

Monitor pilot diagnostische tussentijdse toets

bevindingen o-meting

MONITOR PILOT DIAGNOSTISCHE
TUSSENTIJDSE TOETS
bevindingen 0-meting

- eindrapport -

L. Bertling MSc.
Dr. M.C. Paulussen-Hoogeboom
Drs. J.J. van der Wel
Dr. B. Dekker

Regioplan
Jollemanhof 18
1019 GW Amsterdam
Tel.: +31 (0)20 – 531 53 15

Amsterdam, september 2015
Publicatienr. 15027

© 2015 RegioPlan, in opdracht van het ministerie van Onderwijs, Cultuur en Wetenschap
Het gebruik van cijfers en/of teksten als toelichting of ondersteuning in artikelen, scripties en boeken is toegestaan mits de bron duidelijk wordt vermeld.
Niets uit deze uitgave mag worden veeelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand en/of openbaar gemaakt in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of op enige andere manier zonder voorafgaande schriftelijke toestemming van RegioPlan.
RegioPlan aanvaardt geen aansprakelijkheid voor drukfouten en/of andere onvolkomenheden.

INHOUDSOPGAVE

1	Inleiding	1
1.1	De diagnostische tussentijdse toets	1
1.2	Vraagstelling.....	2
2	Vorbereidingen bij de monitor DTT	5
2.1	Bevindingen uit de literatuurstudie	5
2.2	Bevindingen uit de interviews met betrokkenen.....	8
2.3	Consequenties voor de monitoring.....	9
3	Bevindingen bestands- en deelnemersanalyse	11
3.1	Achtergrondkenmerken.....	11
3.2	Resultaatkenmerken.....	11
4	De bevindingen uit de monitor	13
4.1	Respondentkenmerken.....	14
4.2	Eerste ervaringen met de pilot	16
4.3	Bekendheid pilot onder niet-pilotscholen	19
4.4	Doelgroepen van de toets	19
4.5	Opbrengstgericht werken	21
4.6	Voorwaarden bij diagnostische toetsing	27
4.7	Mening over het nationaal onderwijsbeleid.....	31
5	Conclusies	33
Bijlagen	37
Bijlage 1	Literatuurlijst	39
Bijlage 2	Interviews	41

1 INLEIDING

In de afgelopen jaren is er beleidsmatig veel aandacht geweest voor de verbetering van de leerprestaties, niet alleen bij de zwakkere leerlingen, maar ook bij de potentieel goede leerlingen. Internationaal vergelijkend onderzoek laat zien dat het Nederlandse onderwijs gemiddeld goed scoort, maar ook dat de groep hoog presterende leerlingen in Nederland relatief klein is. Aan het eind van het vorige decennium kwamen daar nog aanwijzingen bij dat de onderwijsprestaties van de Nederlandse leerlingen omlaag dreigden te gaan.

In dezelfde periode werd steeds vaker aandacht gevraagd voor onderwijs dat beter aansluit bij de individuele leerlingen en waarbij meer gebruik wordt gemaakt van informatie uit de leeruitkomsten op individueel en groepsniveau, het zogenaamde 'opbrengstgericht werken' (OGW). Het idee hierbij is dat met opbrengstgericht werken de sterke en zwakke kanten van leerlingen, docenten, secties en de school als geheel in beeld worden gebracht en dat die informatie wordt gebruikt om het onderwijs bij te sturen. Afhankelijk van de analyse van de leeropbrengsten gaat het dan om meer onderwijs op maat, veranderingen in de inhoud of uitvoering van de lessen, aanvullende lesmethoden, bijscholing van de docenten, het meer betrekken van ouders bij het leerproces van de leerlingen et cetera. In het ideale geval ontstaat zo een 'lerende organisatie' waarin beschikbare informatie wordt geanalyseerd en leidt tot aanpassing en verbetering.

1.1 De diagnostische tussentijdse toets

De diagnostische toets als hulpmiddel

De overheid, het toezicht en de scholen hebben zich de afgelopen jaren ingezet voor kwaliteitsverbetering en voor een verdere ontwikkeling van opbrengstgericht werken, maatwerk voor leerlingen. Dit blijkt bijvoorbeeld uit de afspraken in de opeenvolgende sectorakkoorden vo, de prominente plaats in het toezichtskader en de belangstelling van scholen voor het onderdeel opbrengstgericht werken in het programma School aan Zet. De uitvoering in de praktijk is echter niet eenvoudig, zoals blijkt uit de behoefte aan ondersteuning en de constatering van de Inspectie in de recente onderwijsverslagen.

De beperkte beschikbaarheid van goede en betrouwbare informatie speelt hier ook mee. Examengegevens kunnen een hulpmiddel zijn, maar leveren alleen achteraf een beeld. Een tussentijdse toets eerder in de schoolcarrière biedt meer mogelijkheden voor bijsturing. De onderwijsraad adviseerde in 2011 een betrouwbare tussentijdse diagnostische toets te ontwikkelen en schreef daar onder meer over:

“De tussentijdse niveaubepaling moet waar nodig leiden tot een bijstelling van het onderwijsprogramma voor (individuele) leerlingen. Wanneer een leerling onder het verwachte niveau presteert, besteedt de docent extra aandacht aan het behalen van dit niveau; wanneer een leerling boven het niveau presteert, krijgt hij een uitdagender leerstofaanbod. [...] Scholen maken hiervoor bij voorkeur gebruik van een betrouwbare, gevalideerde toets. Belangrijk is dat een dergelijke toets geschikt is om het beheersingsniveau van alle leerlingen te bepalen, óók van de meer getalenteerde leerlingen. De toets moet niet alleen meten of leerlingen een bepaald niveau halen, maar tevens hoever ze boven of onder dat niveau zitten. Alleen dan levert een toets voldoende gegevens voor een school om verbeteringen door te kunnen voeren.”

Bron: Naar hogere leerprestaties in het voortgezet onderwijs. Onderwijsraad (2011)

De pilot diagnostische tussentijdse toets

Het idee voor een diagnostische tussentijdse toets (DTT) is overgenomen en uitgewerkt. Om scholen te ondersteunen is er nu, voor de kernvakken Nederlands, Engels en wiskunde, een pilot diagnostische tussentijdse toets gestart. Het uitgangspunt is dat de toets diagnostisch is, dat wil zeggen dat leerlingen en scholen er niet op ‘afgerekend’ worden, maar dat de toets ondersteunend is bij de vaststelling of een leerling maatwerk nodig heeft en zo ja, welk. Het is belangrijk dat de DTT aansluit bij de wensen van de schoolleiders en docenten en dat het daadwerkelijk als hulpmiddel kan worden ingezet om het onderwijs te verbeteren.

De pilot DTT loopt van september 2014 tot september 2017 en is in opzet geformuleerd als een vrijwillige samenwerking met scholen met wie wordt nagedacht over de opzet en inhoud van de DTT en door wie de toets getest en geïjkt wordt. Zo moet in de drie jaar die de pilot duurt een aantrekkelijke toets worden ontwikkeld die meerwaarde heeft voor leerlingen, ouders, verzorgers, docenten en scholen.

Een monitor- en evaluatieonderzoek maakt deel uit van de pilot. We gaan nu eerst in op het doel en de vraagstelling van dat onderzoek.

1.2 Vraagstelling

Het doel van de pilot DTT is om te komen tot een gedragen instrument dat scholen gebruiken om het onderwijs te verbeteren. Het monitor- en evaluatieonderzoek moet daarbij inzicht geven in de implementatie, het gebruik, de bruikbaarheid en de gevolgen van de DTT. Daarnaast moet ook duidelijk worden welke voorwaarden leiden tot een succesvolle inzet van de DTT en hoe bredere verspreiding ervan onder de scholen mogelijk is. Voor het monitor- en evaluatieonderzoek is de volgende set onderzoeks- en evaluatievragen geformuleerd.

Onderzoeksvragen

- 1. Welke scholen doen mee aan de DTT, welke niet en waarom?**
 - Welke kenmerken hebben scholen die wel/niet meedoen aan de pilot DTT?
 - Welke overwegingen hebben scholen om wel/niet mee te doen?
 - Hangt deelname samen met kenmerken en eerdere opbrengsten van scholen?
- 2. Wat is de werkwijze binnen de scholen en hoe verloopt het proces?**
 - Hoe gebruiken scholen de DTT?
 - Zijn er grote verschillen tussen scholen?
 - Wat gaat goed en wat kan beter (onder andere afname, gebruik, opvolging)?
- 3. Welke impact heeft DTT op de scholen?**
 - Wat is de bijdrage/invloed van de DTT op:
 - opbrengstgericht werken;
 - onderwijs op maat;
 - verbetercultuur;
 - onderwijsinhoud en werkvormen;
 - werkwijze/houding van leerlingen;
 - werkwijze/houding van docenten;
 - eventuele (onvoorziene) neveneffecten.
 - Is er een relatie tussen de impact en schoolkenmerken?
- 4. Voor welk type scholen, afdelingen, klassen en groepen leerlingen heeft de DTT de meeste relevantie en bruikbaarheid?**
 - Wat is de toegevoegde waarde van de DTT voor verschillende (groepen):
 - leerlingen;
 - docenten;
 - vakken;
 - scholen.
- 5. Welke randvoorwaarden bij scholen (schoolleiders, docenten) zijn bepalend voor een effectief gebruik van de DTT?**
- 6. Wat zijn goede voorbeelden van het gebruik van de DTT ter verbetering van het onderwijs of bij het bieden van onderwijs op maat?**
- 7. Hoe kan gestimuleerd worden dat de DTT voor meer scholen, klassen, afdelingen en groepen leerlingen relevant en bruikbaar is?**

De nulmeting is uitgezet onder de deelnemers aan de eerste pretest van de diagnostische tussentijdse toets en onder een controlegroep van scholen die niet deelnemen. Met de nulmeting kunnen we nog geen antwoord geven op al deze vragen. Voor de beantwoording van de vragen over de impact van de DTT is bijvoorbeeld nog te weinig informatie beschikbaar. Daarnaast moeten we ons realiseren dat de toets in deze fase nog in ontwikkeling is.

In dit conceptrapport doen we verslag van de literatuurstudie, de interviews met deskundigen en de bevindingen uit de bestandsanalyse waarin is gekeken in hoeverre pilotdeelnemers afwijken van niet-pilotdeelnemers. Ook gaan we in op de resultaten van het monitoronderzoek dat onder deelnemers (*pilot-scholen*) en niet-deelnemers is uitgezet. Na de zomervakantie is een aantal aanvullende interviews met schoolleiders (van pilot- en niet-pilotscholen) en docenten (van pilotscholen) uitgevoerd. De bevindingen uit die interviews zijn in deze versie van de rapportage toegevoegd.

2 VOORBEREIDINGEN BIJ DE MONITOR DTT

In de literatuurstudie en interviews zochten we onder meer naar randvoorwaarden die op scholen (bij schoolleiders, docenten) bepalend zijn voor een effectief gebruik van de diagnostische tussentijdse toets (DTT) in een verbetercyclus op school. Het doel van deze fase was om zo specifiek mogelijke input te verkrijgen voor het opstellen van de vragenlijsten. Daarbij zochten we onder meer naar gevalideerde instrumenten en schalen, en voorbeelden van uitkomstmaten.

2.1 Bevindingen uit de literatuurstudie

In deze paragraaf lichten we de bevindingen van de literatuurstudie kort toe. Als uitgangspunt voor de literatuurstudie is een bestaand overzicht gebruikt met relevante bronnen. Via de sneeuwbalmethode is verder gezocht naar bronnen die relevant leken voor het onderzoeken en zijn aanvullende studies gezocht via internet. Verder is gekeken naar aanvullende bronnen die naar voren kwamen uit de interviews met betrokkenen.

2.1.1 Randvoorwaarden

De meeste studies richten zich op datagestuurde werken, ook wel onderzoeksmatig werken genoemd. Veel informatie op dit gebied is afkomstig uit buitenlandse studies. In hun review onderscheiden Schildkamp e.a. (2014) drie vormen van formatieve toetsing: opbrengstgericht werken, assessment for learning en diagnostisch toetsen. Formatieve toetsing krijgt redelijk veel aandacht, de meeste studies richten zich echter op de eerste twee vormen. De auteurs vonden weinig relevante artikelen over diagnostisch toetsen. Het merendeel van de randvoorwaarden die we hebben gevonden, hebben niet specifiek betrekking op diagnostisch toetsen. Bij het selecteren van de randvoorwaarden is gekeken naar de relevantie en toepasbaarheid binnen het kader van diagnostisch toetsen.

In de literatuurstudie zijn in totaal ruim tachtig randvoorwaarden gevonden, die hier en daar wat overlap vertonen. De randvoorwaarden kunnen onderverdeeld worden in zes niveaus:

- randvoorwaarden op het niveau van de toets;
- randvoorwaarden op het niveau van de leerlingen;
- randvoorwaarden op het niveau van de docenten;
- randvoorwaarden op het niveau van de schoolleiders;
- randvoorwaarden die betrekking hebben op de context binnen de school;
- randvoorwaarden die betrekking hebben op de context buiten de school.

Toets

De randvoorwaarden op het niveau van de toets hebben betrekking op kenmerken van de toets, het toetsproces, kenmerken van de data die de toets oplevert en het datasysteem waarmee de data kan worden geanalyseerd. Bij kenmerken van de toets moet men denken aan doel en afstemming (bijvoorbeeld op het niveau van de leerlingen, het curriculum, landelijke standaarden). Voorbeelden van randvoorwaarden met betrekking tot het toetsproces zijn uniforme afname en een beperkte nakijklast. Bij kenmerken van de data die de toets oplevert, moet men denken aan de kwaliteit (bijvoorbeeld betrouwbaarheid en validiteit), accuratesse en relevantie. Randvoorwaarden met betrekking tot het datasysteem gaan over de toegankelijkheid (bijvoorbeeld data weergegeven in een begrijpelijk format), feedback en mogelijkheden voor monitoring.

Leerlingen

De randvoorwaarden op het niveau van de leerlingen hebben betrekking op een positieve houding, actieve betrokkenheid en vaardigheden die samenhangen met formatief toetsen.

Interessant in dit kader zijn de bevindingen uit de evaluatie onder een subgroep van bijna vierduizend leerlingen die in het voorjaar van 2015 aan de eerste pretest hebben deelgenomen (DUO e.a., 2015). Bevindingen uit deze evaluatie zijn:

- 48 procent van de responderende leerlingen heeft de wiskundetoets gemaakt, 42 procent de toets Nederlands en 36 procent de toets Engels; gemiddeld ongeveer 1,25 toets per leerling.
- Een klein deel van de leerlingen kampte met technische problemen (6%). Daarnaast geeft een klein deel van de leerlingen (7%) aan dat niet alle vragen het deden.
- Een derde van de leerlingen was van tevoren op de hoogte van de hulpmiddelen (vergrootglas en dergelijke). De hulpmiddelen worden gemiddeld door een zesde van de leerlingen gebruikt, buiten de rekenmachine die veel vaker wordt gebruikt (het gebruik is vermoedelijk vrijwel algemeen in de wiskundetoets). Leerlingen zijn overigens redelijk tevreden over de hulpmiddelen.

De pretest vraagt vrij veel van de leerlingen. Bijna de helft van de leerlingen had moeite om de aandacht er de hele toets bij te houden. Zeventig procent vindt dat de toets wat korter had mogen zijn.

Docenten

De randvoorwaarden op het niveau van de docenten hebben voornamelijk betrekking op houding en vaardigheden. Net als bij leerlingen is een positieve houding van docenten ten opzichte van de toets van belang. Daarnaast komen ook gevoel van '*self-efficacy*', bereidheid tot zelfreflectie, een '*interne locus of control*' en een onderzoekende houding naar voren als belangrijke voor-

waarden.¹ Bij vaardigheden kan men denken aan onderzoeksvaardigheden (zoals het analyseren en interpreteren van data), technologische vaardigheden (omgaan met het datasysteem) en didactische vaardigheden (zoals het begeleiden van leerlingen bij hun leren, bekend zijn met de leerstrategieën die leerlingen gebruiken en zinvolle feedback geven).

Naast houding en vaardigheden komen ook samenwerking tussen docenten en participatie in besluitvorming naar voren als belangrijke voorwaarden.

Schoolleiders

De randvoorwaarden op het niveau van de schoolleiders gaan enerzijds over het ondersteunen en faciliteren van docenten. Hierbij kan men denken aan het respecteren van de autonomie en professionaliteit, het bieden van intellectuele uitdagingen, het stimuleren van een onderzoekende houding, sturen op scholing en het beschikbaar maken van tijd en middelen. Anderzijds hebben deze randvoorwaarden ook betrekking op de vaardigheden van schoolleiders. Hierbij kan een onderscheid worden gemaakt tussen onderzoeksvaardigheden (zoals kennis en vaardigheden met betrekking tot datagebruik en een link kunnen leggen tussen onderwijs en onderzoek) en leiderschapsvaardigheden (zoals leidinggeven aan datateams, veranderingen in het toetsbeleid implementeren, uitkomsten van de data-analyse verbinden aan concrete acties, een cultuur van formatief toetsen creëren en de visie uitdragen). Tot slot is samenwerking tussen schoolleiders, docenten en administratief/ondersteunend personeel naar voren gekomen als een belangrijke randvoorwaarde voor effectief gebruik van formatieve toetsen.

Context binnen de school

Voorwaarden met betrekking tot de context binnen de school kunnen verdeeld worden in praktische voorwaarden en culturele voorwaarden. Voorbeelden van praktische voorwaarden zijn toegang tot methodes, hulpmiddelen en systemen, het vormen van data-teams, de aanwezigheid van coaches/experts die docenten ondersteunen, scholingsmogelijkheden voor zowel docenten als schoolleiders (afgestemd op de context waarin ze werken), en tijd en budget. Voorbeelden van culturele voorwaarden zijn een gezamenlijke visie, normen, doelen en verwachtingen (draagvlak), vertrouwen onder docenten (een niet-bedreigende context) en het datagestuurd werken/formatief toetsen zien als een gezamenlijke verantwoordelijkheid.

Context buiten de school

Bij randvoorwaarden met betrekking op de context buiten de school moet men denken aan netwerken met andere scholen, aandacht voor assessment bij de opleiding voor docenten, heldere benchmarks, aansluiting van het toezicht van de Inspectie en helder, consistent en coherent overheidsbeleid (rondom

¹ Onder self-efficacy (zelfeffectiviteit) verstaan we de mate waarin iemand erin vertrouwt dat hij of zij bekwaam is om taken succesvol uit te voeren. Als er sprake is van een interne locus of control gelooft iemand dat hij of zij zelf (mede) verantwoordelijk is voor het succes (of het falen) van zijn of haar pogingen om een doel te bereiken.

formatief toetsen en op het gebied van scholing en professionaliseringsmogelijkheden voor docenten).

Op basis van de gevonden randvoorwaarden onderscheiden we zeven thema's:

- attitude;
- samenwerking;
- data literacy (technische vaardigheden);
- faciliteiten, tijd en ruimte;
- uniformiteit van de toets;
- afstemming van de toets;
- uitkomsten gebruiken voor verbetering van de leerprestaties.

2.1.2 Schalen en uitkomstmaten

De meeste studies gaan over de effectiviteit en randvoorwaarden van datagestuurde werken. We hebben de indruk dat er nog weinig instrumenten bestaan die zich specifiek richten op het meten van deze randvoorwaarden. Het onderzoek van Krüger (2014) vormt hierop een uitzondering. In deze studie zijn er vragenlijsten afgenomen bij docenten om de invloed van onderzoeksmatig leidinggeven door de schoolleiding op het onderzoeksmatig handelen van docenten in kaart te brengen.

Na een interview met haar ontvingen we van mevrouw Schildkamp (Universiteit Twente) een vragenlijst die zij heeft gebruikt voor een onderzoek naar opbrengstgericht werken in het voortgezet onderwijs (2012). Een aantal van de stellingen in onze vragenlijst is aan deze vragenlijst ontleend.

Tijdens het bestuderen van de literatuur kwamen we enkele uitkomstmaten tegen. Met uitkomstmaten bedoelen we aanpassingen die worden doorgevoerd naar aanleiding van de toetsresultaten, om de leerprestaties van de leerlingen te verbeteren. Deze verbeteracties hebben bijvoorbeeld betrekking op de docent, de schoolleiding en het curriculum.

2.2 Bevindingen uit de interviews met betrokkenen

We hebben met vijf personen gesproken, waarvan twee interviews face to face plaatsvonden en andere telefonisch. Tijdens de interviews is gesproken over de randvoorwaarden die volgens de betreffende personen bepalend zijn voor de school om de diagnostische tussentijdse toets optimaal te kunnen gebruiken.

De indruk uit de interviews is dat de docent een cruciale rol speelt: hoe staat de docent tegenover de diagnostische toets en wat kan de docent vervolgens met de bevindingen uit de toets?

Bij de afname van de diagnostische toets moet de wens om onderwijs op maat te kunnen verzorgen het (dominante) uitgangspunt zijn. De docenten moeten erop kunnen vertrouwen dat de diagnostische toets om die reden wordt ingezet en niet bijvoorbeeld door de schoolleiding als kwaliteitsmaat wordt gebruikt. Verder moeten de docenten ervan uit kunnen gaan dat de toets ook daadwerkelijk een bijdrage kan leveren aan het vormgeven van dat maatwerk. Om onderwijs op maat te kunnen bieden, moet de docent inzicht hebben in wat er naar aanleiding van de toets kan worden gezegd, hoe dat overgebracht kan worden aan de leerlingen en hoe de uitkomsten vertaald kunnen worden naar een aanbod op maat.

Hierbij is het belangrijk dat de schoolleiding vanuit een langetermijnvisie werkt, draagvlak creëert binnen het team en onder de leerlingen, aan vertrouwen onder de docenten en leerlingen bouwt en de docenten faciliteert om hun kennis en vaardigheden te versterken en de analyses uit te kunnen voeren.

2.3 Consequenties voor de monitoring

De vragen van de monitor kunnen langs verschillende lijnen worden geordend. In onze verwerking van de literatuur sluiten we aan bij de ordening zoals de Universiteit Twente deze hanteert. In de eerste aanzet is deze indeling, zij het licht aangepast, bij de monitoring gevolgd en zijn vragen geformuleerd op het de zes niveaus die in paragraaf 2.1.1 zijn beschreven.

Op het moment dat de vragenlijst is ontwikkeld, liepen er verschillende onderzoeken naar de diagnostische toets (door Cito en CvTE onder andere) en waren er verschillende onderzoeken net afgerond. Een deel van de vragen wordt met deze onderzoeken al beantwoord. Zo richt de evaluatie van de pretest zich nadrukkelijk op het niveau van de toets en van de leerlingen (DUO e.a., 2015). In de eerste monitor hebben we, hoewel we niet per se zicht hadden op de uitkomsten van de andere onderzoeken, ervoor gekozen om overlap in informatieverzameling te voorkomen en daarom in de vragenlijsten geen vragen op te nemen over deze niveaus.

Bij de ontwikkeling van de vragenlijsten voor de monitor is zo goed mogelijk rekening gehouden met wat er al bekend is en welke kennis er met ander onderzoek kan worden verworven. Mede daarom hebben we ons in de eerste monitor geconcentreerd op de docenten, de schoolleiders en de context (met name binnen de school).

Voor het eerste monitoronderzoek zijn vragenlijsten ontwikkeld voor schoolleiders en docenten (Nederlands, Engels en wiskunde) van pilotscholen en niet-pilotscholen. In totaal zijn er vier verschillende vragenlijsten opgesteld.²

² Aan het begin van hoofdstuk 4 beschrijven we de onderdelen waaruit de vragenlijsten bestaan.

In de vragenlijst wordt onder meer een aantal stellingvragen voorgelegd (met name in de onderdelen 2 tot en met 4 en onderdeel 6). Deze vragen geven inzicht in de plaats die meetgestuurd, opbrengstgericht en gedifferentieerd werk momenteel binnen de school inneemt en de wijze waarop dat in de praktijk wordt ingevuld (onderdeel 4) en in de mate waarin de school voldoet aan voorwaarden die nodig zijn bij het diagnostisch toetsen (onderdeel 5). Bij dat laatste gaat het om houdingsaspecten, leiderschap, facilitering, vaardigheden en samenwerking.

Bij de stellingvragen hebben we als schaalengte 5-punts Likertschalen (plus een optie 'Ik weet het niet/geen mening') gebruikt. Antwoordmogelijkheden variëren hierbij, bijvoorbeeld, van 'Helemaal mee eens' tot 'Helemaal mee oneens'.

De bevindingen uit de in dit hoofdstuk beschreven activiteiten hebben hun weg gevonden naar de vragenlijsten die in juni 2015 aan de scholen zijn voorgelegd en waarvan we in deze conceptrapportage de eerste bevindingen presenteren. De vragenlijst voor schoolleiders van pilotscholen is als voorbeeld in een bijlage opgenomen.

3 BEVINDINGEN BESTANDS- EN DEELNEMERSANALYSE

Aan het eerste jaar van de pilot doen 100 vestigingen mee (van 85 unieke brinnummers). Op deze scholen (of schoollocaties) is in het voorjaar van 2015 een *pretest* afgenomen onder leerlingen. In het schooljaar 2014-2015 is ruim acht procent van het totale aantal vo-vestigingen bij de pilot betrokken³, komend schooljaar (2015-2016) neemt het aantal deelnemers toe tot ongeveer 250 (stand van zaken eind juli 2015).⁴

Met behulp van informatie over schoolkenmerken en opbrengsten⁵ hebben we, door beide populaties met elkaar te vergelijken, onderzocht of er sprake is van significante verschillen tussen vestigingen die in 2014-2015 zijn begonnen aan de pilot en vestigingen die dat (nog) niet doen.⁶

3.1 Achtergrondkenmerken

Wat betreft de regionale achtergrondkenmerken gemeentegrootte en stedelijkheid verschillen de pilotvestigingen niet van de niet-pilotvestigingen. Ook zijn pilotvestigingen niet vaker dan de niet-pilotvestigingen gevestigd in één van de vier grote steden of in een bepaalde provincie. Ook wat betreft denominatie treden er geen verschillen op.

Wanneer we kijken naar het leerlingaantal per vestiging, zien we dat de pilotvestigingen met gemiddeld 164 leerlingen wat groter zijn dan de niet-pilotvestigingen (gemiddeld 142 leerlingen).

3.2 Resultaatkenmerken

Wanneer we kijken naar het oordeel van de Inspectie over de leeropbrengsten van de vestigingen (zijn deze voldoende of onvoldoende) geldt voor alle schoolsoorten (vmbo-b, vmbo-k, vmbo-gt, havo en vwo) dat hiertussen geen verschillen bestaan tussen de pilotvestigingen en de niet-pilotvestigingen.

Ten slotte hebben we, per schoolsoort, onderzocht of de pilotvestigingen en de niet-pilotvestigingen van elkaar verschillen wat betreft de examenresultaten

³ Het totale aantal vestigingen in het vo bedraagt 1215.

⁴ Informatie afkomstig van het CvTE.

⁵ Het betreft hier de opbrengstoverzichten vo 2014 (ook wel Toezichtkaart 2014 genaamd). Deze bestanden betreffen het schooljaar 2012-2013.

⁶ Het aantal pilotvestigingen in de analyses bedraagt 97 (3 pilotvestigingen waren niet terug te vinden in de vestigingsbestanden van de Inspectie).

(schoolexamen en centraal examen) op de kernvakken. Bij het vmbo gaat het hier om de vakken Nederlands, Engels en wiskunde. Bij havo en vwo gaat het om de vakken Nederlandse taal & cultuur, Nederlandse taal, Engelse taal & cultuur, Engelse taal, wiskunde A, wiskunde B, wiskunde C en wiskunde D.

Tussen de pilotvestigingen en de niet-pilotvestigingen⁷ vonden we binnen geen van de vijf schoolsoorten significante verschillen in de examenresultaten op de kernvakken.

Op basis van deze analyses is onze indruk dat de eerste groep pilotscholen niet noemenswaardig afwijkt van de groep scholen die (nog) niet deelnemen aan de pilot.

⁷ Omdat er bij deze analyse veel toetsen werden uitgevoerd, is hier gekozen voor een strenger significantieniveau ($\alpha < 0,01$ in plaats van $\alpha < 0,05$).

4 DE BEVINDINGEN UIT DE MONITOR

Kort voor het einde van het schooljaar 2014-2015 zijn alle 100 locaties die in het voorjaar van 2015 hebben deelgenomen aan de pretest van de diagnostische tussentijdse toets (in de tekst ook wel pilotscholen genoemd) en een steekproef van ruim 600 scholen die niet deelnemen aan de pilot per e-mail uitgenodigd om mee te werken aan de monitor. De scholen ontvingen daarvoor twee links: één link naar een digitale vragenlijst voor schoolleiders en één link naar een vragenlijst voor docenten (van de secties Nederlands, Engels en wiskunde). Bij de pilotscholen is deze e-mail naar de contactpersonen gestuurd, bij de niet-pilotscholen naar een algemeen e-mailadres met het verzoek om de e-mail naar de juiste persoon door te sturen.

De vragenlijsten⁸ bestaan uit de volgende onderdelen:

- De relevantie van de diagnostische tussentijdse toets voor specifieke doelgroepen (*gelijk voor alle respondentgroepen*).
- De stand van het opbrengstgericht werken op school (*gelijk voor alle respondentgroepen*).
- De voorwaarden bij diagnostische toetsing (*stellingen in grote lijnen gelijk, maar toegespitst op de specifieke respondentgroepen*).
- De ervaringen tijdens de pretest (*alleen voorgelegd aan pilotscholen*).
- Enige stellingen over het nationale onderwijsbeleid (*gelijk voor alle respondentgroepen*).

De vragenlijsten beginnen met enige vragen die op de respondentgroep zijn toegesneden.

Naast de afname van de vragenlijst zijn aanvullende interviews gehouden met een aantal schoolleiders van pilotscholen en van niet-pilotscholen en enkele docenten van pilotscholen.⁹ Met deze interviews wilden we meer zicht krijgen op de keuzemotieven van de scholen, de praktijk van het opbrengstgericht werken en de deelname aan de pilot en de samenhang tussen de verschillende onderwerpen. De informatie uit deze gesprekken is in dit hoofdstuk in boxjes geplaatst.

⁸ In de bijlagen hebben we de vragenlijst voor schoolleiders van pilotscholen opgenomen als voorbeeld.

⁹ Er is gesproken met in totaal zeventien personen: tien schoolleiders van pilotscholen, drie schoolleiders van niet-pilotscholen en vier docenten van pilotscholen. De gesprekken duurden gemiddeld ongeveer een half uur.

4.1 Respondentkenmerken

Ruim 350 respondenten hebben gehoor gegeven aan onze uitnodiging.¹⁰ De respondenten zijn afkomstig van 185 scholen (of locaties). Van 35 scholen beschikken we over informatie van zowel de schoolleiders als de docenten.

Tabel 4.1 Verdeling volledige vragenlijsten over de scholen

	Unieke scholen	Zowel schoolleiders als docenten
Pilotscholen	51 scholen	21 scholen
Niet-pilotscholen	134 scholen	14 scholen

4.1.1 Schoolleiders

De vragenlijsten voor schoolleiders zijn door 89 respondenten ingevuld¹¹: 37 schoolleiders van pilotscholen die aan de pretest hebben deelgenomen en 52 schoolleiders van niet-pilotscholen.¹² De schoolleiders vertegenwoordigen samen in totaal 88 unieke schoollocaties.

Tabel 4.2 Responderende schoolleiders per schoolsoort

	pilotscholen (n=37)	niet-pilotscholen (n=52)
Vmbo	32%	33%
Havo/vwo	22%	31%
Zowel vmbo als havo/vwo	46%	37%

De diagnostische tussentijdse toets is ontwikkeld voor de vakken Nederlands, Engels en wiskunde. In tabel 4.3 geven we per schoolsoort weer welke afdelingen volgens de schoolleiders meedoen aan de pilot. Gemiddeld nemen per school 2,5 afdelingen ('schoolsoorten') deel aan de pilot. De afdelingen basis- en kaderberoepsgerichte leerweg zijn koploper met gemiddeld ruim twee vakken, de havo neemt de toets gemiddeld in 1,4 vakken af.

¹⁰ 82 procent van de personen die de vragenlijst hebben geopend, heeft deze ook volledig ingevuld.

¹¹ In de analyse zijn alle vragenlijsten meegenomen die tot ten minste de vragen over opbrengstgericht werken zijn ingevuld. 82 schoolleiders hebben de vragenlijst volledig ingevuld.

¹² In de definitieve rapportage presenteren we in een bijlage meer gegevens over de respons (en de consequenties daarvan).

Tabel 4.3 Participatie aan de pilot per schoolsoort, per vak (n=37)

	n	Nederlands	Engels	Wiskunde
Vmbo-basisberoepsgerichte leerweg	16	10	12	12
Vmbo-kaderberoepsgerichte leerweg	16	10	10	15
Vmbo-gemengde/theoretische leerweg	19	14	9	10
Havo	20	10	7	10
Vwo	22	9	12	13

4.1.2 Docenten

De beide docentenvragenlijsten zijn door in totaal 217 docenten ingevuld¹³: 48 docenten van pilotscholen die aan de pretest hebben deelgenomen en 169 docenten van niet-pilotscholen. Samen vertegenwoordigen ze 142 unieke schoollocaties, waaronder 37 scholen die aan de pretest hebben deelgenomen.

De docenten geven les in wiskunde (37%), Nederlands (31%) of Engels (29%). De respondenten van pilotscholen wijken hierbij nauwelijks af van de respondenten van niet-pilotscholen.¹⁴

De responderende docenten van pilotscholen geven significant vaker onderwijs aan het vmbo (tabel 4.6). Ze geven ook significant minder vaak les in de bovenbouw (tabel 4.7).

Tabel 4.6 Responderende docenten per schoolsoort

	pilotscholen (n=48)	niet-pilotscholen (n=169)
Vmbo	52%	30%
Havo/vwo	33%	53%
Zowel vmbo als havo/vwo	15%	17%

Tabel 4.7 Responderende docenten per bouw

	pilotscholen (n=48)	niet-pilotscholen (n=169)
Onderbouw	29%	24%
Bovenbouw	8%	27%
Zowel onderbouw als bovenbouw	63%	49%

¹³ Ook bij de docenten zijn in de analyse alle vragenlijsten meegenomen die tot ten minste de vragen over opbrengstgericht werken zijn ingevuld. Van deze 217 docenten hebben er 204 de vragenlijst volledig ingevuld.

¹⁴ Twee docenten geven les in een ander vak (Duits/Frans en natuurkunde).

4.2 Eerste ervaringen met de pilot

4.2.1 Motieven om aan de pilot mee te doen

Als belangrijkste redenen om mee te doen aan de pilot noemen schoolleiders:

- omdat deelname binnen de school een impuls aan opbrengstgericht werken kan geven (84% (zeer) belangrijk);
- omdat ze verwachten dat de diagnostische toets de standaard wordt (78%);
- om de leerlingen nu al meer maatwerk te kunnen bieden (59%);
- om als school invloed te kunnen uitoefenen tijdens de ontwikkeling van de toets (54%).

Ruim een kwart van de schoolleiders (27%) geeft aan dat tegenvallende resultaten op een of meer kernvakken een belangrijke reden voor deelname aan de pilot waren. Eén school merkt op dat ze meedoet omdat ze graag wil dat de belangen van het vso worden meegenomen tijdens de ontwikkeling van de toets.

Deelname aan de pilot is een zaak van de school zelf, het bestuur is zelden direct betrokken bij de beslissing van de school om aan de pilot mee te doen (tabel 4.4).

Tabel 4.4 Rol van het bestuur bij keuze voor deelname (n=37)

	Aantal
Bestuur hield zich afzijdig, maar is wel op de hoogte van deelname	18
Bestuur is niet op de hoogte van deelname aan de pilot	11
Bestuur en schoolleiding vallen samen	6
School heeft in samenspraak met bestuur besloten om aan pilot mee te doen	2

Aan docenten van de pilotscholen is gevraagd wat voor hen belangrijke redenen zijn om deel te nemen aan de pilot. Als belangrijkste worden gezien:

- om de leerlingen nu al meer maatwerk te kunnen bieden (83% (zeer) belangrijk);
- omdat deelname binnen de school een impuls aan opbrengstgericht werken kan geven (79%);
- omdat ze verwachten dat de diagnostische toets de standaard wordt (61%);
- om als school invloed uit te kunnen uitoefenen tijdens de ontwikkeling van de toets (52%).

Ten slotte noemt een relatief groot deel van de docenten (42%) mindere resultaten op (een van de) kernvakken een (zeer) belangrijke reden om mee te doen aan de pilot.

Uit de telefoongesprekken met betrokkenen blijkt dat doorgaans de schoolleiding het initiatief tot deelname aan de pilot neemt; docenten worden er niet of pas later bij betrokken. Bij de afweging die de schoolleiding maakt, speelt meestal een combinatie van motieven. Daarbij wordt er gekeken in hoeverre deelname aan de pilot aansluit bij de ontwikkelingen waar de school mee bezig is. Dat kan inhoudelijk zijn maar ook praktisch ('Hebben we er voldoende tijd voor?'). De motieven om aan de pilot deel te nemen, sluiten aan bij de uitkomsten uit de vragenlijst (zie hierboven). Verder geven verschillende respondenten aan als school ervaring op te willen doen met digitaal toetsen.

Naast de pretest nemen pilotscholen vaak ook andere toetsen af waarmee ze hun leerlingen kunnen diagnosticeren. Voorbeelden daarvan zijn Diataal en Cito-VAS. Verschillende pilotscholen nemen deel aan de pretest omdat ze een andere toets willen vervangen. De adaptiviteit van de DTT wordt daarbij als een belangrijke voordeel gezien.

De ene school kiest ervoor om rustig te beginnen, met de afname van de toets in één van de drie vakken binnen een beperkt aantal schoolsoorten; de andere doet vanaf het begin over de volle breedte mee. Bij de keuze om een specifieke toets af te nemen kan het draagvlak voor de toets binnen de sectie(s) meespelen, maar ook andere overwegingen kunnen een rol spelen. Zo vertelde een schoolleider uit het vmbo dat zijn afdeling voor de wiskundetoets koos om de leerlingen beter voor te bereiden op de rekentoets. De afnameduur van de toets (drie uur) is voor verschillende scholen reden geweest om bij de eerste pretest met minder toetsen mee te doen dan ze zich hadden voorgenomen.

4.2.2 Ervaringen met de pretest

Twee derde van de schoolleiders (68%) is tevreden over de wijze waarop de deelname aan het eerste pilotjaar is verlopen. De rest zegt doorgaans 'niet zo tevreden' te zijn. Ruim de helft van de schoolleiders ervaart dit eerste jaar knelpunten (tabel 4.5).

Tabel 4.5 Beleving van knelpunten door schoolleiders van pilotscholen (n=34)

	Aantal	Percentage
Schoolleider ervaart knelpunten	18	53%
Schoolleider ervaart geen knelpunten	15	43%
Weet niet/geen mening	1	3%

De meest genoemde knelpunten zijn:

- de tijd die de implementatie van de DTT kost;
- technische problemen tijdens de testafname;
- gebrek aan afnamefaciliteiten (computers, computerlokalen).

In de open vraag noemen verschillende schoolleiders het verschil tussen de verwachte en gerealiseerde duur van de toets als knelpunt.

Schoolleiders zijn het meest tevreden over de communicatie vanuit DUO (over Facet), maar de verschillen met de beoordeling van de communicatie vanuit CvTE en OCW zijn niet groot.

De docenten zijn matig tevreden over de wijze waarop de deelname aan het eerste pilotjaar is verlopen. Docenten van havo/vwo zijn significant minder vaak tevreden dan docenten van het vmbo. Docenten zijn ook minder tevreden dan de schoolleiders (zie 4.1.1). Een groot deel van de docenten heeft in dit eerste jaar knelpunten ervaren (tabel 4.8).

Tabel 4.8 Beleving van knelpunten door docenten van pilotscholen (n=48)

	Aantal	Percentage
Docent ervaart knelpunten	21	44%
Docent ervaart geen knelpunten	14	29%
Weet niet/geen mening	13	27%

Meest gehoorde knelpunten zijn:

- de tijd die de implementatie van de DTT kost;
- het draagvlak voor de DTT binnen de school;
- de communicatie over de DTT intern;
- de communicatie over de DTT extern.

De vier hierboven genoemde knelpunten zijn elk in totaal zeven keer genoemd.

Het meest tevreden zijn docenten over de communicatie vanuit CvTE, maar de verschillen met de beoordeling van de communicatie vanuit DUO en OCW zijn niet groot. Bovendien zijn er nogal wat docenten die aangeven geen mening te hebben over de communicatie door deze partijen.

De verwachtingen die respondenten hebben van de DTT zijn vaak hoog: meer zicht op wat leerlingen kunnen en een beter inzicht in wat de leerling nodig heeft. Uiteindelijk moet deelname aan de toets leiden tot verbeterde resultaten. Verschillende respondenten menen dat de toets (in de toekomst) behulpzaam kan zijn om het eigenaarschap van de leerlingen te vergroten. Schoolleiders en docenten van pilotscholen zien daarom uit naar de rapportages op leerlingniveau. Hoewel de DTT daar niet specifiek voor is bedoeld, zien enkele schoolleiders de toets ook als middel om te beoordelen hoe de school het doet ten opzichte van andere scholen.

De ervaringen met de DTT zijn nog erg pril. Bovendien is de intensiteit waarin de school heeft meegedaan aan de pretest soms erg laag. Een afdelingsleider vertelde bijvoorbeeld dat de resultaten van de toets niet zijn gedeeld met de betreffende sectie, onder meer omdat de leerlingen de toets niet erg serieus hebben genomen.

Ondanks de nog beperkte ervaring met de DTT geven schoolleiders tijdens de gesprekken als voorbeelden van 'zachte' effecten de ervaring die de school heeft opgedaan met digitaal toetsen op een grotere schaal en secties die vanwege de toets op eigen initiatief een ondersteuningstraject zijn aangegaan.

Hoewel de scores over het algemeen niet afwijken van wat de scholen weten uit andere toetsen (bijvoorbeeld Cito-VAS) viel het bij één school op dat haar groepen wat achterbleven op een onderdeel in de toets Engels. De sectie heeft dit opgepakt en gaat meer aandacht aan dat onderdeel besteden.

Uit de digitale vragenlijst krijgen we de indruk dat docenten van pilotscholen wat minder tevreden zijn over de deelname aan de pretest dan de schoolleiders van deze scholen. We hebben te weinig docenten gesproken om meer hierover te kunnen zeggen.

4.3 Bekendheid pilot onder niet-pilotscholen

Bijna driekwart van de schoolleiders van niet-pilotscholen zegt (redelijk) goed op de hoogte te zijn van de diagnostische tussentijdse toets, slechts een enkeling heeft er niet eerder van gehoord. Schoolleiders kennen de toets veelal (76%) omdat ze erover hebben gelezen in een publicatie van de overheid (CvTE, OCW, et cetera). Een derde van de schoolleiders kent de toets omdat de school heeft overwogen om deel te nemen aan de pilot.

Uit de gesprekken met enkele schoolleiders van niet-pilotscholen valt op dat deze, wanneer ze overwogen hebben om deel te nemen aan de pilot, dezelfde afwegingen maken als schoolleiders die hun school uiteindelijk wel hebben aangemeld voor de pilot. De intensiteit van andere ontwikkelingen binnen de school kan een reden zijn om niet aan de pilot deel te nemen. Ook niet-pilotscholen nemen toetsen af waarmee ze hun leerlingen kunnen diagnosticeren. Deelname aan de pilot kan ook daarom minder prioriteit binnen de school hebben.

Ook aan docenten van scholen die niet deelnemen aan de pilot is gevraagd in hoeverre ze bekend zijn met de diagnostische tussentijdse toets. Dertig procent van hen is (redelijk) goed op de hoogte. Ongeveer een derde van deze docenten (34%) is onbekend met de toets. Wanneer docenten ervan gehoord hebben, is dat meestal binnen de school (39%) of hebben ze erover gelezen (37%) in een publicatie van de overheid (CvTE, OCW, et cetera).

4.4 Doelgroepen van de toets

De diagnostische tussentijdse toets is bedoeld voor leerlingen in het laatste jaar van de onderbouw van het voortgezet onderwijs. Met de toets kunnen vanaf volgend jaar de kennis en vaardigheden van individuele leerlingen worden geanalyseerd. De diagnose die aan de hand van deze analyse kan worden gesteld, maakt het voor de docenten mogelijk om – samen met de leerlingen – keuzes te maken voor onderwijzen en leren op maat. Tijdens de pretest zijn de opbrengsten overigens nog alleen op groepsniveau teruggekoppeld.

De diagnostische tussentijdse toets kent geen specifieke doelgroep (buiten dat de toets wordt afgenomen in het laatste jaar van de onderbouw), de toets is adaptief en in principe geschikt voor leerlingen van uiteenlopend niveau. Dit betekent niet dat de respondenten geen doelgroep voor ogen kunnen hebben.

Vandaar dat hun gevraagd is hoe relevant de diagnostische toets volgens hen is voor verschillende groepen leerlingen.

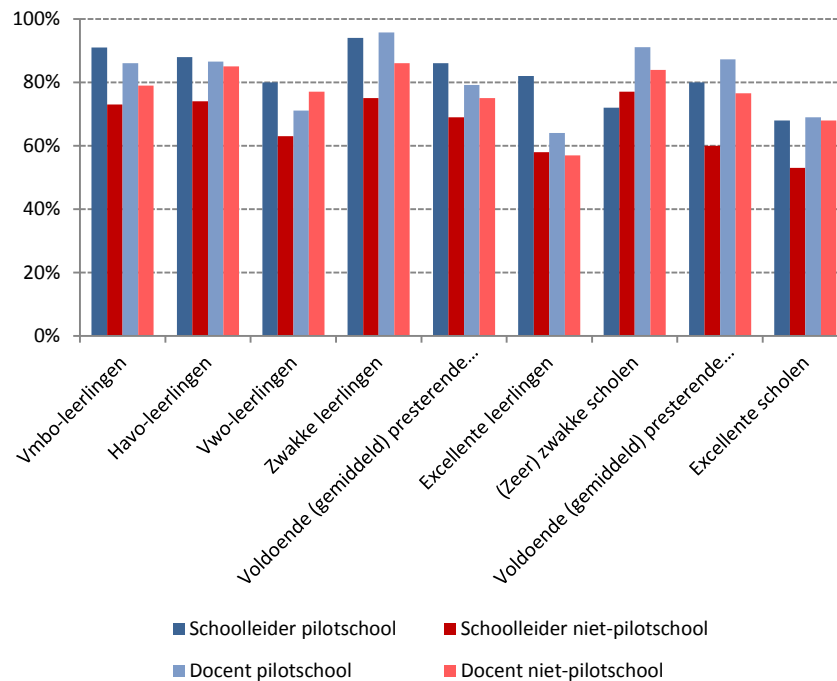
Het eerste dat opvalt aan de reacties van de respondentgroepen op deze vraag is het grote aantal personen dat 'Ik weet het niet/geen mening' invult. Een derde van de docenten geeft bij ten minste één van de voorgelegde doelgroepen aan geen antwoord te weten of geen mening te hebben. Met name bij de schoolsoorten (vmbo-, havo- of vwo-leerlingen) gebeurt dit vaak. Docenten van niet-pilotscholen vullen gemiddeld bij meer groepen 'Ik weet niet/geen mening' in dan docenten van pilotscholen.

Bij de schoolleiders heeft veertig procent ten minste éénmaal aangegeven niet te weten hoe relevant de toets voor doelgroepen is. Bij de schoolleiders gebeurt dat het vaakst bij de havo-leerlingen en bij de vwo-leerlingen.

In figuur 4.1 presenteren we per doelgroep het oordeel van de respondenten die hierover *wel* een uitspraak hebben gedaan. De verschillen tussen de respondentgroepen zijn doorgaans niet significant (te kleine respondentgroepen). De indruk is dat de respondenten het in belangrijke mate eens zijn dat de toets relevant is voor zwakke leerlingen en wat minder relevant is voor excellente scholen. Verder lijken docenten (van pilotscholen en niet-pilotscholen) meer op één lijn te zitten dan de schoolleiders, de verschillen tussen pilot- en niet-pilotscholen zijn bij de schoolleiders groter (en slechts bij enkele groepen significant¹⁵).

¹⁵ Significante verschillen zien we bij Zwakke leerlingen en Voldoende (gemiddeld) presterende scholen.

Figuur 4.1 Relevantie van de DTT voor verschillende groepen (alleen respondenten die een uitspraak hebben gedaan)



Hoewel respondenten relatief vaak 'Ik weet het niet/geen mening' als antwoord gaven, is de indruk dat schoolleiders en docenten bij de bepaling van de relevantie van de toets onderscheid maken tussen de verschillende groepen leerlingen. Een relatief klein deel van de respondenten (31% van de schoolleiders en 30% van de docenten) vindt dat de toets relevant is voor *alle* voorgestelde groepen leerlingen.

4.5 Opbrengstgericht werken

De diagnostische tussentijdse toets is een specifieke toepassing van opbrengstgericht werk. Onder opbrengstgericht werken verstaan we het *systematisch, doelgericht werken aan het maximaliseren van prestaties*. Volgens de Inspectie van het Onderwijs bestaat opbrengstgericht werken idealiter uit de volgende fasen:

*Het meten en analyseren van behaalde leerlingresultaten, het stellen van doelen op het gebied van de leerlingresultaten, het uitvoeren van maatregelen om de resultaten te verbeteren en de evaluatie daarvan.*¹⁶

Om de stand van het opbrengstgericht werken op de scholen in grote lijnen in kaart te brengen, zijn de in totaal negen aspecten, waarin de beide fasen door de Inspectie zijn onderverdeeld, als stelling voorgelegd aan de schoolleiders

¹⁶ Inspectie van het Onderwijs (2015). *De staat van het onderwijs. Onderwijsverslag 2013/2014*. Utrecht: Inspectie van het Onderwijs. Voor de negen deelaspecten en de beoordeling daarvan door de Inspectie, zie pagina 100.

en docenten. De respondenten is gevraagd aan te geven in welke mate het aspect opgaat voor de eigen school.

De antwoorden van de respondenten hebben we gehercodeerd naar een waarde, waarbij een hogere waarde aangeeft dat de stelling in sterkere mate wordt onderschreven. De antwoorden 'Ik weet het niet/geen mening' zijn hierbij buiten beschouwing gelaten.¹⁷ De bevindingen uit deze 'zelfevaluatie' presenteren we hieronder in een aantal figuren.

Schoolleiders over opbrengstgericht werken

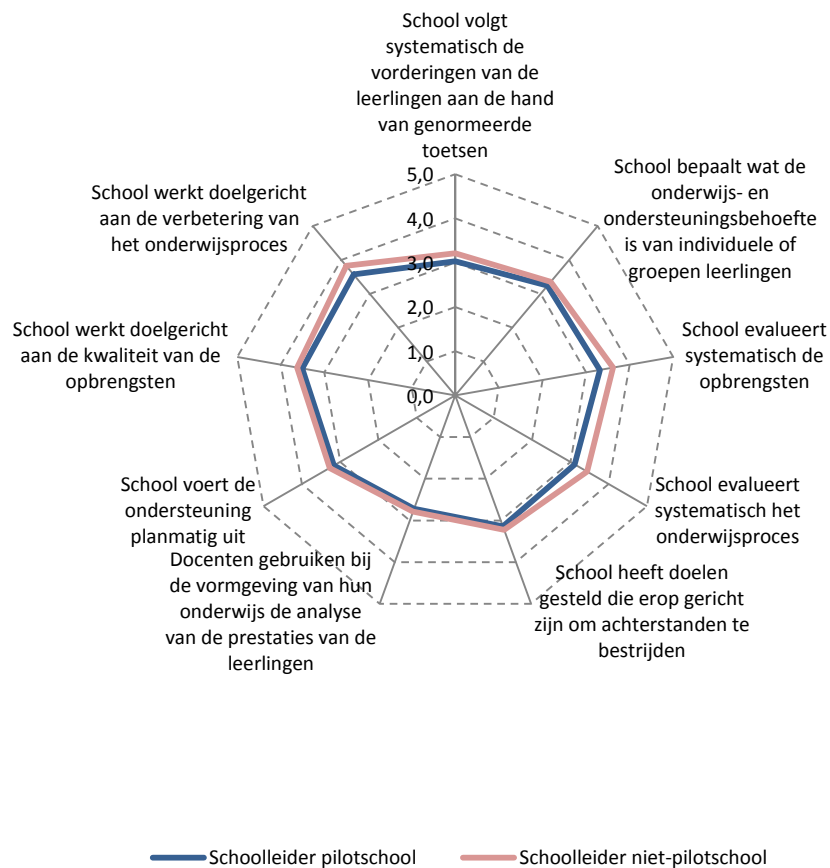
De sterkste punten van de school binnen het opbrengstgericht werken zijn volgens de schoolleiders dat de school:

- doelgericht werkt aan de kwaliteit van de opbrengsten;
- doelgericht werkt aan de verbetering van onderwijsproces;
- in iets mindere mate de opbrengsten systematisch evalueert (figuur 4.2).

Verschillen tussen pilotscholen en niet-pilotscholen zijn hierbij niet significant. Tussen de scores van de respondenten op deelaspecten zit samenhang: hoge scores op het ene deelaspect correleren met hoge scores op andere deelaspecten.

¹⁷ Deze antwoorden zijn behandeld als ontbrekende waarden (*missings*).

Figuur 4.2 Schoolleiders over de stand van het OGW op hun school*



* De maximale score is 5. Hoe hoger de waarde, hoe sterker de mate waarin de stelling wordt onderschreven.

Twee derde van de schoolleiders van pilotscholen meent dat binnen hun school hooguit de helft van de secties het opbrengstgericht werken voldoende op orde heeft. Bij de schoolleiders van niet-pilotscholen geldt dat voor 80 procent van de respondenten. Dat sluit aan bij de indruk die figuur 4.2 wekt: gemiddeld genomen menen schoolleiders van niet-pilotscholen dat hun school iets verder is met opbrengstgericht werken dan de schoolleiders van pilotscholen.

Schoolleiders, met name van de pilotscholen, menen dat er in het vmbo van hun school sterker opbrengstgericht wordt gewerkt dan in havo en vwo.

Een groot deel van de schoolleiders van pilotscholen (84%) gaf in de vragenlijst aan dat ze hopen dat deelname aan de pilot de invoering van opbrengstgericht werken binnen de school een impuls geeft (zie paragraaf 4.2.1). Op basis van de telefoongesprekken met schoolleiders en docenten zien we geen verschillen tussen pilot- en niet-pilotscholen in de mate waarin de scholen opbrengstgericht werken.

Tegenover scholen waar OGW behoorlijk ontwikkeld lijkt, staan scholen waar dat veel minder het geval is. Schoolleiders constateren, soms grote, verschillen tussen afdelingen en secties. OGW is voor vrijwel alle schoolleiders dan ook een belangrijk aandachtspunt. Verschillende schoolleiders menen dat individuele docenten niet beschikken over de kennis en vaardigheden om opbrengstgericht te werken.

Deelname aan de pilot DTT is niet de enige manier om zich op dit terrein te ontwikkelen. Tijdens onze gesprekken horen we bijvoorbeeld vanuit beide groepen scholen (pilot en niet-pilot) dat ze bezig zijn met (de implementatie van) RTTI.¹⁸ Daarnaast noemen scholen enkele andere instrumenten en programma's die als doel hebben het opbrengstgericht werken op school te versterken. Soms zijn kritische opmerkingen over de evaluatie van de opbrengsten door de onderwijsinspectie reden om aan de slag te gaan met OGW.

Een van de geïnterviewde schoolleiders (van een niet-pilotschool) gaf aan dat OGW geen doel is van de school. De school richt zich op het bieden van kwaliteit in de 'breedst mogelijke zin'.

Docenten over opbrengstgericht werken

Hoewel de verschillen zelden significant zijn, lijken de docenten van de pilotscholen de stand van het opbrengstgericht werken op hun school over de hele linie juist wat hoger in te schatten dan de docenten van niet-pilotscholen op hun school (figuur 4.3).¹⁹ Ook bij de docenten zien we een sterke samenhang tussen de scores op de deelaspecten.

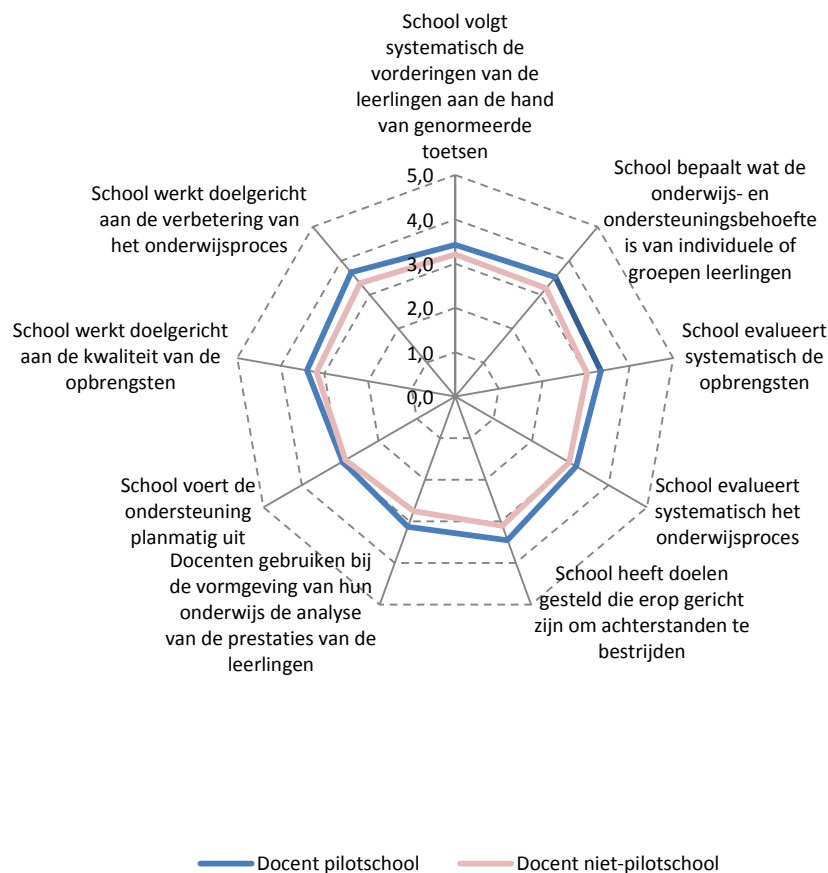
Docenten noemen als sterkste punten van de school binnen het opbrengstgericht werken:

- het doelgericht werken aan de verbetering van het onderwijsproces;
- het bepalen wat de onderwijs- en ondersteuningsbehoefte is van individuele of groepen leerlingen;
- het systematisch volgen van de vorderingen van de leerlingen aan de hand van genormeerde toetsen;
- het doelgericht werken aan de kwaliteit van de opbrengsten.

¹⁸ In de RTTI-systematiek worden op cognitief gebied vier niveaus onderscheiden: reproductie, toepassingsgericht niveau 1, toepassingsgericht niveau 2 en inzicht. De doelstelling is om zo meer zicht te krijgen op de leerontwikkeling van individuele leerlingen.

¹⁹ Bij de gemiddelde schaalscore is het verschil tussen docenten van pilotscholen en niet-pilotscholen duidelijk, maar net niet significant.

Figuur 4.3 Docenten over de stand van het OGW op hun school*



* De maximale score is 5. Hoe hoger de waarde, hoe sterker de mate waarin de stelling wordt onderschreven.

Opbrengstgericht werken op de pilotscholen

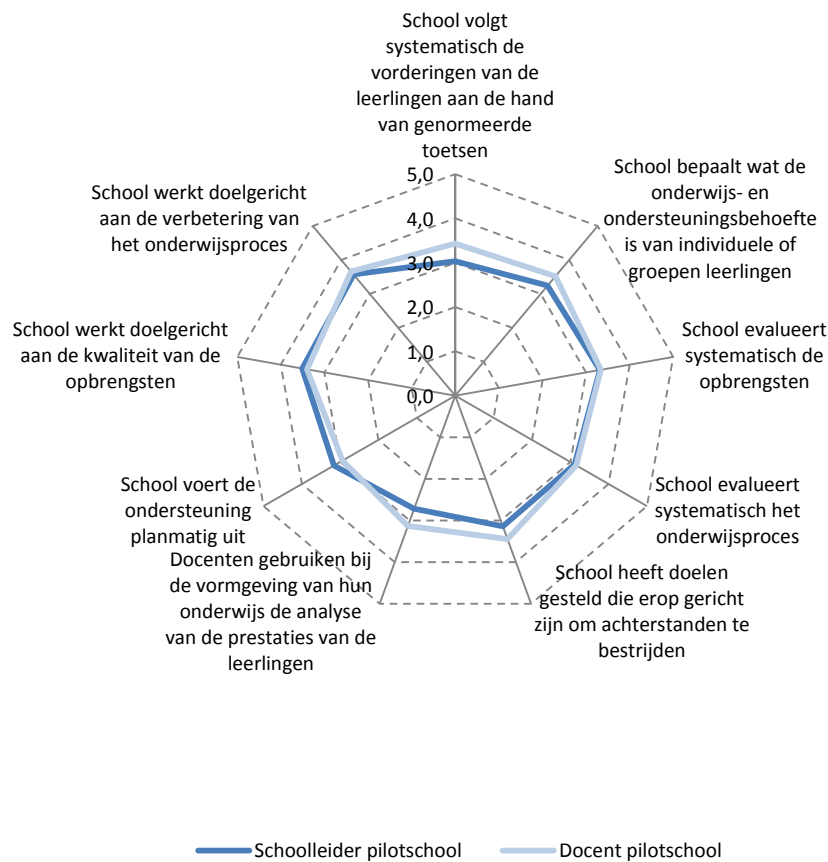
Docenten van pilotscholen lijken iets meer vertrouwen te hebben in de stand van het opbrengstgericht werken dan de schoolleiders (figuur 4.4). Dit geldt met name voor de deelaspecten waarbij ze zelf wat nauwer betrokken zijn:

- systematisch volgen van vorderingen van leerlingen;
- bepalen van onderwijs- en ondersteuningsbehoefte;
- het gebruiken van analyse van de prestaties bij de vormgeving van hun onderwijs.

Overigens zijn ook hier de verschillen niet significant.²⁰

²⁰ Bij de vergelijking van de scores van schoolleiders en docenten van niet-pilotscholen zien we dat schoolleiders significant positiever zijn over de stand van het OGW op hun school dan de docenten van deze scholen.

Figuur 4.4 Respondenten over de stand van het OGW op hun school*



* De maximale score is 5. Hoe hoger de waarde, hoe sterker de mate waarin de stelling wordt onderschreven.

Vergelijking met de bevindingen van de Inspectie

Overigens zijn volgens de Inspectie de sterke punten van de 125 scholen die zij in haar onderzoek in schooljaar 2013-2014 heeft meegenomen:

- het bepalen van de onderwijs- en ondersteuningsbehoefte van individuele of groepen leerlingen (97% van de onderzochte afdelingen voldoet hieraan in schooljaar 2013-2014);
- de systematische evaluatie van de opbrengsten (94% voldoet);
- de planmatige uitvoering van de ondersteuning (91% voldoet).²¹

Het percentage afdelingen waar volgens de Inspectie leraren gebruikmaken van de analyses van de prestaties van de leerlingen bij de vormgeving van hun onderwijs blijft hierbij duidelijk achter (slechts 27% van de onderzochte afdelingen voldoet hieraan in schooljaar 2013-2014). Anders dan bij de Inspectie springen deze aspecten er in de monitor in het kader van de pilot DTT niet uit; de scores in de monitor zijn gelijkmatiger en laten minder pieken en dalen zien.

²¹ Inspectie van het Onderwijs (2015): p. 100.

De Inspectie onderscheidt bij opbrengstgericht werken *Fase 1* ('Meten en analyseren', de eerste vier aspecten, beginnend met 'School volgt systematisch de vorderingen van de leerlingen aan de hand van genormeerde toetsen') en *Fase 2 en 3* ('Doelen stellen en uitvoering/inrichting', de vijf aspecten vanaf 'School heeft doelen gesteld die erop gericht zijn om achterstanden te bestrijden'). Van de respondenten in de monitor vindt 52 procent dat alle vier aspecten van Fase 1 in (ten minste redelijk) sterke mate opgaan voor hun school, 43 procent vindt dat van alle vijf aspecten van Fase 2 en 3.

Docenten van pilotscholen lijken wat meer overtuigd van de mate waarin hun school opbrengstgericht werkt dan schoolleiders van de pilotscholen, bij de niet-pilotscholen is dat precies omgekeerd.

4.6 Voorwaarden bij diagnostische toetsing

Aan alle respondenten is een reeks stellingen voorgelegd die betrekking hebben op mogelijke randvoorwaarden bij diagnostisch toetsen. Ze konden aangeven in hoeverre ze het met deze stelling eens waren. Opnieuw zijn de antwoorden gehercodeerd naar een waarde waarbij een hogere waarde aangeeft dat de stelling in sterkere mate wordt onderschreven. Het resultaat presenteren we in tabel 4.9.

De bevindingen in hoofdlijnen zijn:

- *Houding tegenover DTT*
Respondenten die afkomstig zijn van de pilotscholen verwachten meer effecten van de DTT op opbrengstgericht werken en het leveren van maatwerk dan de respondenten van niet-pilotscholen. Schoolleiders zijn daarbij nog wat positiever dan docenten.
- *Leiderschap*
Voor zowel pilotscholen als niet-pilotscholen geldt dat schoolleiders de aanwezigheid van leiderschapsaspecten binnen de school in sterkere mate onderschrijven dan de docenten.
- *Facilitering*
Verschillen tussen schoolleiders en docenten bij pilotscholen lijken bij dit onderdeel wat minder groot dan tussen schoolleiders en docenten van niet-pilotscholen. De verklaring hiervoor ligt deels in de verschillende formulering van de stellingen (naar DTT of naar OGW). Daarnaast is het bijvoorbeeld mogelijk dat docenten van de pilotscholen door de deelname aan de pilot daarvoor worden gefaciliteerd.
- *Kennis en vaardigheden*
De scores liggen voor alle respondentgroepen (pilotscholen/niet-pilotscholen; schoolleiders/docenten) dicht bij elkaar. Het vertrouwen in de vaardigheden van de docenten is binnen de verschillende respondentgroepen ongeveer even sterk.

- *Samenwerking*
Bij de mate waarin resultaten worden besproken (met docenten, met leerlingen) scoren de respondenten van pilotscholen duidelijk lager dan de respondenten van niet-pilotscholen. Dit lijkt toe te schrijven aan de formulering van de laatste twee stellingen onder dit onderdeel. In de vragenlijst voor respondenten van de pilotscholen is hierin een relatie gelegd met de DTT en niet met het wat bredere OGW zoals in de vragenlijst voor respondenten van niet-pilotscholen. Tijdens de eerste pretest zijn er bijvoorbeeld nog geen resultaten op leerlingniveau beschikbaar die met de leerlingen kunnen worden besproken. De lage score op dat aspect bij de docenten van pilotscholen is daarom geen verrassing.

In algemene zin lijkt de beantwoording van de stellingvragen over de diagnostische tussentijdse toets door de respondenten van de pilotscholen nog erg gekleurd door de prille, maar ook nog wat wisselende ervaringen met de toets tijdens de pretest waardoor het, zeker in deze 0-meting, lastig is om in de pilotscholen en de niet-pilotscholen met elkaar te vergelijken.

Tabel 4.9 Voorwaarden bij diagnostische toetsing*

	Schoolleider pilotschool	Schoolleider niet-pilotschool	Docent pilotschool	Docent niet-pilotschool
Houding tegen over DTT				
De DTT draagt bij aan opbrengstgericht werken#	4,15	3,47	3,98	3,62
De DTT maakt meer gepersonaliseerd onderwijs aan leerlingen mogelijk#	4,00	3,32	3,72	3,51
De DTT draagt bij aan de ontwikkeling van een verbetercultuur op de school#	4,09	3,20	3,83	3,39
De DTT is vooral bedoeld om (groepen van) leerlingen verder te helpen#	3,82	3,23	3,87	3,58
Leiderschapsaspecten				
Schoolleiding maakt op een goede wijze gebruik van verschillende datagegevens	3,79	3,96	3,53	3,24
Schoolleiding moedigt het gebruik van de DTT/OGW actief aan binnen de school	4,18	3,88	3,80	3,52
Schoolleiding bespreekt data met de docenten	3,74	4,04	3,11	3,35
Visie op het gebruik van de DTT/OGW heeft meer dan voldoende draagvlak binnen de school	3,41	3,74	3,11	3,24
Facilitering				
Docenten krijgen voldoende tijd om zich te bekwamen in het analyseren en gebruiken van de data (bijvoorbeeld uit de DTT)	3,24	3,42	2,78	2,74
Binnen de school is iemand aanwezig die de docenten kan ondersteunen bij de analyse en het gebruik van data (bijvoorbeeld uit de DTT)	3,55	3,69	3,14	2,97
Schoolleiding roostert specifiek tijd in voor datagebruik (bijvoorbeeld uit de DTT)	2,44	2,54	2,23	1,97

* De maximale score is 5. Hoe hoger de waarde, hoe sterker de mate waarin de stelling wordt onderschreven.

Stelling is voor alle respondentgroepen gelijk. Bij alle andere (niet gemarkeerde) stellingen is de formulering aangepast aan functie en/of aan deelname.

Vervolg Tabel 4.9*

	Schoolleider pilotschool	Schoolleider niet-pilotschool	Docent pilotschool	Docent niet- pilotschool
Kennis en vaardigheden				
Docenten beschikken, gemiddeld genomen, over de vaardigheden om de opbrengsten van en resultaten van toetsen te kunnen analyseren	3,48	3,45	3,41	3,59
Docenten zijn, gemiddeld genomen, in staat om met de data uit de toets de individuele leerbehoeften van de leerlingen vast te stellen	3,38	3,45	3,42	3,64
Docenten hebben, gemiddeld genomen, voldoende ervaring met de toepassing van formatieve evaluaties in de les	3,09	3,09	3,11	3,24
Samenwerking				
Schoolleiding, docenten en leerlingen werken goed met elkaar samen#	3,76	4,06	3,89	3,56
Docenten delen en bespreken de resultaten van de DTT/toetsen met andere docenten	3,07	3,76	3,26	3,88
Docenten delen en bespreken resultaten van de DTT/toetsen met de leerlingen	2,64	3,88	2,90	4,02

* De maximale score is 5. Hoe hoger de waarde, hoe sterker de mate waarin de stelling wordt onderschreven.

Stelling is voor alle respondentgroepen gelijk. Bij alle andere (niet gemarkeerde) stellingen is de formulering aangepast aan functie en/of aan deelname.

4.7 Mening over het nationaal onderwijsbeleid

Aan het einde van de vragenlijst hebben we de respondenten gevraagd in hoeverre ze het eens zijn met een aantal stellingen over (de invloed op) het nationaal onderwijsbeleid. De respondenten kregen willekeurig tien stellingen of een vaste keuze van vijf van deze tien stellingen voorgelegd.

Ook hier hebben we de antwoorden gehercodeerd naar een waarde waarbij een hogere waarde aangeeft dat de stelling in sterkere mate wordt onderschreven.

De belangrijkste bevindingen zijn:

- Schoolleiders van pilotscholen menen dat ze een sterkere invloed hebben op het onderwijsbeleid en de ontwikkeling daarvan dan de andere respondentgroepen. Het verschil dat we zien tussen schoolleiders van pilotscholen en niet-pilotschoolleiders treedt niet op bij de docenten.
- Docenten worden, zeker in de ogen van de schoolleiders, betrokken bij de ontwikkeling en implementatie van het onderwijsbeleid op school. Docenten van niet-pilotscholen lijken hiervan wat minder overtuigd dan docenten van pilot-scholen.
- Schoolleiders en docenten van pilotscholen zijn positiever over de langetermijneffecten van het huidige onderwijsbeleid dan hun collega's op niet-pilotscholen. Dit geldt in mindere mate voor de effecten op leerlingniveau.

Bij nadere analyse van de beantwoording van deze stellingvragen zien we een correlatie tussen onder meer:

- de eerste twee stellingen onderling (eens met 'Invloed op vorming van nationaal onderwijsbeleid is beperkt');
- de eerste twee stellingen negatief met de vijfde stelling (oneens met 'Vrijheid om te bepalen hoe met onderwijsbeleid om te gaan') en negatief met de laatste vier stellingen (oneens met 'Beleid heeft positief effect op onderwijs');
- de laatste vier stellingen onderling (eens met 'Beleid heeft positieve effect op het onderwijs');
- de laatste vier stellingen negatief met de eerste twee stellingen (oneens met 'Invloed op vorming van nationaal onderwijsbeleid is beperkt');
- de laatste drie stellingen met de vijfde stelling (eens met 'Vrijheid om te bepalen hoe met onderwijsbeleid om te gaan').

Tabel 4.10 Mening over het nationaal onderwijsbeleid*

	Schoolleider pilotschool	Schoolleider niet-pilotschool	Docent pilotschool	Docent niet-pilotschool
Schoolleiders en docenten hebben volgens mij veel te weinig macht om onderwijsbeleid te beïnvloeden	2,88	3,21	3,62	3,69
Schoolleiders en docenten kunnen de ontwikkeling van onderwijsbeleid op nationaal niveau (minister en staatssecretaris, ministerie van Onderwijs, Tweede Kamer) helemaal niet beïnvloeden#	2,72	3,19	3,53	3,72
Op mijn school denken docenten – via werkgroepen of vergaderingen – mee over de uitwerking van onderwijsbeleid#	4,00	4,15	3,71	3,62
Bij de invoering van onderwijsbeleid wordt er op mijn school helemaal niet naar docenten geluisterd	1,82	1,72	2,00	2,41
Ik heb vrijheid om te bepalen hoe ik met onderwijsbeleid omga#	3,42	3,31	3,15	3,11
Als ik onderwijsbeleid in uitvoering breng, moet ik me aan strakke procedures houden	2,56	2,96	2,70	2,79
Over het algemeen denk ik dat het huidige onderwijsbeleid op de lange termijn bijdraagt aan verbetering van de onderwijskwaliteit.	3,19	2,92	3,04	2,66
Alles bij elkaar genomen denk ik dat het huidige onderwijsbeleid leidt tot verbetering van de onderwijskwaliteit#	3,19	2,88	3,16	2,69
Over het algemeen kan ik dankzij het huidige onderwijsbeleid de problemen van mijn leerlingen beter oplossen#	2,78	2,55	2,58	2,50
Over het algemeen zorgt het huidige onderwijsbeleid ervoor dat ik leerlingen efficiënter kan helpen	2,69	2,44	2,69	2,48

* De maximale score is 5. Hoe hoger de waarde, hoe sterker de mate waarin de stelling wordt onderschreven.

Stelling is aan alle respondenten voorgelegd.

5 CONCLUSIES

In dit hoofdstuk lopen we de onderwerpen die aan bod komen in de onderzoeksvragen na, voor zover we er uitspraken over kunnen doen. Vragen over de impact van de diagnostische tussentijdse toets kunnen we nog niet of nauwelijks beantwoorden, al is het maar omdat het voor de meeste aspecten te vroeg is om de effecten daarop te meten. Om die reden zijn op dit moment bijvoorbeeld ook nog geen goede voorbeelden uitgewerkt.

Deelname aan de pilot

In het voorjaar van 2015 hebben 100 vestigingen (van 85 unieke brinnummers) deelgenomen aan de eerste *pretest* van de diagnostische tussentijdse toets. Deze scholen zijn op een aantal kenmerken vergeleken met scholen die niet aan de pretest hebben deelgenomen. Daarbij komen geen noemenswaardige afwijkingen aan het licht.

In de vragenlijst noemen de schoolleiders van de pilotscholen als belangrijkste redenen voor deelname de impuls die deelname aan het opbrengstgericht werken in de school kan geven en de verwachting dat de toets de standaard wordt. Docenten van pilotscholen onderschrijven de eerstgenoemde reden, maar noemen de mogelijkheid die de pilot biedt om de leerlingen nu al maatwerk te kunnen leveren nog net even iets vaker. Eerdere (tegenvallende) resultaten worden door een kwart van de schoolleiders van pilotscholen genoemd als reden om aan de pilot mee te doen.

Onbekendheid met de diagnostische tussentijdse toets (en de pilot) lijkt geen reden dat scholen niet aan de pilot meedoen. Driekwart van de schoolleiders van scholen die niet aan de pilot meewerken, zegt (redelijk) goed op de hoogte te zijn van de toets. Een derde van deze schoolleiders heeft overwogen om de school voor de pilot aan te melden. Onder docenten van niet-pilotscholen is de bekendheid met de toets minder groot. Ongeveer een derde zegt (redelijk) goed op de hoogte te zijn. Nog eens een derde kent de toets niet.

De verwachting is dat in het schooljaar 2015-2016 ongeveer 250 vestigingen aan de tweede pretest meedoen. Ongeveer twintig procent van de vestigingen in het voortgezet onderwijs is dan bij de pilot betrokken.

Werkwijze binnen de pilotscholen

In het voorjaar is een eerste versie van de toets afgenomen onder leerlingen. Uit een evaluatie die door het CvTE is uitgezet onder (een deel van) deze leerlingen, maken we op dat bijna de helft van hen de wiskundetoets heeft gemaakt, 42 procent de toets Nederlands en 36 procent de toets Engels. De indruk uit de evaluatie is dat de eerste pretest vrij veel van de leerlingen heeft gevraagd. Een enkele keer was er sprake van technische problemen bij de afname van de toets, hetgeen ook door enkele respondenten van de pilot-

scholen in de vragenlijst is genoemd. In de vragenlijst wordt er verder op gewezen dat het een behoorlijke logistieke operatie is om de toets af te nemen.

Verschillende schoolleiders (53%) en docenten (44%) ervaren knelpunten tijdens dit eerste pilotjaar. Naast problemen met techniek en afnamefaciliteiten zijn dat de tijd die de afname van de toets kost en, met name genoemd door docenten, de interne en externe communicatie over de toets. Verschillende schoolleiders en docenten gaven aan het jammer te vinden dat de bevindingen van de toets nog niet op leerlingniveau konden worden geleverd. Van 'opvolging', bijvoorbeeld in de vorm van maatwerk op leerlingniveau, is daarom nog maar in beperkte mate sprake.

De genoemde knelpunten passen bij het stadium waarin de pilot verkeert.

De impact van de deelname aan de pilot

Wat betreft het gebruik van de diagnostische tussentijdse toets zitten de scholen nog in het beginstadium. De eerste honderd scholen hebben deelgenomen aan een eerste pretest van de toets die nog volop in ontwikkeling is. De mogelijkheden die de toets in de toekomst wellicht gaat bieden, kunnen nog niet volledig worden benut. Wanneer we over de impact van de toets op het onderwijs spreken, hebben we het in deze fase voornamelijk over de verwachte effecten van de toets.

In de monitor spreken schoolleiders en docenten de verwachting uit dat de toets bijdraagt aan de versterking van het opbrengstgericht werken binnen de school. Op basis van hun eigen perceptie van de mate waarin de school momenteel opbrengstgericht werkt, zit daarin ruimte voor verbetering. Zowel schoolleiders als docenten van pilotscholen geven aan dat het opbrengstgericht werken binnen hun school nog niet in de volle breedte geldt (zie de aspecten die de Inspectie hanteert). De schoolleiders zijn hierbij overigens nog wat kritischer dan de docenten. Tijdens interviews met schoolleiders van pilotscholen noemen enkele van hen 'zachte' effecten als bijvoorbeeld de ervaring die de school heeft opgedaan met digitaal toetsen op een grotere schaal en secties die naar aanleiding van de deelname aan de pilot op eigen initiatief een ondersteuningstraject zijn aangegaan.

Op dit moment beschikken we uit de 0-meting van deze monitor nog over onvoldoende informatie om al uitspraken te doen over de (verwachte) impact van de deelname van de toets op deelaspecten.

Relevantie en bruikbaarheid voor verschillende doelgroepen

Gezien het grote aantal respondenten dat aangeeft niet te weten of de toets relevant is voor een specifieke doelgroep van leerlingen, bestaat er veel onduidelijkheid over de relevantie van de toets voor de verschillende groepen leerlingen. Deze onduidelijkheid is sterker bij respondenten van niet-pilotscholen.

Opmerkelijk is dat met name bij de vraag naar de relevantie voor leerlingen van de voorgestelde schoolsoorten (vmbo, havo en vwo) veel respondenten kiezen voor de antwoordmogelijkheid 'Ik weet het niet/geen mening'. Gemiddeld doet bijna een kwart van de respondenten dat.

Krap een derde van de respondenten die wel een idee over de relevantie voor doelgroepen hebben, vindt dat de toets relevant is voor *alle* voorgestelde groepen leerlingen. Groepen waarover de meeste consensus is, zijn Zwakke leerlingen (sterke relevantie) en (leerlingen van) Excellente scholen (matige relevantie).

De diagnostische tussentijdse toets, in haar verschillende varianten, is een adaptieve toets. Dat wil zeggen dat het niveau van de vragen die aan de leerlingen worden voorgelegd, meebeweegt met het niveau van de beantwoording van de eerdere vragen door de leerling. Door deze aanpassing van het niveau van de vragen versterkt de toets de relevantie voor de verschillende doelgroepen. Uit de beantwoording van de vragen naar relevantie maken we op dat dit niet per se duidelijk is voor de respondenten.

Randvoorwaarden voor een effectief gebruik van de DTT

In de monitor is aan de respondenten een aantal stellingen voorgelegd die betrekking hebben op de randvoorwaarden. Thema's die daarbij worden aangesneden, zijn:

- houding tegenover DTT;
- leiderschap;
- facilitering;
- kennis en vaardigheden (van docenten);
- samenwerking.

Op dit moment kunnen we de mate waarin de respondenten de stellingen onderschrijven nog niet koppelen aan de effecten van de toets, omdat we nog geen effecten kunnen meten. Wel zien we dat de respondenten van pilot-scholen positiever zijn over de verwachte opbrengsten van de toets. De indruk is dat dit in nog wat sterkere mate geldt voor schoolleiders dan voor docenten van deze scholen.

BIJLAGEN

BIJLAGE 1

Literatuurlijst

- Brink, Sara (2015). *Notitie Evaluatie pilot DTT jaar 1 (2014-2015). Versie 1A*. Utrecht: CvTE.
- Cito (2012). *De diagnostische tussentijdse toets: Tussenstand in ontwikkeling*. Arnhem: Cito.
- CvTE (2014). *Publieksversie toetswijzer DTT voor Nederlands, Engels en Wiskunde*. Utrecht: CvTE.
- CvTE (2015). *Evaluatie pretest DTT kandidaten (concept)*. Utrecht: CvTE.
- DUO/DSO/Onderzoek (2015) *Facet pretest DTT feb 2015 (versie 9-3-2015)*. Groningen: DUO.
- Heemskerk, I., Verbeek, F., Kuiper, E., Oomens, M., Linden, J. van der, & Hilbink, E. (2014). *Opbrengstgericht werken in het voortgezet onderwijs: Beeld en beleid*. Amsterdam: Kohnstamm Instituut.
- Inspectie van het Onderwijs (2015). *De staat van het onderwijs. Onderwijsverslag 2013/2014*. Utrecht: Inspectie van het Onderwijs.
- Jansen, L., Maas, J., Muns, W., & Richters, J. (2013). *Van diagnose naar leeractiviteiten: Een casestudy. Aanbod en instructie op basis van de analyse van toetsresultaten*. Amersfoort: CPS Onderwijsontwikkeling en advies.
- Krüger, M.L. (2014). *De invloed van schoolleiderschap op het onderzoeksmatig handelen van leraren in veranderingsprocessen*. Amsterdam: Kenniscentrum Onderwijs en Opvoeding van de Hogeschool van Amsterdam.
- Ledoux, G., Blok, H., & Boogaard, M. (2009). *Opbrengstgericht werken: Over de waarde van meetgestuurd onderwijs*. Amsterdam: Kohnstamm Instituut.
- Onderwijsraad. (2011). *Naar hogere leerprestaties in het voortgezet onderwijs*.
- Ronka, D., Lachat, M., Slaughter, R., & Meltzer, J. (2008). Answering the questions that count [Elektronische versie]. *Educational Leadership*, 66(4), 18-24.
- Schildkamp, K., Heitink, M., Kleij, F. van der, Hoogland, I., Dijkstra, A., Kippers, W., & Veldkamp, B. (2014). *Voorwaarden voor effectieve formatieve toetsing: Een praktische review*. Enschede: Universiteit Twente.
- Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes and promoting and hindering factors [Elektronische versie]. *Teaching and teacher education*, 26, 482-496.
- Schouwstra, Sanneke (2015) *De ontwikkeling van de diagnostische tussentijdse toets: van ontwerp tot diagnostisch profiel (presentatie)*. Arnhem: Cito.
- Slavin, R.E., Cheung, A., Holmes, G., Madden, N., & Chamberlain, A. (2011). *Effects of a data-driven district reform model*. Baltimore: Success for all foundation.

- Stam, M. (2015). Diagnostische tussentijdse toets in het onderwijs: Niet méér maar slimmer toetsen. *Vantwaalf tot achttien*, 1, 30-32.
- Trimbos, Bas en Van der Hoeven, Monique (2015). *Opbrengstgericht werken met en zonder de DTT (presentatie)*. Enschede: SLO.
- Tweede Kamer der Staten-Generaal (2012-2013). 33 661 *Wijziging van de Wet op het voortgezet onderwijs, de Wet voortgezet onderwijs BES, de Wet college voor toetsen en examens alsmede de Wet op de expertisecentra in verband met onder meer de invoering in het voortgezet onderwijs van een leerlingvolgsysteem, een diagnostische tussentijdse toets en verplichte deelname aan internationaal vergelijkend onderzoek (leerlingvolgsysteem en diagnostische tussentijdse toets voortgezet onderwijs)*. Verkregen op 13 april, 2015, van <https://zoek.officielebekendmakingen.nl/kst-33661-3.html>
- Verbeek, F., Ledoux, G., & Glaudé, M. (m.m.v. H. Blok, J. Roeleveld & C. Felix) (2012). *Op weg naar opbrengstgericht leiderschap. Evaluatie van het project 'Versterking kwaliteit bestuur en management'*. Amsterdam: Kohnstamm Instituut.

BIJLAGE 2**Interviews**

Organisatie	Naam respondent	Functie respondent
Cito	Mw. Anne-Marie Anthonissen	Clustermanager Centrale Eindtoets PO / Diagnostische Tussentijdse Toets bij Cito
Cito	Mw. Sanneke Schouwstra	Senior onderzoeker (inhoudelijk projectleider)
CvTE	Dhr. Jacob Raap	Sectormanager PO/diagnostische tussentijdse toets VO
Universiteit Twente	Mw. dr. Kim Schildkamp	Universitair hoofddocent (projectleider)
SLO	Dhr. Bas Trimbos	Leerplanontwikkelaar moderne vreemde talen. Secretaris van de toetswijzercommissie Engels voor de Diagnostische tussentijdse toets (DTT)

Regioplan Beleidsonderzoek

Jollemanhof 18

1019 GW Amsterdam

T 020 531 531 5

E info@regioplan.nl

I www.regioplan.nl