (2014). Beoordelen als fundament voor goed opleiden in het beroepsonderwijs: Analyse van toetsprogramma's in het mbo en hbo. *Pedagogische Studiën*, 91(1), 54-68.
- Baartman, L., & Gulikers, J. (2017). Assessment in Dutch vocational education: Overview and tensions of the past 15 years. In E. de Bruijn, S. Billett & J. Onstenk (eds.), *Enhancing Teaching and Learning in the Dutch Vocational Education System* (245-266). Springer International Publishing.
- Basoski, I, & Koot, T. (2014). *De teamleider in positie. Advies over de inhoud en implementatie van een beroepsprofiel voor leidinggevenden van onderwijsteams in het mbo*. Geraadpleegd op 6 december 2018 via <https://www.mborraad.nl/publicaties/profiel>.
- Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3), 287-302.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347-364.
- Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1(1), 5-22.
- Black, P.J. (1998). *Testing: friend or foe? The theory and practice of assessment and testing*. London: Falmer Press.
- Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.
- Bologna Working Group (2005). *A Framework for Qualifications of the European Higher Education Area*. Bologna Working Group Report on Qualifications Frameworks. Copenhagen: Ministry of Science, Technology and Innovation.
- Bosker, R., & Van de Vorle, R. (2014). *Advies over de uitwerking van de referentieniveaus 2F en 3F voor rekenen in toetsen en examens*. Enschede: SLO.
- Bouwer, R., Goossens, M., Mortier, A, Lesterhuis, M., & De Maeyer, S. (2018). Een comparatieve aanpak voor peer assessment: leren door te vergelijken. In D. Sluijsmans & M. Segers (red.), *Toetsrevolutie: naar een feedbackcultuur in het hoger onderwijs* (92-106). Phronese.
- Brookhart, S.M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12.
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210-216.
- Canton, J., Aler, T., Heemskerk, K., Van der Westen, W., & Willemsen, K. (2013). *Effecten van sturing op discrepantie tussen de cijfers van het centraal examen. Onderzoek naar de sturing door schoolleiders en de effecten daarvan op het taalonderwijs op havo en vwo en het schoolexamen bij de talen*. Onderzoek in opdracht van de Vereniging van Levende Talen uitgevoerd door Onderwijs Innovatie Groep met subsidie van het Ministerie van OCW.
- Carbonneau, N., Vallerand, R.J., & Lafrenière, M.A.K. (2012). Toward a tripartite model of intrinsic motivation. *Journal of Personality*, 80(5), 1147-1178.

- Centrale Examencommissie Vaststelling Opgaven (2006). *De verhouding tussen het schoolexamen en het centraal examen*.
- College voor Toetsen en Examens (2015). *Samenvatting opbrengsten onderzoeken commissie Bosker*. Geraadpleegd op 6 december 2018 via <https://www.cvte.nl/documenten/rapporten/2015/07/01/samenvatting-bosker-onderzoeken>.
- College voor Toetsen en Examens (2018a). *Brochure kaders rekentoets VO 2019*. Geraadpleegd op 6 december 2018 via <https://www.cito.nl/onderwijs/voortgezet-onderwijs/centrale-examens-voortgezet-onderwijs/rekentoets-vo>.
- College voor Toetsen en Examens (2018b). *Combinatiecijfer (havo/vwo)*. Geraadpleegd op 6 december 2018 via <https://www.examenblad.nl/onderwerp/combinatiecijfer-havo-vwo/2019>.
- College voor Toetsen en Examens (2018c). *De digitale adaptieve Centrale Eindtoets 2019 gaat niet door*. Geraadpleegd op 6 december 2018 via <https://www.cvte.nl/actueel/nieuws/2018/11/29/de-digitale-adaptieve-centrale-eindtoets-2019-gaat-niet-door>.
- Commissie Externe Validering Examenkwaliteit Hoger Beroepsonderwijs (2012). *Vreemde ogen dwingen*. Den Haag: HBO-raad, vereniging van hogescholen.
- Commissie Verantwoorde Invoering Rekentoets (2014). *Doordacht doorzetten naar een hoger rekenniveau*. Enschede: SLO.
- Conseil Supérieur de l'Éducation Quebec (2014). *Toward the Continuous Improvement of the Curriculum and Programs of Study*. Geraadpleegd op 27 november 2018 via <http://webcache.googleusercontent.com/search?q=cache:-wWyg97N9RYJ:www.cse.gouv.qc.ca/fichiers/documents/publications/CEBE/50-0199Summary.pdf+&cd=2&hl=en&ct=clnk&gl=nl&client=firefox-b>.
- De Boer, H., Bosker, R.J., & Van der Werf, M.P.C. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology*, 102(1), 168-179.
- De Graaff, R. & Bodde, M. (2006). Professionele talentdocenten blijven zich ontwikkelen. *Levende Talen Magazine*, 2009(6), 26-29.
- De Groot, A. (1966). *Vijven en zessen. Cijfers en beslissingen: het selectieproces in ons onderwijs*. Groningen: Wolters.
- DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education*, 64(4), 356-372.
- Dijkstra, J., Van der Vleuten, C.P.M., & Schuwirth, L.W.T. (2010). A new framework for designing programmes of assessment. *Advances in Health Sciences Education*, 15(3), 379-393.
- Dochy, F.J.R.C., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-350.
- Dronkers, J., Elffers, L., Karsten, S., & Van der Velden, R. (2016). Groeiende ongelijkheid vraagt om ingrijpen. *Socialisme en Democratie*, 73(2), 48-53.
- Expertgroep BKE/SKE (2013). *Verantwoord toetsen en beslissen in het hoger beroepsonderwijs*. Geraadpleegd op 28 november 2018 via <http://www.vsnu.nl/files/documenten/Domeinen/Onderwijs/Conferentie%20Examencommissies/IV%20achtergrondinformatie%20toetsbekwaamheid%20-%20Hanze.pdf>.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008). *Over de drempels met taal en rekenen*. Enschede: SLO.
- Faber, M., Faber, J.M., Visscher, A.J., & Schut, W.G.C. (2015). *Opbrengstgericht werken in het primair onderwijs: competenties, uitvoering en resultaten*. Enschede: Universiteit Twente.
- Feron, E., Schils, T. & Ter Weel, B. (2016). Does the teacher beat the test? The value of the teacher's assessment in predicting student ability. *The Economist*, 164(4), 391-418.
- Fisher, D., & Frey, N. (2009). Feed up, back, forward. *Educational Leadership*, 67(3), 20-25.
- Glas, C.A.W., Emons, W.H.M., & Berding-Oldersma, P.K. (2016). *Scenario's voor ijking van de eindtoetsen op de referentieniveaus*. Utrecht: Expertgroep PO.

- Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review*. Washington DC: National Council on Teacher Quality.
- Gulikers, L., & Baartman, J. (2016). *Doelgericht professionaliseren: formatieve toetspraktijken met effect! Wat DOET de docent in de klas?* Wageningen University & Research/Hogeschool Utrecht.
- Gulikers, J., Sluijsmans, D.M.A., & Baartman, L. (2008). The power of assessment in teacher education. In A. Swennen & M. van der Klink (Eds.), *Becoming a Teacher Educator. Theory and Practice for Novice Teacher Educators* (175-191). Dordrecht, Netherlands: Kluwer Academic Publishers Group.
- Hargreaves, A., Boyle, A., & Harris, A. (2014). *Uplifting leadership: How organizations, teams, and communities raise performance*. John Wiley & Sons.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- HBO-raad vereniging van hogescholen (2011). *Generieke Kennisbasis. Tweedegraads lerarenopleidingen*. Den Haag: HBO-raad, vereniging van hogescholen.
- Hollenberg, J., Keuning, J. & Meijer, J. (2016). *Gebruik van toetsen bij het plannen van leerroutes*. Amsterdam: Kohnstamm Instituut/Arnhem: Cito.
- Huygen, M. (2016). *Citotoets moet weer grotere rol gaan spelen bij schoolkeus*. Geraadpleegd op 6 december 2018 via <https://www.nrc.nl/nieuws/2016/04/14/citotoets-moet-weer-grotere-rol-gaan-spelen-bij-s-1611711-a324871>.
- Inspectie van het Onderwijs (2009). *Werken aan een beter rendement. Case study naar uitval en rendement in het hoger onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2010). *Opbrengstgericht werken in het basisonderwijs: Een onderzoek naar opbrengstgericht werken bij rekenen-wiskunde in het basisonderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2014). *De staat van het onderwijs 2013-2014*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2015a). *Opbrengstgericht werken aan taal en rekenen in het voortgezet onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2015b). *Verdere versterking: onderzoek naar het functioneren van examencommissies in het hoger onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2016a). *Taal en rekenen aan het einde van de basisschool*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2016b). *De kwaliteit van toetsing in het hoger onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2016c). *Recht van spreken: het functioneren van opleidingscommissies in het bekostigd hoger onderwijs in 2016*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2017). *De staat van het onderwijs 2015-2016*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2018a). *De staat van het onderwijs 2016-2017*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2018b). *Naar een nieuw resultatenmodel primair onderwijs*. Geraadpleegd op 6 december 2018 via <https://www.onderwijsinspectie.nl/onderwerpen/onderwijsresultaten-primair-onderwijs/naar-een-nieuw-onderwijsresultatenmodel>.
- Inspectie van het Onderwijs (2018c). *Onderzoekskader 2017 voor het toezicht op het middelbaar beroepsonderwijs*. Geraadpleegd op 6 december 2018 via <https://www.onderwijsinspectie.nl/documenten/rapporten/2018/07/13/onderzoekskader-2017-voor-het-toezicht-op-het-middelbaar-beroepsonderwijs>.
- Inspectie van het Onderwijs (2018d). *Peil.onderwijs*. Geraadpleegd op 6 december 2018 via <https://www.onderwijsinspectie.nl/onderwerpen/peil-onderwijs>.

- Inspectie van het Onderwijs (2018e). *Vershil SE-CE en de handhaving op de examenlicentie*. Geraadpleegd op 6 december 2018 via <https://www.onderwijsinspectie.nl/onderwerpen/onderwijsresultatenmodel-vo/nieuw-model/verschil-se-ce-en-de-handhaving-op-de-examenlicentie>.
- Inspectie van het Onderwijs (2018f). *VMBO Maastricht*. Geraadpleegd op 6 december 2018 via <https://www.onderwijsinspectie.nl/onderwerpen/vmbo-maastricht>.
- Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen*. Fontys publicatiebank.
- Karsten, S. (2016). *De hoofdstroom in de Nederlandse onderwijsdelta: een nuchtere balans van het mbo*. Apeldoorn: Garant.
- Kennispunt MBO Onderwijs en examinering (2017). *Kwaliteitsborging van examinering in de beroepspraktijk*.
- Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212-218.
- Krüger, M.L., & Andersen, I. (2017). *De lerende schoolleider: effecten van professionalisering*. Utrecht: VO-Academie/Den Haag: NRO.
- Lange, M.D., & Dronkers, J. (2007). *Groeide de ongelijkwaardigheid van het eindexamen tussen scholen verder in 2005? Discrepanties tussen de cijfers voor het schoolonderzoek en het centraal examen in het voortgezet onderwijs: 2005 versus 1997-2004*. Paper voor de Onderwijsresearchdagen 2007 in Groningen, 6-8 juni 2007.
- Larman, C., & Basili, V.R. (2003). Iterative and incremental developments. A brief history. *Computer*, 36(6), 47-56.
- Lenhard, W. & Schröppel, D. (2014). Prediction of academic performance prior to intersections within a multi-tiered school system. *Educational Research and Evaluation*, 20(6), 454-468.
- Lepper, M.R., Greene, D., & Nisbett, R.E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129-137.
- Lynch, R., & Hennessy, J. (2017). Learning to earn? The role of performance grades in higher education. *Studies in Higher Education*, 42(9), 1750-1763.
- Mandinach, E.B., & Gummer, E.S. (2016). *Data literacy for teachers: Making it count in teacher preparation and practice*. New York: Teachers College Press.
- McMillan, J.H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research & Evaluation*, 7(8).
- MBO Raad (2015). *Examenagenda mbo 2015-2020*. Woerden: MBO Raad.
- MBO Raad (2018). *Afspraken valide exameninstrumenten mbo*. Woerden: MBO Raad.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2015). *Stand van zaken invoering referentieniveaus taal en rekenen*. Brief Minister en Staatssecretaris OCW aan Voorzitter Tweede Kamer, 6 oktober 2015. Kamerstukken II 2015-2016, 31332, 52.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2017a). *Cijfer rekentoets telt niet meer mee voor het behalen van het diploma*. Brief Minister OCW aan Voorzitter Tweede Kamer, 19 december 2017. Kamerstukken II 2017-2018, 31332, 85.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2017b). *Toetsing en examinering in het voortgezet onderwijs*. Brief Minister OCW aan Voorzitter Tweede Kamer, 20 december 2017. Kamerstukken II 2017-2018, 31289, 357.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2018a). *Intrekking wetsvoorstel leerlingvolgsysteem en diagnostische tussentijdse toets voortgezet onderwijs*. Brief Minister OCW aan Voorzitter Tweede Kamer, 24 januari 2018. Kamerstukken II 2017-2018, 33661, 12.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2018b). *Start ontwikkelfase curriculumherziening primair en voortgezet onderwijs*. Brief Minister OCW aan Voorzitter Tweede Kamer, 23 februari 2018. Kamerstukken II 2017-2018, 31293, 386.

- Ministerie van Onderwijs, Cultuur en Wetenschap (2018c). *Toekomst van rekenen in het vo en mbo*. Brief Ministers OCW aan Voorzitter Tweede Kamer, 9 november 2018. Kamerstukken II 2017-2018, 31332, 88.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2018d). *Uitvoering regeerakkoord t.a.v. kleuter-toetsen*. Brief Minister OCW aan Voorzitter Tweede Kamer, 6 juli 2018. Kamerstukken II 2017-2018, 31293, 406.
- Moonen-van Loon, J., Overeem, K., Donkers, H.H., Van der Vleuten, C.P.M., & Driessen, E.W. (2013). Composite reliability of a workplace-based assessment toolbox for post-graduate medical education. *Advances in Health Sciences Education*, 18(5), 1087-1102.
- Nederlands-Vlaamse Accreditatieorganisatie (2016). *Beoordelingskader accreditatiestelsel hoger onderwijs Nederland*. NVAO: Den Haag.
- Newton, P.E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 14(2), 149-170.
- Newton, P.E. (2017). There is more to educational measurement than measuring: The importance of embracing purpose pluralism. *Educational Measurement: Issues and Practice*, 36(2), 5-15.
- Onderwijscoöperatie (2016). *Herijking bekwaamheidseisen*. Geraadpleegd op 28 november 2018 via <https://www.onderwijscooperatie.nl/wp-content/uploads/sites/2/2016/03/Verantwoording-Herijking-Bekwaamheidseisen.pdf>.
- Onderwijs Innovatie Groep (2013). *Effecten van sturing op discrepantie tussen de cijfers van het centraal examen en het schoolexamen bij de talen*. Utrecht: Vereniging van Leraren in Leven-de Talen.
- Onderwijsraad (2010a). *Een diploma van waarde*. Den Haag: Onderwijsraad.
- Onderwijsraad (2010b). *Ontwikkeling en ondersteuning van onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2011a). *Stimuleer innovatiekracht hoger onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2011b). *Een stevige basis voor iedere leerling*. Den Haag: Onderwijsraad.
- Onderwijsraad (2013). *Een smalle kijk op onderwijskwaliteit*. Den Haag: Onderwijsraad.
- Onderwijsraad (2014) *Toegevoegde waarde*. Den Haag: Onderwijsraad.
- Onderwijsraad (2015a). *Maatwerk binnen wettelijke kaders: eindtoetsing als ijkpunt voor het fundamenteel onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2015b). *Kwaliteit in het hoger onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2016a). *De volle breedte van onderwijskwaliteit*. Den Haag: Onderwijsraad.
- Onderwijsraad (2016b). *Een ander perspectief op professionele ruimte in het onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2017a). *Doordacht digitaal*. Den Haag: Onderwijsraad.
- Onderwijsraad (2017b). *De leerling centraal?* Den Haag: Onderwijsraad.
- Onderwijsraad (2017c). *Vluchtelingen en onderwijs. Naar een efficiëntere organisatie, betere toegankelijkheid en hogere kwaliteit*. Den Haag: Onderwijsraad.
- Onderwijsraad (2018a). *Briefadvies curriculumvernieuwing*. Den Haag: Onderwijsraad.
- Onderwijsraad (2018b). *Doorgeschoten differentiatie in het onderwijsstelsel*. Den Haag: Onderwijsraad.
- Organisation for Economic Co-operation and Development (2010). *Strong performers and successful reformers in education: lessons from PISA for the United States*. Paris: OECD.
- Organisation for Economic Co-operation and Development (2015). *OECD supports the assessment of learning outcomes in higher education*. Geraadpleegd op 6 december 2018 via <http://www.oecd.org/education/skills-beyond-school/ahelo-main-study.htm>.
- Popham, W. J. (2001). Teaching to the Test? *Educational Leadership*, 58(6), 16-21.
- Pulfrey, C., Darnon, C., & Butera, F. (2013). Autonomy and task performance: Explaining the impact of grades on intrinsic motivation. *Journal of Educational Psychology*, 105(1), 39-57.

- Rekers-Mombarg, L.T.M., Harms, G.J. & Van der Werf, M.P.C. (2009). Discrepanties tussen school-examen- en centraal examencijfers bij allochtone leerlingen. Omvang en verklaringen. *Pedagogische Studiën*, 86(6), 425-442.
- Rijksoverheid (2018). *Van Engelshoven stelt paal en perk aan bindend studieadvies*. Geraadpleegd op 6 december 2018 via <https://www.rijksoverheid.nl/actueel/nieuws/2018/09/03/van-engelshoven-stelt-paal-en-perk-aan-bindend-studieadvies>.
- Rouwhorst, Y., Sinke, D., Van Kessel, L., & Zondervan, E. (2017). *Sucstress: De stress om succes*. Utrecht: Interstedelijk Studenten Overleg.
- Scheerens, J. (2017). *Opportunity to Learn, Curriculum Alignment and Test Preparation*. Springer International Publishing.
- Scheerens, J., & Exalto, R. (2017). *Teaching to/from the test. Een verkennende studie naar het in lijn brengen van doelen, toetsen, curriculum en onderwijsaanbod*. Utrecht: Oberon.
- Schildkamp, K., Heitink, M.C., Van der Kleij, F., Hoogland, I., Dijkstra, A.M., Kippers, W.B., & Veldkamp, B.P. (2014). *Voorwaarden voor effectieve formatieve toetsing: een praktische review*. Enschede: Universiteit Twente
- Schönwetter, D.J., Sokal, L., Friesen, M., & Taylor, K.L. (2002). Teaching philosophies reconsidered: A conceptual model for the development and evaluation of teaching philosophy statements. *International Journal for Academic Development*, 7(1), 83-97.
- Schuwirth, L.W., & Van der Vleuten, C.P.M (2011). Programmatic assessment: from assessment of learning to assessment for learning. *Medical Teacher*, 33(6), 478-485.
- Sluijsmans, D., Joosten-ten Brinke, D., & Van Schilt-Mol, T. (2015). *Kwaliteit van toetsing onder de loep. Handvatten om de kwaliteit van toetsing in het hoger onderwijs te analyseren, verbeteren en borgen*. Antwerpen-Apeldoorn: Garant.
- Sluijsmans, D., Peeters, A., Jakobs, L., & Weijzen, S. (2012). De kwaliteit van toetsing onder de loep. *Onderwijsinnovatie*, 4, 17-25.
- Surssock, A., Smidt, H., & Davies, H. (2010). *Trends 2010: A decade of change in European Higher Education (Vol. 1)*. Brussels: European University Association.
- Teurlings, C., & Uerz, D. (2009). *Professionalisering van roc-docenten: zoeken naar verbinding*. Tilburg: IVA.
- Timmermans, A.C., Kuiper, H., & Van der Werf, M.P.C. (2013). *Schooladviezen en onderwijsloopbanen. Voorkomen, risicofactoren en gevolgen van onder- en overadvisering*. Groningen: GION.
- Trappenburg, M. (2011). Waarom het allemaal niet lukt. In J. van Tol, I. Helsloot en F. Mertens (red.), *Veiligheid boven alles? Essays over oorzaken en gevolgen van de risico-regelreflex* (37-53). Den Haag: Boom Lemma uitgevers.
- Van de Veen, E., Pol, M., & Moerkerke, G. (2018). De toetscommissie: borgen of zorgen? *Hoger Onderwijs Management*, 2018(2),10-14.
- Van de Werfhorst, H.G, Elffers, L., & Karsten, S. (2015). *Onderwijsstelsels vergeleken: leren, werken en burgerschap*. Amsterdam: University of Amsterdam.
- Van den Berg, F. (2018). *De examencommissie en de opleidingscommissie: samen op weg naar betere toetskwaliteit*. Poster gepresenteerd tijdens het SKE-symposium in Enschede op 9 juli 2018.
- Van den Bergh, H. , Rohde, E., & Zwarts, M. (2003). Is het ene examen het andere? Over de stabiliteit van schoolonderzoek en centraal examen. *Pedagogische Studiën*, 80(3), 176-191.
- Van der Vleuten, C.P.M., Norman, G.R., & Graaff, E. (1991). Pitfalls in the pursuit of objectivity: issues of reliability. *Medical Education*, 25(2), 110-118.
- Van der Vleuten, C.P.M., Schut, S. & Heeneman, S. (2018). Programmatisch toetsen als motor voor professioneel leren in het hoger onderwijs. In D. Sluijsmans, & M. Segers (red.), *Toetsrevolutie. Naar een feedbackcultuur in het hoger onderwijs* (124-136). Phronese.

- Van der Vleuten, C.P.M., & Schuwirth, L.W.T. (2005). Assessing professional competence: from methods to programmes. *Medical Education*, 39(3), 309-317.
- Van der Vleuten, C.P.M., Schuwirth, L.W.T., Driessen, E.W., Govaerts, M.J.B., & Heeneman, S. (2015). Twelve tips for programmatic assessment. *Medical Teacher*, 37(7), 641-646.
- Van Schilt-Mol, T., Sluijsmans, D., Peters, M., De Beer, F., & Jakobs, L. (2016). De touwtjes in handen: toetskwaliteit in het hbo. *Onderwijsinnovatie*, 2016(4), 15-24.
- Velon (2018). *Kennisbasis 2016-2019*. Geraadpleegd op 6 december 2018 via <http://www.lerarenopleider.nl/velon/kennisbasis/actualisering-kennisbasis/>.
- Verhoeven, B.H., Snellen-Balendong, H.A.M., Hay, I.T., Boon, J.M., Van Der Linde, M.J., Blitz-Lindeque, J.J., ... & Van der Vleuten, C.P.M. (2005). The versatility of progress testing assessed in an international context: a start for benchmarking global standardization? *Medical teacher*, 27(6), 514-520.
- VO Academie (2014). *Beroepsstandaard Schoolleiders VO*. Utrecht: VO-raad.
- Volman, M., Raban, A., Heemskerk, I., Ledoux, G., & Kuiper, E. (2018). *Toekomstgericht onderwijs*. Amsterdam: Kohnstamm Instituut.
- VO-raad (2015). *Diploma op maat. Ruimte voor talent in het voortgezet onderwijs*. Utrecht: VO-raad.
- Voor- en tegenstanders over de Cito kleutertoets (2018). Geraadpleegd op 6 december 2018 via <https://educatie-en-school.infonu.nl/diversen/120960-voor-en-tegenstanders-over-de-cito-kleutertoets.html>.
- White, R.W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, 66(5), 297-333.
- Wiliam, D. (2016). The secret of effective feedback. *Educational Leadership*, 73(7), 10-15.
- Yeh, S. (2010). The cost effectiveness of 22 approaches for raising student achievement. *Journal of Education Finance*, 36(1), 38-75.

Geraadpleegde deskundigen

De heer A. Béguin	Cito
Mevrouw A. Blok	Cito
De heer. W. Broekkamp	Bonarius Bedrijven
De heer J. Bruin	Universiteit Leiden
De heer B. Bulder	Inspectie van het Onderwijs
De heer R. Diris	Maastricht University
De heer A. Flierman	Nederlandse-Vlaamse Accreditatieorganisatie
Mevrouw A. Gielen	Landelijk Actie Komitee Scholieren
Mevrouw C. Goulmy	Aventuslyceum
Mevrouw J. Gulikers	Wageningen UR
De heer R. Hageman	Nederlandse-Vlaamse Accreditatieorganisatie
De heer T. Heemskerk	Ministerie van Onderwijs, Cultuur en Wetenschap
De heer F. Hooglandt	Inspectie van het Onderwijs
De heer H. Huizenga	Stenden Hotel Management School
De heer J. Kastelein	College voor Toetsen en Examens
Mevrouw R. Korthals	Inspectie van het Onderwijs
De heer W. Kuiper	Stichting Leerplanontwikkeling
Mevrouw N. Lafeber	MBO Raad
De heer M. Leegwater	Bonarius Bedrijven
De heer G. van Lonkhuyzen	College voor Toetsen en Examens
Mevrouw M. van der Lubbe	Inspectie van het Onderwijs
De heer P. Lucas	VO-raad
De heer S. Meershoek	EP-Nuffic
De heer P. van der Molen	Cito
De heer E. Mulder	Ministerie van Financiën
De heer F. Munk	Universiteit Utrecht
Mevrouw H. van Oostrom	VO-raad
Mevrouw M. Pol	Inspectie van het Onderwijs
De heer J. Raap	College voor Toetsen en Examens
De heer J. Rijpma	Hyperion Lyceum
De heer W. Rosier	Stichting Leerplanontwikkeling
De heer P. Sanders	Cito
Mevrouw T. Schils	Maastricht University
Mevrouw G. van Silfhout	Stichting Leerplanontwikkeling
De heer C. Sluijter	Cito
De heer J. Smits	Ministerie van Onderwijs, Cultuur en Wetenschap
De heer B. Teunis	PO-Raad
Mevrouw E. Verheijen	MBO Raad
Mevrouw M. Vogelzang	Inspectie van het Onderwijs
De heer D. van der Wateren	Filosofische praktijk De Verwondering
De heer B. Wilbrink	zelfstandig toetsdeskundige
De heer O. van Wilgenburg	EP-Nuffic
Mevrouw N. de Winter	Interstedelijk Studenten Overleg
Mevrouw A. de Wolff	Ministerie van Onderwijs, Cultuur en Wetenschap
Mevrouw S. Wools	Cito

Panel

Mevrouw H. Ackerman

Mevrouw A. van Beek

De heer B. Dankloff

Mevrouw J. Groenendijk

De heer T. Pijls

Mevrouw M. Straatemeier

De heer S. de Valk

De heer B. Veldkamp

De heer E. Welp

Mevrouw P. Willemse

Koninklijke Vereniging MKB-Nederland

Consortium Beroepsonderwijs

Stichting Examenkamer

Prove2Move

CINOP

Universiteit van Amsterdam

De GEU

Universiteit Twente

Kennisnet

Tilburg University

Sprekers op de Onderwijsdialoog

De heer J. Artz

Mevrouw J. Berckenkamp

Mevrouw A. Bergsma

Mevrouw L. Breek

De heer J. Drentje

Mevrouw K. Heij

Mevrouw R. van Schoonhoven

De heer F. Siero

Mevrouw D. Sluijsmans

Mevrouw E. Thier

ROC de Leijgraaf

Snijdersschool

Anna van Rijn College

Petrus Canisius College

Deltion Sprint Lyceum

zelfstandig toetsdeskundige

Actis Advies

Hogeschool Rotterdam

Zuyd Hogeschool

Stichting aan Zet

Onderwijs zonder toetsen en examens is nauwelijks denkbaar. Toetsresultaten geven leerlingen en docenten inzicht in het onderwijsleerproces, vormen de basis voor belangrijke beslissingen over leerlingen en studenten, en zijn van belang in het toezicht op de onderwijskwaliteit.

De huidige toets- en examenpraktijk is niet evenwichtig. Er wordt te weinig gebruikgemaakt van formatief toetsen, er worden te veel kwantitatieve toetsen gehanteerd, en er ligt veel nadruk op centrale vormgeving van toetsen en examens. Toetsen en examens dragen hierdoor niet optimaal bij aan de kwaliteit van het onderwijs.

De raad werkt twee aanbevelingen uit die meer balans in de toets- en examenpraktijk kunnen aanbrengen. Ten eerste roept hij de rijksoverheid en onderwijsinstellingen op om een heldere visie te formuleren. De rijksoverheid dient duidelijke kaders te bieden voor onderwijsinstellingen. Onderwijsinstellingen dienen de ruimte binnen de kaders te benutten vanuit keuzes op belangrijke toetsvraagstukken. Ten tweede adviseert de raad om de toetsdeskundigheid van onderwijsprofessionals te vergroten, zodat zij ook daadwerkelijk de ruimte die ze hebben, goed kunnen benutten.