



**rijksuniversiteit
groningen**

faculteit rechtsgeleerdheid

vakgroep bestuursrecht
& bestuurskunde

De selectiviteit, efficiency en effectiviteit van het risicogerichte toezicht op grond van de WOT

Januari 2015

1 Inleiding

1.1 Aanleiding

De Wet op het onderwijstoezicht (WOT) regelt de wijze waarop in Nederland toezicht wordt gehouden op de kwaliteit van het onderwijs. Dit toezicht wordt uitgeoefend door de Inspectie van het Onderwijs (hierna: de Inspectie). Dit toezicht staat in de praktijk bekend onder de noemer 'risicogericht toezicht'. Tijdens de behandeling in de Eerste Kamer van het wetsvoorstel tot wijziging van de WOT op 24 januari 2012 heeft de minister van OCW toegezegd in 2014 een evaluatie van het risicogerichte toezicht aan het parlement te sturen. Het onderzoek richt zich op twee aspecten, te weten de effectiviteit van het toezicht en de efficiency ervan. Het ultieme einddoel van het toezicht door de Inspectie is dat de kwaliteit van het onderwijs verbetert. Als dat gebeurt, is het toezicht effectief. De efficiency van het toezicht betreft de tijd, het geld en de menskracht die het toezicht de Inspectie kost én de toezichtslasten voor de scholen. Des te beperkter de kosten voor de Inspectie en de toezichtslasten voor de scholen zijn, des te efficiënter is het toezicht.

Kern van het toezicht is dat scholen worden beoordeeld op de kwaliteit van het onderwijs en dat dat oordeel openbaar wordt gemaakt. Scholen die niet aan de normen voldoen, worden gekwalificeerd als zwak of zeer zwak. Deze scholen krijgen maximaal twee jaar de tijd de kwaliteit van het onderwijs weer op een aanvaardbaar niveau te brengen. In deze verbeterfase voert de inspectie geïntensiveerd toezicht uit.

Sinds 2007 werkt de Inspectie met het systeem van risicogericht toezicht. Dat wil zeggen dat de aard en de omvang van het toezicht worden afgestemd op de risico's op een opleiding of school. Bij de risicoanalyse worden onder meer CITO-scores, examenresultaten, in-, door-, en uitstroomgegevens, jaarstukken, signalen (zoals klachten) en nalevingsaspecten geanalyseerd. Op grond van deze analyse wordt bepaald of een school risico's loopt ten aanzien van de kwaliteit van het onderwijs. Op basis hiervan worden het niveau en de intensiteit van het toezicht bepaald. Het toezicht is er op gericht te waarborgen dat de onderwijsprestaties van scholen zich op een voldoende niveau (gaan) bevinden. Dan geldt als arrangement het basistoezicht, hetgeen inhoudt dat een school eens in de vier jaar door de inspectie wordt bezocht.

Het risicogerichte toezicht spreekt het bestuur van scholen aan. De gedachte is dat het bestuur actief, gericht en effectief optreedt waar het gaat om de kwaliteit van het onderwijs. Het schoolbestuur wordt gecontroleerd door een intern toezichthouder en treedt in dialoog met de omgeving, waaronder ouders en andere direct belanghebbenden.

De Inspectie heeft onder de noemer 'Toezicht 2020' de ambitie om te komen tot een bredere kwaliteitsverbetering van het onderwijs door ook scholen die voldoende presteren, te stimuleren zich te ontwikkelen in de richting van excellentie.

Toen in de jaren negentig van de vorige eeuw bleek dat haar oudste zoon op de openbare Montessorischool in het centrum van Amsterdam een grote leerachterstand had opgelopen, ging Karina Schaapman de strijd aan met de gemeente Amsterdam om vergoeding van noodzakelijke bijlessen. Onderzoek toonde vervolgens aan dat meer kinderen op de betreffende school leerachterstanden hadden opgelopen. Kern van het probleem was dat de school, de gemeente en de Inspectie geen idee hadden van de gerealiseerde onderwijskwaliteit en de daarop toe te passen normen. Het geschil tussen de gemeente Amsterdam en Karina Schaapman werd uiteindelijk in hoger beroep in een vonnis van 26 mei 1999 van de rechtbank Amsterdam in het voordeel van Schaapman beslecht. Dat leidde ertoe dat meer ouders bijleskosten vergoed kregen en dat meer kinderen bijgespijkerd werden. In 2000 publiceerde Karina Schaapman een boek over de affaire: 'Schoolstrijd. Ouders op de bres voor beter onderwijs.' De casus leidde uiteindelijk tot ingrijpende veranderingen in het onderwijstoezicht en de regels over governance binnen het onderwijs.

1.4 Dit rapport

In hoofdstuk 2 wordt het model van het risicogerichte toezicht beschreven. Daarbij komen de verschillende toezichtskaders voor het primair en het voortgezet onderwijs aan de orde. In hoofdstuk 3 staat de zelfevaluatie van de Inspectie centraal. In hoofdstuk 4 komen we tot een evaluatie van het systeem van onderwijstoezicht. De rapportage sluit in hoofdstuk 5 af met een beschouwing waarin ook aandacht wordt besteed aan de ambities van de Inspectie met 'Toezicht 2020'.

raad van bestuur) staat centraal in het toezichtsmodel. Het bestuur is als enige bevoegd tot het nemen van maatregelen gericht op de kwaliteit van het onderwijs en is daarvoor dus ook verantwoordelijk. Het bestuur verantwoordt zich extern voor die kwaliteit. Dat betekent dat aan de belanghebbenden, maar ook aan de toezichthouder informatie over de onderwijskwaliteit wordt verstrekt. Dat het schoolbestuur zich extern moet verantwoorden leidt, zo wordt verondersteld, tot versterking van het interne toezicht en tot de ontwikkeling van een model voor kwaliteitszorg. De externe verantwoordingsinformatie wordt volgens het model door de Inspectie gevalideerd met behulp van onafhankelijke bronnen. Op grond daarvan wordt vervolgens de risicoanalyse uitgevoerd. Voor het beoordelen van scholen hanteert de Inspectie toezichtskaders die alle relevante aspecten van het te beoordelen onderwijs beogen te omvatten en die in overleg met het onderwijsveld tot stand zijn gekomen.

2.3 Toezichtskaders

2.3.1 Inleiding

De Inspectie stelt toezichtskaders vast om de eenheid van handelen van de inspecteurs op de werkvloer te bevorderen. Het voordeel van een geëxpliciteerd kader is ook dat de werkwijze van de Inspectie kenbaar is voor betrokkenen en belanghebbenden. In 2009 zijn de toezichtskaders PO en VO geactualiseerd. In 2009 is gekozen voor een geïntegreerd toezichtskader PO/VO. Het toezichtskader bevat twee elementen: de werkwijze die de Inspectie hanteert (het hoe) en de kwaliteitsaspecten en indicatoren (het waarderingskader of kernkader) waarop scholen worden beoordeeld (het wat). Het kernkader PO omvat op dit moment 54 indicatoren en het kernkader VO 46 indicatoren waar bij een schoolbezoek op gelet wordt.

Het toezichtskader PO/VO van 2009 is in 2011 aangepast (onder meer naar aanleiding van de in de Wet 'Goed onderwijs, goed bestuur' opgenomen bevoegdheid tot het geheel of gedeeltelijk beëindigen van de bekostiging van een school door de minister en de komende aanpassing van de standaard-verbetertermijn voor zeer zwakke scholen van twee naar 1 jaar). In juli 2012 zijn enkele technische wijzigingen doorgevoerd in het toezichtskader PO/VO. Dat kader is vanaf het schooljaar 2012/2013 gehanteerd.

Op 20 september 2012 is met name op het onderdeel waarderingskader een grondige aanpassing van het toezichtskader 2011 vastgesteld voor het VO.⁴ Vanaf 1 januari 2013 geldt het toezichtskader PO/VO 2012 voor het PO en vanaf datzelfde moment het toezichtskader VO 2013 voor het VO, behalve voor VO-scholen die voor 1 januari 2013 in een aangepast toezichtstraject zaten. Deze scholen worden nog volgens het geïntegreerde toezichtskader benaderd. Er is dus momenteel sprake van twee toezichtskaders, één voor PO en één voor VO.

De werkwijze van de Inspectie kent vier uitgangspunten:

- het toezicht kent een bestuursgerichte aanpak;
- het toezicht is risicogericht en op maat;
- het toezicht is er mede op gericht te voorkomen dat scholen zwak of zeer zwak worden;
- het toezicht is mede gericht op (programmatisch) handhaven.

Met de invoering van de WOT is een kernthema van het toezicht het aanspreken van het bevoegde gezag op zijn verantwoordelijkheid voor de onderwijskwaliteit. Dat betreft in eerste instantie de behaalde onderwijsresultaten en, pas bij onvoldoende resultaten, de verrichte inspanningen. Daarbij worden scholen niet meer belast dan nodig is (art. 4, lid 2 WOT).

⁴ Het waarderingskader wordt regelmatig op onderdelen aangepast. Zo werd op 5 september 2013 een beleidsregel van de Inspecteur-generaal van het Onderwijs, nr. 4160445, gepubliceerd, houdende het instellen van een beslisregel bij het Toezichtkader VO 2013 en het van toepassing verklaren van het Toezichtkader VO 2013 op scholen als bedoeld in artikel 56 Wet op het voortgezet onderwijs. Op 19 december 2013 is het waarderingskader opnieuw op onderdelen aangepast.

Behalve bezoeken aan scholen die op basis van een analyse van de opbrengsten (zeer) zwak dreigen te worden, worden ook reguliere schoolbezoeken afgelegd. Elke school wordt in beginsel eens in de vier jaar bezocht. Tijdens dit bezoek wordt in principe geen gesprek met het schoolbestuur gevoerd.

2.3.3 Waarderingskaders en beslisregels

De grondslagen voor het waarderingskader zijn te vinden in de sectorwetten, de Wet op het primair onderwijs en de Wet op het voorgezet onderwijs en de WOT 2012.

De waarderingskaders van het toezichtskader VO 2013 en van het toezichtskader PO 2012 zijn gebaseerd op artikel 11, lid 2 en 3, WOT. Van belang zijn (lid 2):

- a. leerresultaten;
- b. voortgang in de ontwikkeling van leerlingen;
- c. het gevoerde personeelsbeleid, voor zover daar op grond van artikel 6a, eerste en tweede lid, aanleiding toe bestaat.

Als er een vermoeden bestaat dat het onderwijs tekortschiet, stelt de inspectie nader onderzoek in, waarbij tevens de oorzaken van het tekortschieten worden onderzocht. Bij dit nader onderzoek worden volgens lid 3 van artikel 11 WOT een of meer van de volgende aspecten van kwaliteit betrokken, te weten:

- a. het leerstofaanbod
- b. de leertijd
- c. het pedagogisch klimaat
- d. het schoolklimaat
- e. de effectiviteit van het didactisch handelen van het onderwijspersoneel
- f. de leerlingenzorg
- g. de inhoud, het niveau en de uitvoering van de toetsen, tests, opdrachten of examens
- h. voor wat betreft de kwaliteit van het onderwijspersoneel:
 - 1°. de eisen van bevoegdheid en bekwaamheid,
 - 2°. het personeelsbeleid, voor zover het betreft de duurzame borging van de kwaliteit van het onderwijspersoneel”.

Daarnaast vermeldt de WOT dat de kwaliteitszorg van de school bij het toezicht betrokken wordt (artikel 12, eerste lid, WOT).

Waarderingskader PO

Het waarderingskader van de Inspectie bestaat uit 54 indicatoren voor het PO. Het waarderingskader PO is onderverdeeld in vijf domeinen:

- opbrengsten,
- onderwijsleerproces,
- (leerlingen)zorg en begeleiding,
- kwaliteitszorg en
- wet- en regelgeving.

Binnen het domein opbrengsten is sprake van twee normindicatoren: de beoordeling van de eindresultaten (in de praktijk vrijwel uitsluitend: de CITO-score) en de beoordeling van de tussenresultaten. De eerste indicator betreft de resultaten van leerlingen aan het eind van de basisschool; die moeten ten minste op het niveau liggen dat op grond van kenmerken van de leerlingenpopulatie mag worden verwacht. De tweede indicator betreft de resultaten van de leerlingen voor Nederlandse taal en rekenen en wiskunde tijdens de schoolperiode; die moeten ten minste op het niveau liggen dat op grond van de kenmerken van de leerlingenpopulatie mag worden verwacht.

Voor de domeinen onderwijsleerproces en leerlingenzorg worden acht normindicatoren gehanteerd. Wanneer een school niet beschikt over eindresultaten van leerlingen en niet over

3 Zelfevaluatie

3.1 Inleiding

Ten behoeve van de evaluatie van het risicogerichte toezicht heeft de Inspectie een zelfevaluatie opgesteld. Die betreft onder meer de selectiviteit, de efficiency en de effectiviteit van het risicogerichte toezicht. In dit hoofdstuk wordt die zelfevaluatie besproken. In enkele gevallen wordt daarbij aanvullende informatie betrokken die naar aanleiding van de zelfevaluatie door de onderzoekers is opgevraagd. In het volgende hoofdstuk wordt het risicogerichte toezicht geëvalueerd. Op die plaats komen elementen van de zelfevaluatie opnieuw aan de orde. Achtereenvolgens worden hier eerst besproken de selectiviteit (par. 3.2), de efficiency (par. 3.3) en de effectiviteit van de risicoanalyse (par. 3.4).

3.2 Selectiviteit risicoanalyse

3.2.1 Inleiding

Om de betrouwbaarheid van de risicoanalyse te toetsen, heeft de Inspectie de risico-inschatting, gebaseerd op analyses van de opbrengsten, vergeleken met de beoordeling van de onderwijskwaliteit door de inspecteurs op basis van schoolbezoeken. Het ingeschatte risiconiveau is vergeleken met het arrangement zoals vastgesteld na afloop van een schoolbezoek.

3.2.2 Primair onderwijs

In tabel 3.1 zijn de uitkomsten weergegeven van de risicoanalyse op basis van opbrengsten en het uiteindelijke arrangement dat de inspectie zou moeten toekennen in onderzoeken over 2007 en 2008 bij in totaal 640 scholen voor primair onderwijs. Een voorbeeld ter illustratie: bij 281 scholen zijn geen risico's gesignaleerd en is het uiteindelijk toegekende arrangement 'basis'. Het gaat hier om een aselechte steekproef van scholen die allemaal door een inspecteur onderzocht en beoordeeld zijn.

TABEL 3.1: RISICOANALYSE VERSUS FEITELIJK ARRANGEMENT 2007/2008 (N=640)

	Toegekend arrangement		
	Basis	Zwak	Zeer zwak
Geen risico (N=286)	98,3% (N=281)	1,4% (N=5)	0,3% (N=1)
Risico (N=303)	93,4% (N=283)	5,6% (N=17)	1,0% (N=3)
Ernstig risico (N=51)	51,0% (N=26)	41,2% (N=21)	7,8% (N=4)

'Geen risico' betekent in dit verband dat de scholen positief scoren op de opbrengsten, bij 'risico' zijn er onvoldoende gegevens om dat vast te stellen. Bij 'ernstig risico' zijn de opbrengsten onvoldoende.

De tabel laat volgens de zelfevaluatie zien dat de risicoanalyse "zeer voorzichtig was afgesteld". Daarmee wordt bedoeld dat in relatief veel gevallen waar op grond van de risicoanalyse sprake zou (kunnen) zijn van risico, dit uiteindelijk volgens het toegekende arrangement niet het geval bleek te zijn. Hiervoor is de rij 'ernstig risico' bepalend. Van de 51 scholen die het betreft blijkt dat in ongeveer de helft van de gevallen waarin op basis van de analyse van opbrengsten sprake zou zijn van ernstig risico, de scholen toch in het basis toezichtsarrangement terecht zijn gekomen. Het gaat dus om 'false negatives'; deze scholen zijn in de risicoanalyse ten onrechte als risicovol bestempeld.

de opbrengstenanalyse gesignaleerde risico's in de latere analyse vaker blijken mee te vallen, zodat toch een basisarrangement wordt toegekend. De Inspectie stelt hierover zelf dat die toename het gevolg kan zijn van het na 2011 gebruiken van de septemberlevering van de examengegevens. Scholen die volgens de risicoanalyse, gebaseerd op de examens van jaar T, zwak scoren, maar waarvan uit de septemberlevering van de examendata van jaar T+1 blijkt dat de onvoldoende opbrengsten van het centraal eindexamen en/of het verschil tussen schoolexamens en het centraal eindexamen mogelijk zijn opgeheven, krijgen vooruitlopend op de volgende risicoanalyse het basisarrangement toegekend.

3.2.3 Conclusie

Concluderend wordt in de zelfevaluatie gesteld dat de toetsing van de indicatoren aan de waargenomen werkelijkheid laat zien dat de keuze voor de risico-indicatoren goed lijkt. Tegelijkertijd wordt in de zelfevaluatie gesteld dat veel indicatoren vooral indicaties zijn van onderpresteren. Daarmee wordt bedoeld dat onderpresteren wordt waargenomen nadat eerst onvoldoende opbrengsten zijn vastgesteld. De zelfevaluatie koppelt dat aan de observatie dat in de gevolgde systematiek vroegtijdige signalering nog niet altijd aan de orde is.

3.3 Efficiency van het toezicht

3.3.1 Inleiding

De invoering van het risicogerichte toezicht heeft op verschillende manieren gevolgen gehad. In de oorspronkelijke inrichting van het toezicht volgens de WOT (dus voor het risicogerichte toezicht) was sprake van een bezoek per jaar, gecombineerd met ten minste één uitgebreid onderzoek in de vier jaar. In de nieuwe inrichting is het jaarlijkse bezoek vervangen door de risicoanalyse. Als scholen dan risico's of tekortkomingen vertonen, krijgen zij vervolgens te maken met een veel zwaardere belasting dan in het verleden. Scholen waar niets aan de hand is, krijgen eenmaal in de vier jaar een lichte vorm van onderzoek. In totaliteit is er sprake van een reductie van de lasten voor zowel de scholen als de Inspectie. In de volgende subparagrafen wordt nader ingegaan op de toezichtskosten voor de Inspectie en de toezichtskosten voor de scholen.

3.3.2 Toezichtskosten voor de Inspectie

In de onderstaande tabel is de ontwikkeling te zien van het aantal fte's dat werkzaam is bij de Inspectie, per jaar, uitgesplitst per sector. De formatie is geleidelijk teruggelopen als resultaat van de taakstellingen met betrekking tot de uitgaven van de rijksoverheid.

TABEL 3.5: FORMATIE PER SECTOR VOOR TOEZICHT

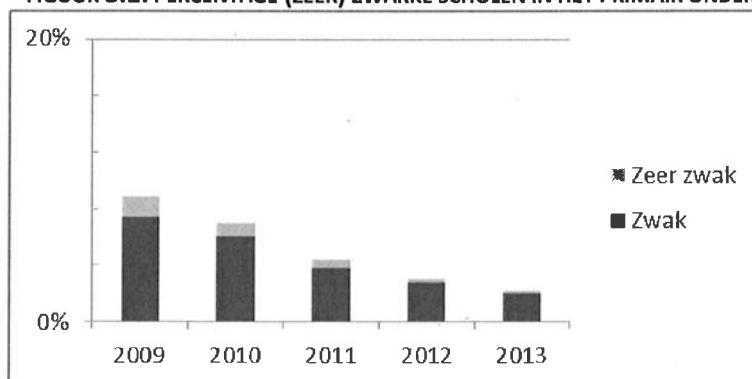
	PO	Vo
2010	133	46,8
2011	131	46,8
2012	131,1	46,1
2013	131,1	44,6
2014	126,8	44,6

Uit de beschikbare formatie concludeert de Inspectie dat er geen sprake kan zijn van toename van de toezichtlast voor de scholen omdat voor het toezicht minder inspecteurs beschikbaar zijn en het aantal scholen in belangrijke mate stabiel is gebleven.

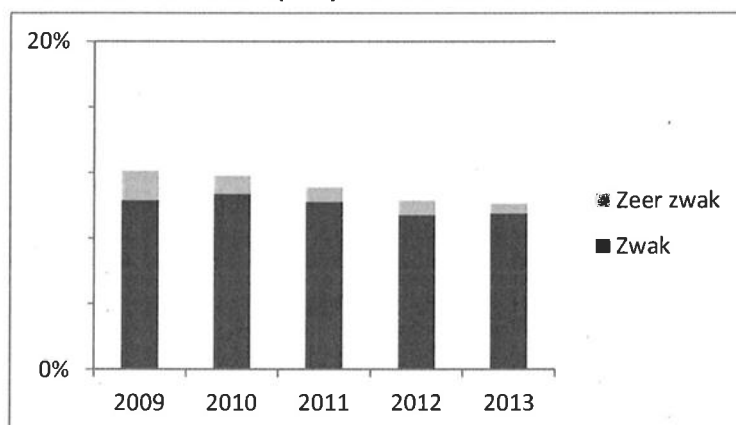
Of de Inspectie de beschikbare capaciteit effectief heeft ingezet, komt aan de orde in de par. 3.4.

Een van de manieren om naar de effectiviteit van het toezicht te kijken, is de ontwikkeling in het aantal (zeer) zwakke scholen te bestuderen. In de figuren 3.1 en 3.2 is de ontwikkeling van het percentage (zeer) zwakke scholen tussen 2009 en 2013 voor primair en voortgezet onderwijs weergegeven.

FIGUUR 3.1: PERCENTAGE (ZEER) ZWAKKE SCHOLEN IN HET PRIMAIR ONDERWIJS



FIGUUR 3.2: PERCENTAGE (ZEER) ZWAKKE SCHOLEN IN HET VOORTGEZET ONDERWIJS



Vanaf 2009 is er bij zowel het primair als het voortgezet onderwijs sprake van een afname van het aantal (zeer) zwakke scholen, zo blijkt uit bovenstaande figuren. In het voortgezet onderwijs is de teruggang van het aantal (zeer) zwakke scholen geringer dan in het primair onderwijs. Op grond van de ontwikkeling van het aantal zwakke en zeer zwakke scholen zou, zo wordt in de zelfevaluatie gesteld, kan worden geconcludeerd dat het risicogerichte toezicht zeer effectief is geweest in het waarborgen van de basiskwaliteit.

3.4.2 Doorlooptijd

De effectiviteit van het toezicht kan, aldus de Inspectie, ook worden afgemeten aan de snelheid waarin scholen zich ontwikkelen van (zeer) zwak naar basis. In grafiek 3.1 is deze doorlooptijd weergegeven. Hierbij heeft de Inspectie geen onderscheid gemaakt tussen primair en voortgezet onderwijs.

3.4.4 Bestuursinsteek

Besturen worden in de huidige werkwijze nadrukkelijk aangesproken op hun verantwoordelijkheden voor de kwaliteit van het onderwijs. Door het bestuur (de raad van bestuur of de bestuurder) nadrukkelijk als kwaliteitsbewaker aan te spreken, wordt daar de verantwoordelijkheid gelegd en zal het bestuur minder snel geneigd zijn geen of minimale activiteiten te ondernemen op het terrein van kwaliteitsbewaking en verbetering. De vraag naar de daadwerkelijke effectiviteit van de bestuursinsteek is op zijn inhoudelijke gevolgen niet te toetsen. Er kan, zo stelt de zelfevaluatie, wel gekeken worden naar randvoorwaardelijke effecten en naar percepties vanuit de besturen zelf. Een merkbaar gevolg ervan zou zijn dat de besturen zich veel meer dan tevoren op de hoogte willen stellen van het functioneren van hun scholen. Het leidt tot een sterkere betrokkenheid van de besturen bij de kwaliteitszorg. Vooral besturen met een of meer zeer zwakke scholen, zeggen zich veel bewuster te zijn geworden van hun verantwoordelijkheid op dit punt. Of dat echter ook geldt voor besturen waar zich geen grote problemen hebben voorgedaan, is niet bekend.

3.4.5 Conclusie

Concluderend stelt de zelfevaluatie dat de ontwikkeling van het aantal (zeer) zwakke scholen laat zien dat het risicogerichte toezicht zeer effectief is geweest in het waarborgen van de basiskwaliteit van scholen. Die effectiviteit blijft evenwel beperkt tot de categorie (zeer) zwakke scholen. Scholen die voldoen aan de norm van basiskwaliteit, blijken zich niet verder te verbeteren. Als wordt gekeken naar de doorlooptijd, kan geconcludeerd worden dat het risicogerichte toezicht van de inspectie effectief lijkt om scholen van zeer zwak naar basis te krijgen, maar minder effectief als het gaat om zwakke scholen. De effectiviteit van het toezicht als het gaat om het onderwijsproces lijkt niet groot. De effectiviteit van de bestuursinsteek is niet te toetsen.

4.3 Selectiviteit

De essentie van het model van risicogericht onderwijstoezicht is dat door het toezicht te richten op een selectie van scholen waarvan de onderwijskwaliteit onvoldoende is of dreigt te worden, de belasting van scholen beperkt wordt tot een minimum en de kosten van het toezicht voor de Inspectie eveneens beperkt zullen zijn. De risicoselectie wordt gebaseerd op een analyse van de opbrengsten aan de hand van een aantal indicatoren en daarop gebaseerde beslisregels. De risicoselectie werkt optimaal wanneer het aantal 'false negatives', de scholen die ten onrechte als risicovol worden aangemerkt, en het aantal 'false positives', de scholen die ten onrechte als risicoloos worden aangemerkt, beperkt blijft. In beide gevallen geschiedt de toets van de risicoanalyse tijdens het schoolbezoek dat volgt op het oordeel 'ernstig risico', dan wel tijdens het reguliere schoolbezoek, dat in beginsel op iedere school eens in de vier jaar plaats vindt.

De Inspectie claimt in de zelfevaluatie, zo hebben we in par. 3.2 beschreven, dat de risicoanalyse goed werkt en dat die ook steeds selectiever is geworden. Een vergelijking tussen de risicoanalyse 2007/2008 en die van 2013 laat zien dat het aantal scholen in het PO met een ernstig risico op basis van de opbrengsten, dat toch in een basisarrangement terecht is gekomen, is gedaald. Hier kan op twee manieren worden gekeken. Wordt het aantal scholen met ernstig risico afgezet tegen de totale populatie in de tabellen 3.1 en 3.2, dan is sprake van een afname van 4% (26 van 640) naar ruim 1% (89 van 6991). Wordt het aantal scholen met ernstig risico dat in een basisarrangement is terecht gekomen afgezet tegen het totale aantal scholen met ernstig risico, dan is nog steeds sprake van een daling, zij het dat die gering is (van 51% naar 48,9%).

Een ander punt is de 'false positives', de scholen die op basis van de opbrengstenanalyse geen risico leken te hebben, maar die op basis van het reguliere schoolbezoek toch als (zeer) zwak werden gekwalificeerd. Vanuit het oogpunt van bescherming van de leerlingen is dit een nog belangrijker punt. Hier doet de risicoanalyse het in het PO ook goed, gemeten naar het aantal scholen dat ten onrechte niet in de risicoanalyse naar boven is gekomen. Opnieuw is dit cijfer op twee manieren te bepalen. Als deel van het totaal gaat het in de cijfers over 2007/2008 om (26 van 640): 4%. In 2013 betreft het 62 op een totaal van 6991 scholen: 1%. Als deel van de totale populatie '(zeer) zwakke' scholen in de tabellen 3.1 en 3.2 gaat het nog steeds om een daling: van 50% (26 van 51) naar 31% (62 van 197).

De kans op 'false negatives' is volgens de cijfers dus gedaald, net als die op 'false positives'. De risicoanalyse laat overigens tussen 2007/2008 en 2013 een duidelijke afname zien van het aantal scholen met een ernstig risico op basis van de opbrengsten: dat aantal daalt van 7% naar 2%. De Inspectie spreekt van het 'scherper afstellen' van de risicoanalyse. Dat lijkt echter een verwarrende term: in werkelijkheid is de analyse juist minder kritisch geworden.

Bij bovenstaande vergelijkingen tussen de tabellen 3.1 en 3.2 moet wel worden opgemerkt dat die strikt genomen moeilijk te maken zijn. De indeling in categorieën verschilt, de normen die zijn aangelegd zijn ook gewijzigd in de loop van de tijd en bovendien zijn – zeker in tabel 3.2 – scholen meegenomen die niet zijn bezocht en die dus vrijwel per definitie in het basisarrangement vallen. In feite zou de conclusie dus moeten zijn dat niet te achterhalen is of de risicoanalyse 'scherper' is afgesteld.

De cijfers over het VO laten tussen 2011 en 2013 een toename zien van het aantal 'false negatives' van 1,5% (35 van 2442) naar 3% (80 van 2527). De 'false positives' blijven stabiel op ruim 1,5% (37 van 2442 in 2011 en 43 van 2527 in 2013). De risicoanalyse in het VO lijkt het dus (iets) minder goed te doen dan in het PO. Ook hiervoor geldt overigens dat veel scholen die wel zijn meegenomen in de tabellen, niet zijn bezocht waardoor ze vrijwel per definitie een basisarrangement hebben gekregen. Om een betere inschatting te kunnen maken van de 'false positives' en 'false negatives' zouden alleen scholen in de analyse betrokken moeten worden die daadwerkelijk bezocht zijn.

Bij een bespreking van de mate van selectiviteit van het onderwijstoezicht hoort ook een opmerking over de gestelde ondergrens voor kwaliteit. Het model van risicogericht toezicht slaagt er goed in de scholen met problemen in hun onderwijskwaliteit te selecteren op basis van een risicoanalyse van de opbrengsten. Daarmee is echter nog niet gezegd dat de juiste maatstaf wordt gehanteerd voor kwaliteit. Er zijn signalen die er op wijzen dat de voor kwaliteit gehanteerde maatstaf niet zeer streng is. Inspecteurs stellen in dit verband wel dat er heel wat mis moet zijn op een school voordat aan de onderwijskwaliteit het oordeel zeer zwak wordt toegekend. Dat neemt niet weg dat gegeven de gedefinieerde normen voor kwaliteit het risicogerichte toezicht er redelijk goed in lijkt te slagen de ondermaats presterende scholen te selecteren.

4.4 Efficiency

Bij efficiency gaat het om een onderwerp dat in het verlengde van selectiviteit ligt. Immers, hoe selectiever de risicoanalyse, hoe meer de Inspectie er in kan slagen de inzet te richten op de onderkant van de scholen die onvoldoende onderwijskwaliteit laten zien. De Inspectie claimt in de zelfevaluatie dat het risicogerichte toezicht efficiënt is. Die claim is in algemene zin te baseren op de systematiek van het risicogerichte toezicht: in plaats van elke school elk jaar te bezoeken wordt jaarlijks een risicoanalyse op basis van de opbrengsten uitgevoerd. Alleen die scholen met een 'ernstig risico' worden vervolgens bezocht. Daarnaast wordt net als voorheen elke school ten minste eens in de vier jaar bezocht. Het risicogerichte toezicht is dus efficiënter wanneer de kosten voor de Inspectie van het schoolbezoek van scholen die als risicovol zijn aangemerkt geringer zijn dan de kosten die voorheen gemoeid waren met de jaarlijkse bezoeken. Daarvan lijkt inderdaad sprake. De Inspectie heeft meer capaciteit in de analyse van de cijfers moeten investeren, maar de kosten van schoolbezoeken nemen af. Die kosten, in termen van de tijd die is gemoeid met een schoolbezoek aan een school met onvoldoende kwaliteit, nemen weliswaar toe, maar doordat daartegenover een sterke daling staat van de kosten van de reguliere schoolbezoeken (eens per vier jaar), is per saldo toch van een grotere efficiency sprake. De zelfevaluatie laat dat ook zien. Uit de cijfers van de Inspectie blijkt dat sprake is van minder fte's bij de Inspectie, maar tegelijkertijd slaagt de Inspectie er met die geringere capaciteit in die schoolbezoeken uit te voeren die geïndiceerd zijn op grond van de risicoanalyse.

Een andere vraag is hoe de toezichtslasten zich voor de scholen ontwikkelen. De Inspectie spreekt in de ex ante evaluatie van het risicogerichte toezicht van een beoogde halvering van de toezichtslasten. Duidelijk is dat de introductie van de systematiek er voor heeft gezorgd dat de reguliere schoolbezoeken niet langer elk jaar, maar eens in de vier jaar worden bezocht. De 'goede' scholen worden dus minder belast. Uit de cijfers blijkt dat de schoolbezoeken die volgen op het oordeel 'ernstig risico' veel intensiever zijn dan het jaarlijkse schoolbezoek voorafgaand aan de introductie van het risicogerichte toezicht. De toezichtslasten voor de (zeer) zwakke scholen zijn dus duidelijk toegenomen, maar omdat het gaat om een (zeer) klein deel van de totale populatie zijn per saldo de totale toezichtslasten voor de scholen afgenomen. Deze cijfers gelden voor het PO, voor het VO zijn die niet beschikbaar. Dat zich daar eenzelfde ontwikkeling heeft voorgedaan is niet onaannemelijk.

Behalve door de schoolbezoeken worden scholen ook belast door de verzoeken om informatie die de Inspectie tot ze richt. Deze bevragsingslast is door de Inspectie in kaart gebracht. Uit de zelfevaluatie blijkt dat die nagenoeg gelijk is gebleven voor het PO en licht gedaald is voor het VO. Het cijfer voor de bevragsingslast representeert het aantal vragenlijsten dat bij de scholen is uitgezet. Daarmee is nog geen specificatie beschikbaar voor de tijd die elke afzonderlijke vragenlijst kost. Een sluitende analyse hiervan is moeilijk te maken. Op basis van de gegevens van de Inspectie staat niet vast hoeveel tijd daadwerkelijk met de bevraging is gemoeid. Jaarlijks zijn er andere vragenlijsten en de met de vragenlijsten benodigde invultijd wordt op een steeds andere wijze ingeschat. Zo wordt de invultijd in 2010 en 2011 door de Inspectie ingeschat, terwijl de invultijd in 2012 en 2013 door de scholen diende te worden aangegeven. De representativiteit van de (gemiddelde) door scholen aangegeven invultijd is ook twijfelachtig.

een gunstige invloed heeft op de onderwijskwaliteit in brede zin, kan echter niet worden vastgesteld. De zelfevaluatie biedt daarvoor ook onvoldoende informatie. Dat scholen op voldoende niveau scoren in de risicoanalyse wil niet zeggen dat zij niet beter zouden kunnen. Juist om die reden is de Inspectie bezig met de ontwikkeling van Toezicht2020, dat zich richt op 'stimulerend toezicht' in de breedte van het onderwijs.

Op dit moment gaat van het toezicht dat de Inspectie uitoefent niet of nauwelijks een kwaliteitsimpuls uit die gericht is op scholen met een basisarrangement. De conclusie is dan ook dat het toezicht van de Inspectie slechts een beperkte invloed heeft op de algehele ontwikkeling van de kwaliteit van het onderwijs. Het signaal van de Inspectie voor scholen waarvan de onderwijskwaliteit voldoende is, is dat de Inspectie hen met rust laat. Dat kan voor die scholen goed zijn en ze de ruimte geven zich op eigen kracht te ontwikkelen. Het is echter ook mogelijk dat scholen die net voldoende scoren, op dat niveau blijven hangen; het onderwijstoezicht is voor hen geen prikkel zich te verbeteren. De Inspectie zou hierover de komende jaren informatie moeten zien te verzamelen, in aanloop naar de introductie van Toezicht2020, zodat onomstotelijk vast staat wat de invloed is van het toezicht op de kwaliteit van het onderwijs, ook wanneer de onderwijskwaliteit op zich voldoende is.

Overigens geven opbrengsten niet altijd de werkelijke meerwaarde van een school voor kinderen weer. Een gemiddelde CITO-score van 535 of een cijfer 6,5 op een centraal eindexamen kan voor de ene school een bovengemiddelde prestatie zijn terwijl het voor een andere school lager is dan verwacht mag worden. Weliswaar worden resultaten gewogen aan de hand van het opleidingsniveau van de ouders, maar er zijn meer factoren die de 'uitgangssituatie' van een leerling bepalen. IQ, sociaal-economische achtergrond, bevolkingsopbouw, een stimulerende thuissituatie, kennis van de Nederlandse taal, grootte van klassen (en de invloed op extremen op het gemiddelde), het zijn factoren die een gemiddelde CITO-score beïnvloeden maar die niet (volledig) verdisconteerd zijn in de generieke wegingsfactor, die uitsluitend het opleidingsniveau van de ouders meeweegt.⁷ Bovendien zijn de marges binnen de wegingsfactoren ruim; er zit een verschil tussen ouders die gemiddeld bijvoorbeeld aan de onderkant van MBO-niveau zitten en die aan de bovenkant zitten, maar er wordt dezelfde wegingsfactor op toegepast.

4.6 Neveneffecten

De vraag is of de door de Inspectie gehanteerde systematiek ook (ongewenste) neveneffecten kent. Dat lijkt zeker het geval. Er is bij de Inspectie een focus op opbrengsten, omdat een school bij onvoldoende opbrengsten in drie opeenvolgende jaren vrijwel per definitie (zeer) zwak wordt en de opbrengsten de basis vormen voor de risico-analyses. Door te focussen op opbrengsten ontstaat bij scholen de neiging zich in het onderwijs sterk op de opbrengsten te concentreren. Jaarlijks zijn er slechts enkele gevallen waarin het onderwijsproces dermate positief werd beoordeeld dat het de onvoldoende opbrengsten compenseerde en de school alsnog in het basisarrangement terecht kwam. Scholen vertonen daardoor de neiging de opbrengsten te overaccentueren, zeker als de school al in de gevarenzone zit en voldoende opbrengsten op een gegeven moment noodzakelijk zijn om niet zwak te worden verklaard. Hierdoor kunnen perverse prikkels optreden. Scholen voegen zich naar de normen van de Inspectie. Ze richten zich helemaal op het behalen van voldoende opbrengsten, hetgeen ten koste kan gaan van andere activiteiten.

Daarnaast zijn veel betrokkenen (schoolbesturen, schooldirecteuren, inspecteurs) het erover eens dat opbrengsten te manipuleren zijn. In het PO werd voorheen wel getracht de resultaten de zwakkere leerlingen buiten beschouwing te laten. Nadat de Inspectie daartegen is gaan optreden (door de CITO-score 517 toe te kennen aan leerlingen die zonder onderbouwing buiten de scores zijn gebleven) is een neveneffect dat tegenwoordig breed optreedt, zeker bij (zeer)

⁷ Wel geldt sinds augustus 2014 een nieuwe werkinstructie voor het herberekenen van de eindresultaten. Daarbij kunnen leerlingen buiten de beoordeling worden gelaten wanneer ze toelaatbaar zijn tot het praktijkonderwijs of het voortgezet speciaal onderwijs, wanneer leerlingen in groep 7 of 8 zijn ingestroomd en wanneer leerlingen sinds kort in Nederland verblijven en om die reden het Nederlands minder goed beheersen.

vervolgopleidingen waar de leerlingen naar toe gaan'), is er intern geen sterke traditie van onderlinge uitwisseling. Enige tijd geleden was nog wel sprake van een beleid dat zich richtte op regelmatige uitwisseling van medewerkers, maar momenteel is vooral sprake van specialisatie. Dat leidt tot verschillen in werkwijze, methodiek en cultuur. Naar verluidt is er ook sprake van verschillen in gestrengheid in het land. De verschillen in methodiek komen ook tot uitdrukking in de uiteenlopende kaders. De toezichtskaders zijn goed vindbaar, maar door de voortdurende wijzigingen die daarin worden aangebracht, maakt het geheel niet altijd een goed toegankelijke indruk. Aanvankelijk was sprake van geïntegreerde kaders voor PO en VO, sinds 2013 zijn er afzonderlijke kaders. Wat ook niet helpt is dat de terminologie niet consistent is; de ene keer heet het waarderingskader ook wel kernkader, dan weer is het kernkader een specifiek deel van de indicatoren binnen het waarderingskader. Wat een en ander nog lastiger te doorgronden maakt is dat de beslisregels niet (steeds) zijn opgenomen in de toezichtskaders, maar elders moeten worden gevonden. Het waarderingskader bestaat de ene keer uit vijf domeinen, dan weer uit drie domeinen. De wijze waarop de cijfers worden gepresenteerd (en verzameld?) wijzigt ook voortdurend, waarbij andere uitgangspunten worden gehanteerd en een afwijkende indeling in categorieën aan de orde is. De vergelijkbaarheid in de tijd is daardoor soms een probleem waar ook in dit onderzoek tegenaan is gelopen.

Om verschillen tussen inspecteurs te voorkomen wordt intern veel over casuïstiek gesproken, worden over 'twijfelgevallen' overlegd, worden de teams die op schoolbezoeken gaan wisselend samengesteld en is er overleg tussen inspecteurs over hun waarnemingen tijdens de schoolbezoeken. Die uitwisseling is van groot belang om interbeoordelaarsbetrouwbaarheid te bevorderen, hoewel die nooit kan worden gegarandeerd.

Een punt van kritiek dat de methodiek van het onderwijstoezicht betreft heeft betrekking op de werkwijze waarbij schoolbezoeken tevoren worden aangekondigd. Uit de gesprekken met zowel inspecteurs als schoolbesturen komt naar voren dat dit er toe leidt dat scholen zich op die bezoeken voorbereiden. Zo worden leraren extra gewezen op de eisen die de inspectie stelt aan bijvoorbeeld de opbouw en inhoud van een les of blijven mindere leraren soms een dagje thuis. Het is dan ook de vraag of het beeld dat tijdens de schoolbezoeken bij de inspecteurs ontstaat van het onderwijsproces en de kwaliteitszorg representatief is voor de algemene gang van zaken binnen de school. Uiteraard is een geoefend inspecteur in staat door de werkelijkheid heen te kijken, maar een element van vertekening blijft altijd bestaan.

4.8 Algemeen: concluderend

Onze algemene conclusie is dat het risicogerichte toezicht door de Inspectie er op een efficiënte wijze aan bijdraagt de onderwijskwaliteit op scholen op voldoende niveau te brengen. Het aantal (zeer) zwakke scholen is de afgelopen jaren gedaald en de tijd die het scholen kost van een aangepast arrangement in het basisarrangement terecht te komen is korter geworden. Dat wil echter niet zeggen dat geen kanttekeningen te plaatsen zijn bij de systematiek van risicogericht toezicht.

De eerste kanttekening betreft de focus op opbrengsten. De toespitsing op opbrengsten in de risicoanalyse leidt tot een efficiënte inzet van menskracht bij de Inspectie en tot relatief lage toezichtslasten voor de scholen. Tegelijkertijd kan de focus op de opbrengsten ten koste gaan van het onderwijsproces voor de onderwijskwaliteit. Dat is principieel het geval omdat de analyse op opbrengsten de intensiteit van het toezicht bepaalt, maar is ook vervolgens aan de orde omdat schoolbezoeken altijd in het teken staan van de tekortschietende opbrengsten.

De tweede kanttekening is dat door de focus op opbrengsten het ongewenste neveneffect optreedt dat het voor scholen aantrekkelijker wordt de resultaten te manipuleren. Er zijn veel signalen die er op wijzen dat daarvan inderdaad sprake is.

De derde kanttekening betreft de hoogte van de normen die worden aangelegd. Er wordt wel gesuggereerd dat de Inspectie 'te soft' zou zijn. Dat zou onder meer tot uitdrukking komen in de

5 Conclusie

In deze evaluatie is het risicogerichte toezicht onder de loep genomen. De centrale vraag in deze evaluatie richt zich op het succes van het risicogerichte toezicht in termen van selectiviteit, efficiency en kwaliteit. Hierna wordt deze vraag samenvattend beantwoord.

Het risicogerichte toezicht slaagt er in de risicoanalyse goed in de scholen met (zeer) zwakke onderwijskwaliteit te achterhalen. De systematiek is niet waterdicht: er zijn 'false positives', maar het aantal scholen dat het betreft is beperkt en is acceptabel. Het aantal 'false negatives' is iets groter. Dat is ook onontkoombaar, tenzij de risicoanalyse een zeer ruime tolerantiegrens zou krijgen. De vraag of het onderwijstoezicht ook effectief is, is uiteraard lastig te beantwoorden. Alternatieve verklaringen voor de impact van het risicogerichte toezicht kunnen niet uitgeschakeld worden. Toch is niet onaannemelijk dat het toezicht een belangrijke bijdrage heeft geleverd aan het verminderen van het aantal (zeer) zwakke scholen in ons land. Scholen doen er gemiddeld 1,5 jaar over van het aangepaste toezichtsarrangement 'zwak' weer in het basisarrangement terecht te komen. Zeer zwakke scholen doen daar uiteraard wat langer over. Dat neemt niet weg dat het opvallend is dat scholen gemiddeld relatief (1,5 tot 2 jaar) snel weer in het basisarrangement terecht komen. Het beeld bestaat dat de onderwijskwaliteit van scholen relatief snel onder de kritische grens terecht kan komen, maar ook relatief snel weer op niveau kan geraken. Dat onderstreept het belang van een systematiek waarbij voortdurend, dat wil zeggen van jaar tot jaar, de ontwikkeling van de onderwijskwaliteit op scholen wordt gevolgd.

Een bijzonder kenmerk van het risicogerichte toezicht is dat het vrijwel geheel de nadruk legt op de scholen aan de onderkant van het spectrum. Wanneer een school voldoende scoort komt die in het basisarrangement en is er niet of nauwelijks contact met de toezichthouder. Die komt volgens het algemene uitgangspunt eens in de vier jaar wel een kijkje nemen, maar een intensief schoolbezoek is dat niet. Op aandrang van de Tweede Kamer en ook op initiatief van de Inspectie gaat dat mogelijk veranderen. Toezicht2020 stelt zich ten doel het toezicht ook voor scholen die voldoende presteren te activeren. Dat betekent dat door 'stimulerend toezicht' of 'gedifferentieerd toezicht' toe te passen ook verbetering van voldoende scholen richting excellentie wordt bevorderd.

De vraag is hoe de Inspectie kan bevorderen dat scholen in het basisarrangement het beter gaan doen. De Inspectie lijkt hier bijna onvermijdelijk op twee gedachten te hinken. Enerzijds is aan het wezen van toezicht op de naleving verbonden dat de Inspectie geen inhoudelijke adviezen kan geven voor de door het schoolbestuur te nemen acties, anderzijds moeten scholen die in het basisarrangement zitten toch aangezet worden tot verbetering van de onderwijskwaliteit. Een van de instrumenten die dan ingezet kan worden is de verdere differentiatie van de onderwijskwaliteit in kwaliteitscategorieën. Waar nu (zeer) zwak en voldoende de onderscheidende grootheden zijn, roept de weg naar 'gedifferentieerd toezicht' de behoefte aan verdere specialisatie van de oordelen op. In het gedifferentieerde toezicht zullen scholen die duidelijkheid ook wensen. Deze weg kent risico's. De verdere detaillering van de waarderingskaders en beslisregels kan leiden tot ongewenste jurisdisering: regulering die steeds verdere verfijning van normen oproept. En dat roept weer het risico op van het ongewenste neveneffect van inzet op de indicatoren die bepalend zijn in de categorisering.⁸ De Inspectie stelt hierover dat de insteek aan de 'bovenkant' een wezenlijk andere moet zijn dan die aan 'de onderkant'. Gedacht moet worden aan dialoog, die tot stand komt tussen professionals. Hoe dat precies vorm moet krijgen is onderwerp van een zoektocht, waarin de Inspectie zich momenteel bevindt. Het is verstandig voor die experimenteerfase de tijd te nemen tegen de achtergrond van de hierboven geschetste risico's van gedifferentieerd toezicht.

De Inspectie is in haar zelfevaluatie zeer positief over de efficiency van het toezicht, positief over de risico-analyses en gematigd positief over de effectiviteit van het toezicht. In onze analyses hebben we hier kanttekeningen bij gezet, de zelfevaluatie was naar ons oordeel soms iets te

⁸ Zie ook Ferdinand Mertens, Naar gedifferentieerd toezicht: elke school in een vakje! Mei 2014.

Bijlage 1: Geraadpleegde documenten

Documenten

CDA, PvdA en ChristenUnie (2007), Coalitieakkoord.
Gemeente Amsterdam (2009), Kwaliteitsaanpak Basisonderwijs Amsterdam.
Inspectie van het Onderwijs, Onderwijsverslagen 2006 t/m 2013.
Inspectie van het Onderwijs, Jaarverslagen 2007 t/m 2013.
Inspectie van het Onderwijs, Jaarwerkplannen 2007 t/m 2014.
Inspectie van het Onderwijs (2006) Onderzoek zeer zwakke scholen.
Inspectie van het Onderwijs (2007) De WOT verwerkelijk. Nieuw toezicht voor de toekomst.
Inspectie van het Onderwijs (2008) Beslisschema voor het bepalen van een aangepast arrangement (zwakke of zeer zwakke school).
Inspectie van het Onderwijs (2008) Implementatieplan taakstelling, mei 2008.
Inspectie van het Onderwijs (2009) Toezichtkader PO VO 2009. Utrecht, juli 2009.
Inspectie van het Onderwijs (2009) Toezichtkader BVE 2009.
Inspectie van het Onderwijs (2009) Aanvulling voor 2010 op het toezichtkader BVE 2009.
Inspectie van het Onderwijs (2009) Beslisschema's aangepast arrangement PO.
Inspectie van het Onderwijs (2010) Evaluatie Risicogericht Toezicht.
Inspectie van het Onderwijs (2010) Risicoanalyse: Risico's op scholen in beeld.
Inspectie van het Onderwijs (2011) Toezicht op zwakke en zeer zwakke scholen: zo werkt het.
Inspectie van het Onderwijs (2012) De beoordeling van opbrengsten in het basisonderwijs.
Inspectie van het Onderwijs (2012) Ontwikkeling bevragslast
Inspectie van het Onderwijs (2012) Tevredenheidsonderzoek besturen primair onderwijsschooljaar 2011/2012.
Inspectie van het Onderwijs (2012) Toezichtkader PO/VO 2012.
Inspectie van het Onderwijs (2013) Analyse en waarderingen van opbrengsten PO.
Inspectie van het Onderwijs (2013) Bevragslast: stand van zaken 2012.
Inspectie van het Onderwijs (2013) De kwaliteit van het bestuurlijk handelen in het funderend onderwijs.
Inspectie van het Onderwijs (2013) Lessen uit het toezicht op ISG Ibn Ghaldoun.
Inspectie van het Onderwijs (2013) Onderzoek bestuurlijk handelen Islamitische Scholengemeenschap Ibn Ghaldoun.
Inspectie van het Onderwijs (2013) Toezichtkader VO 2013.
Inspectie van het Onderwijs (2013) Uitgangspunten werkwijze waarderingskader.
Inspectie van het Onderwijs (2013) Waarderingskader Voortgezet Onderwijs 2013.
Inspectie van het Onderwijs (2014) Bevragslast: stand van zaken 2013.
Inspectie van het Onderwijs (2014) Evaluatie toezicht 2013: het toezicht van PO in getallen.
Inspectie van het Onderwijs (2014) Kwaliteit nader beschouwd alle onderwijssoorten.
Inspectie van het Onderwijs (2014) Meerjaren opbrengsten VO 2014: toelichting.
Inspectie van het Onderwijs (2014) Onderzoek bestuurlijk handelen Stichting voor evangelische scholen (sVES).
Inspectie van het Onderwijs (2014) Oordelen OV 2013 – Kwaliteit meerjarenoverzicht.
Inspectie van het Onderwijs (2014) Resultaten tevredenheidsonderzoek Inspectie van het Onderwijs: Schooljaar 2012/2013.
Inspectie van het Onderwijs (2014) Resultaten tevredenheidsonderzoek Inspectie van het Onderwijs 2012/2013.
Inspectie van het Onderwijs (2014) Tevredenheidsonderzoek Expertisecentra schooljaar 2012/2013.
Inspectie van het Onderwijs (2014) Tevredenheidsonderzoek Primair Onderwijs schooljaar 2012/2013.
Inspectie van het Onderwijs (2014) Tevredenheidsonderzoek Voortgezet Onderwijs schooljaar 2012/2013.
J/M voor ouders (2009) Onderzoek vertrouwen in onderwijs onder ouders.
J/M voor ouders (2012) Onderzoek vertrouwen in onderwijs onder ouders.
Ministerie van OCW (2006), Kaderstellende visie op toezicht: toezicht in vertrouwen, vertrouwen in toezicht.



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

ZELFEVALUATIE RISICOGERICHT TOEZICHT

THEORIE EN PRAKTIJK VAN
HET ONDERWIJSTOEZICHT

INHOUD

Samenvatting 5
Inleiding 7

1 Aanleiding tot het onderzoek 10

- 1.1 Achtergrond 10
- 1.2 Onderzoeksvragen 10
- 1.3 Methodiek 11

2 De beleidscontext 12

- 2.1 Het theoretisch model 12
 - 2.1.1 Het model onder het risicogerichte toezicht 12
- 2.2 De rol van de belangrijkste actoren 14
 - 2.2.1 Het intern toezicht en de belanghebbenden 14
 - 2.2.2 De rol van de externe toezichthouder 16

3 Praktijkevidentie over de efficiency en effectiviteit van het risicogestuurde toezicht 20

- 3.1 Efficiency van het toezicht 21
 - 3.1.1 De risicoanalyse: betrouwbaarheid, validiteit en detectiesnelheid 21
 - 3.1.2 Ontwikkeling toezichtlast voor scholen 25
 - 3.1.3 Ontwikkeling tijd, middelen en taken van de inspectie 27
- 3.2 Draagvlak en gedragsverandering in het veld 27
 - 3.2.1 Ervaren effectiviteit 27
 - 3.2.2 Bestuursinsteek en gedrag besturen 28
 - 3.2.3 Kwaliteitsverbetering scholen door oordelen en interventies 30
- 3.3 Verbetering onderwijskwaliteit 31
 - 3.3.1 (Zeer) zwakke scholen 31
 - 3.3.2 Doorlooptijd zwakke en zeer zwakke scholen 32
 - 3.3.3 Kwaliteit onderwijsproces 33
 - 3.3.4 Leerprestaties 35
 - 3.3.5 Onderwijskwaliteit algemeen 36
 - 3.3.6 Maatschappelijk vertrouwen 37
- 3.4 Conclusie efficiency en effectiviteit risicogericht toezicht 38

4 Conclusie 39

- 4.1 De theorie achter het vernieuwde toezicht 39
- 4.2 De bestuursinsteek 40
- 4.3 De risicoanalyse 40
- 4.4 Proportioneel en selectief toezicht 41
- 4.5 Meer en sterkere interventies 41
- 4.6 Steviger positioneren van besturen en andere actoren 42
- 4.7 Discussie 42

BIJLAGE(N)

Samenvatting

Dit rapport beantwoordt de vraag in hoeverre het risicogerichte toezicht heeft geleid tot veranderingen in de effectiviteit en efficiency van het toezicht. Hiertoe is de theorie achter het risicogerichte toezicht uitgewerkt en vervolgens getoetst aan recente wetenschappelijke inzichten en de huidige Nederlandse praktijk. Dit rapport fungeert als zelfevaluatie. Op basis van dit rapport zal een onafhankelijke deskundige een toetsing uitvoeren.

Het rapport schetst eerst uitvoerig de beleidstheorie rond de governance met betrekking tot het Nederlandse onderwijs en de positie van de externe toezichthouder daarin. De besturen zijn zelf verantwoordelijk voor het onderhoud en de verbetering van de kwaliteit van het onderwijs. Dat doen zij in een omgeving waarin zij gecontroleerd worden door intern toezicht en waarin zij de dialoog aangaan met hun omgeving, in het bijzonder de direct belanghebbenden bij het onderwijs.

De externe toezichthouder, de inspectie, treedt op als waarborgende instantie waar de kwaliteit onder de ondergrens raakt of dreigt te raken. Zij intervenueert dan om ervoor te zorgen dat elke leerling onderwijs krijgt dat aan de standaarden voldoet. Het risicogerichte toezicht is erop gericht dat risico tijdig te onderkennen, terwijl gelijktijdig de bemoeienis van de inspectie met de scholen beperkt blijft. De besturen worden daarbij nadrukkelijk op hun verantwoordelijkheid aangesproken.

Aan de hand van bevindingen uit het toezicht en tussentijdse evaluaties daarvan heeft de inspectie in de praktijk getoetst of de veronderstellingen van het risicogerichte toezicht zich ook in de praktijk voordoet. Uit die bevindingen is gebleken dat het aantal zwakke en zeer zwakke scholen sinds de introductie van het risicogerichte toezicht aanmerkelijk is teruggelopen. Minder scholen worden (zeer) zwak en scholen blijven dat korter. Besturen lijken zich meer bewust van hun verantwoordelijkheid, speciaal als zij scholen hebben die onder de grens zakken.

Gelijktijdig blijkt dat de effecten zich vooral voordoen in en onder het grensgebied tussen voldoende (basistoezicht) en onvoldoende. Scholen boven de streep blijken zich niet verder te verbeteren. Voorts blijkt dat de horizontale dialoog van besturen met hun belanghebbenden nog maar weinig van de grond komt. Het veronderstelde mechanisme uit de beleidstheorie, werkt daar onvoldoende.

Geconcludeerd kan worden dat het risicogerichte toezicht zijn doelstellingen heeft gerealiseerd. Voor een verdere verbetering van het onderwijs kan het evenwel niet zorgen.

Inleiding

Recente vernieuwing onderwijstoezicht; risicogericht + bestuursinsteek

Sinds 2007 werkt de Inspectie van het Onderwijs met een vernieuwde vorm van toezicht. Er is sprake van twee belangrijke wijzigingen in de werkwijze. Ten eerste is een traject ingezet van risicogericht toezicht, dat wil zeggen afstemming van de omvang en aard van het toezicht op de risico's op een opleiding op school. De tweede wijziging is het nadrukkelijker aanspreken van het bestuur op zijn verantwoordelijkheid voor de kwaliteit van het onderwijs op zijn scholen. De vernieuwde werkwijze staat beschreven in "De WOT verwerkelijkt". De nieuwe werkwijze gaf vorm aan uitgangspunten die geformuleerd waren in de Kaderstellende Visie op Toezicht (2005) en de uitwerking daarvan in de Visie op Toezicht van OCW (Toezicht in vertrouwen, vertrouwen in toezicht 2006). In februari 2007 volgde daarop het coalitieakkoord, waarin deze uitgangspunten nader werden bevestigd en aangescherpt. Tenslotte verscheen in 2007 de evaluatie van de WOT.

Uit deze geschriften komen de volgende doelen van het risico- en bestuursgerichte toezicht naar voren:

- Toezicht vanuit de overheid moet zoveel mogelijk gebaseerd zijn op vertrouwen.
- Toezicht is transparant en legt aan de buitenwereld verantwoording af over zijn handelen.
- Onderwijsinstellingen krijgen meer mogelijkheden invulling te geven aan het onderwijs.
- De kwaliteit van het onderwijs moet worden bevorderd en gegarandeerd.
- Wat leerlingen en studenten moeten kennen en kunnen wordt duidelijk vastgelegd, evenals de maatschappelijke doelstellingen van het onderwijs.
- Onderwijsinstellingen leggen over de resultaten verantwoording af aan ouders, studenten en de minister.
- Als de kwaliteit tekortschiet, moet de minister snel en effectief kunnen ingrijpen.
- Besturen zijn verantwoordelijk voor de kwaliteit van het onderwijs en het overige functioneren van de scholen.
- Toezicht op onderwijskwaliteit en financieel toezicht worden geïntegreerd.
- Bij goed presteren van onderwijsinstellingen zal sprake zijn van vermindering van het toezicht.
- In het algemeen zal sprake zijn van het terugdringen van de toezichtlast met 25 procent.

In "De WOT verwerkelijkt" zijn deze uitgangspunten uitgewerkt in een vernieuwd toezichtmodel. De belangrijkste veranderingen in dit nieuwe model zijn (a) de risicogerichte aanpak en (b) de bestuursinsteek. De gehele nieuwe werkwijze wordt gemakshalve het 'risicogerichte toezichtmodel' genoemd. Dit toezichtmodel is gebaseerd op een regelmatige risicoanalyse die input levert voor het toezichtproces. In principe krijgen scholen waar geen risico's zijn vastgesteld, geen verder toezichtonderzoek. In het Algemeen Overleg met de vaste kamercommissie op 3 oktober 2007, waar de risicogerichte werkwijze is bediscussieerd, is daarover afgesproken dat in elk geval ook goede scholen minstens eenmaal in de vier jaar in enigerlei vorm worden onderzocht.

Hoofdstuk 3 bevat een toetsing aan de huidige praktijk. Dit hoofdstuk toetst of de feitelijke inspectiepraktijk de aannames van het risicogerichte toezicht ondersteunt, of juist niet. We beantwoorden hierin de vragen of het risicogerichte toezicht en de bestuursinsteeek wel of niet effectiever en efficiënter zijn, en welke onderdelen nog kritische elementen zijn.

Conclusies worden getrokken in hoofdstuk 4. Hierbij gaan we met name in op de mate waarin de informatie uit de praktijk de nieuwe werkwijze ondersteunt. We geven hierbij antwoord op de vraag of het risicogerichte toezicht wel of niet effectiever en efficiënter is geworden en waar verbeteringen wenselijk of mogelijk zijn.

1.3

Methodiek

Reconstructie en evaluatie van een beleidstheorie is een methode die vaker gebruikt wordt voor evaluatie van overheidshandelen. Ze bestaat uit drie stappen, namelijk:

1. reconstructie van het theoretisch model; uitschrijven van de onderliggende assumpties en het construeren van een causaal model
2. evaluatie van de onderliggende assumpties
3. conclusie opstellen aan de hand van een weging van de evaluatiebevindingen

Het is een theoriegestuurde evaluatiemethodiek die vaak ook gebruikt wordt voor ex-ante evaluaties. De methodiek staat beschreven in Pawson en Tilley (1997), Weiss (2000) en Leeuw (2003, p7).

Recente voorbeelden van reconstructie en evaluatie van beleidstheorieën op het terrein van onderwijs en toezicht zijn te vinden in Ehren, Leeuw en Scheerens (2006) en Janssens en De Wolf (2009). In beide gevallen gaat het om een (ex-ante) evaluatie waarbij de (beleids)assumpties worden getoetst aan wetenschappelijke inzichten. In de studie in 2010 hebben we de assumpties getoetst aan de wetenschappelijk inzichten en aan recente en oude evaluaties van de Nederlandse inspectiepraktijk. In deze studie zullen wij ons beperken tot dit laatste, in het bijzonder op grond van de ervaringen sinds 2010.

In het model zijn drie actoren onderscheiden. Centraal staat, in de middelste kolom, het *autonoom bestuur*; het bevoegd gezag van scholen of instellingen. Het bevoegd gezag is primair verantwoordelijk voor de borging en verbetering van de kwaliteit van het onderwijs. Over zijn resultaten op dat punt (en ook op een aantal andere punten) verantwoordt het bevoegd gezag zich extern¹. Deze externe verantwoording is een centraal punt in de gehele theorie, zowel voor het bevoegd gezag zelf als voor de overige actoren. Op grond van deze verantwoording voert het bevoegd gezag verbeteractiviteiten uit.

Hierbij kent het bevoegd gezag twee ondersteunende mechanismen, namelijk het (interne) kwaliteitszorgsysteem en het intern toezicht. Beide zijn in nauw verband met elkaar in het model opgenomen. Zij hebben een samenhangende functie, waarbij het kwaliteitszorgsysteem primair bedoeld is als hulpmiddel voor het bevoegd gezag zelf om zijn instelling of zijn verschillende scholen te controleren en te sturen, terwijl het intern toezicht er is om, in samenhang daarmee, het bevoegd gezag te controleren. De vorm van het mechanisme van interne kwaliteitszorg kan zeer uiteenlopen en kan centraal of per school georganiseerd zijn.

De tweede actor wordt gevormd door de *belanghebbenden*, weergegeven in de rechterkolom. Dit zijn leerlingen en ouders, maar ook het afnemend veld (vervolgsscholen, werkgevers). De belanghebbenden hebben verwachtingen met betrekking tot (onderdelen van) de school of instelling en hebben klachten indien niet aan deze verwachtingen wordt voldaan. Deze wensen of klachten uiten ze primair aan de school, maar klachten delen zij vaak ook met de toezichthouder (zeker als de school onvoldoende reageert). Aan klachten kunnen zij gevolgen verbinden door school- of opleidingskeuze, door middel van "voice" (het aanspreken van het bevoegd gezag) of door vertrek.

De laatste actor is de *externe toezichthouder*, in dit geval de Inspectie van het Onderwijs. Deze staat, omcirkeld, weergegeven in de linkerkolom van het model. Op deze actor richt deze studie zich. De inspectie gebruikt als referentie voor haar toezicht een toezichtkader, samengesteld uit wettelijke voorschriften en een samenhangende opvatting over onderwijskwaliteit. Dit toezichtkader is opgesteld in samenspraak met belanghebbende partijen in het veld.

Het toezicht van de inspectie is proportioneel en selectief². Daarbij baseert de inspectie zich allereerst op de externe verantwoording door de instelling³ (als die er is). De inspectie valideert deze externe verantwoording met behulp van onafhankelijke bronnen en voert daarmee een risicoanalyse uit. Op grond hiervan bepaalt zij of er toezichtonderzoek nodig is om vast te stellen of er tekortkomingen zijn, dan wel of zij vertrouwen kan hebben in de kwaliteit van het onderwijs. In beide gevallen leidt dat tot een openbaar oordeel. Indien het oordeel negatief is (er zijn belangrijke tekortkomingen vastgesteld), leidt dat tevens tot een interventie die beschreven wordt bij het openbare oordeel.

Het mechanisme om selectiviteit te realiseren, is de risicoanalyse. Deze risicoanalyse op zichzelf vergt regelmatige controle op haar juiste werking. Hiervoor gebruikt de inspectie controleonderzoek dat past bij de wens van de Tweede Kamer om elke school minstens eenmaal in de vier jaar te onderzoeken, ook als er geen risico is.

¹ Het begrip verantwoording suggereert eenrichtingsverkeer. Zoals de Algemene Rekenkamer opmerkt in Kaders voor Toezicht en Verantwoording (Algemene Rekenkamer, 2008, p. 29) moet hierin echter ook de dialoog worden begrepen, met name met de belanghebbenden.

² Volgens de memorie van toelichting van de oude WOT impliceert proportionaliteit ook selectiviteit (MvT WOT, 2001, paragraaf 2.8)

³ WOT, artikel 12 lid 2. In de oude WOT verwijst dit naar een mogelijke zelfevaluatie, in de nieuwe WOT naar de (verplichte) externe verantwoording. Een zelfevaluatie was niet verplicht (en daarmee schaars).

Kwaliteitsinstrumenten zijn geen vorm van toezicht of verantwoording. Zij kunnen wel – mits de betrouwbaarheid voldoende is gewaarborgd – een belangrijk onderdeel zijn van het geheel van checks-and-balances rond een instelling. Zij kunnen immers voor de instelling een prikkel tot beter presteren vormen. Ze kunnen ook een bron van informatie zijn voor kritische klanten. Het verticale toezicht kan in een aantal gevallen gebruik maken van de uitkomsten van deze kwaliteitsinstrumenten. In andere gevallen hebben deze instrumenten juist alleen effect als ze kunnen worden toegepast zonder dat de uitkomsten meteen leiden tot oordelen en interventie door een toezichthouder.”

Er zijn verschillende studies die kanttekeningen plaatsen bij het functioneren van het intern toezicht en de meervoudige publieke verantwoording en daarmee ook bij de aannamen die daarover zijn geformuleerd.

Hooge e.a. (2006) geven aan dat de raden van toezicht onvoldoende professioneel zijn om extern toezicht gedeeltelijk te kunnen “vervangen” in de vorm van proportioneel toezicht. Daar komt bij dat ze nadrukkelijk een onderdeel zijn van de instelling zelf en daardoor ook verbonden met de belangen van de instelling. De Algemene Rekenkamer (2008, p.15) stelt daarom ook dat intern toezicht door raden van toezicht of visitaties niet in de plaats kan komen van extern toezicht waar het gaat om het waarmaken van de ministeriële verantwoordelijkheid voor het onderwijs (zie ook paragraaf 2.3).

De WRR (2014) merkt, in navolging van de Commissie Governance, Handhaving en Codes (Commissie Hooge) (2013) op dat er de afgelopen jaren een duidelijke vooruitgang heeft plaatsgevonden in het interne toezicht, maar dat er bij veel raden van toezicht nog altijd ambitie, scherpste en een cultuur van hoge verwachtingen over de eigen inzet worden gemist. De WRR stelt vast dat er, naast het bestuur en de raad van toezicht, een derde partij nodig is om de beide eerste bij de les te houden.

Daarmee zou het gat worden gevuld dat is ontstaan door de zogenaamde “verstichtelijking” van de publieke dienstverlening in de laatste decennia, waarmee de verbinding met het maatschappelijk veld (voorheen vaak de leden van een vereniging) is verdwenen. Bij de invulling van deze derde partij zou moeten worden voldaan aan enkele belangrijke randvoorwaarden. De relatie zou voldoende werkzaam moeten zijn (voldoende bevoegdheden en expertise, voorzien van incentives en vasthoudend) maar tegelijk werkbaar waardoor de organisatie niet wordt verlamd. Vooral rolvastheid is in dat kader van belang.

De WRR (2014) beschrijft een aantal partijen die invulling kunnen geven aan een dergelijke driehoeksrelatie. Dat varieert van professionele partijen zoals controllers, formele inspraakorganen zoals ondernemings- en medezeggenschapsraden tot min of meer geïnstitutionaliseerde vormen van inspraak door maatschappelijke belanghebbenden. Naast deze bestaande opties poneert de WRR enkele mogelijkheden tot instelling van organen van maatschappelijke belang- en aandeelhouders, waaronder een mogelijke terugkeer naar een ledenraad of ledenvergadering.

Maatschappelijke belanghebbenden bij scholen omvatten onder meer ouders en leerlingen. Janssens en De Wolf (2009) merken op dat ouders/leerlingen slechts in zeer beperkte mate de dialoog voeren (‘voice’) over de kwaliteit van het gebodene, zelfs als daar expliciet de gelegenheid toe wordt geboden, en dat ze weinig beïnvloed worden in hun keuze naar de school te gaan of de school te verlaten.

toezicht moeten de elementen informatieverzameling, oordeelsvorming en interventie goed op elkaar zijn afgestemd.

Drie onderdelen van het school- en instellingstoezicht van de externe toezichthouder in het model zijn mogelijk effectief, namelijk:

- a. proportioneel/selectief toezicht (op basis van verantwoording en onafhankelijke bronnen)
- b. toezichtonderzoek (of school/instellingsonderzoek) met een toezichtkader
- c. openbare oordelen/interventies

Voor het *proportioneel of selectief toezicht* op scholen en instellingen geldt dat dit in theorie efficiënter en effectiever wordt geacht dan traditionele vormen van toezicht. Deze aanname is deels gebaseerd op de veronderstelling dat de inspectie maximaal gebruik maakt van de externe verantwoording van het bestuur, en daarmee de verantwoordelijkheid voor het onderwijs bij het bestuur neerlegt. De beleidstheorie gaat er ook vanuit dat de toezichthouder de externe verantwoording van het bestuur op waarde kan schatten door het gebruik van onafhankelijke bronnen. Wanneer belanghebbenden klachten hebben, laten ze dit in het model ook aan de toezichthouder weten, die hiermee een extra bron voor toetsing van de verantwoording in handen heeft. Met deze signalen, de verantwoording en aanvullende, onafhankelijke bronnen kan de toezichthouder detecteren welke scholen en instellingen risicovol zijn. Hiermee ziet de toezichthouder risico's en mogelijke tekortkomingen sneller. Wanneer de capaciteit vervolgens op deze risicoscholen wordt ingezet, wordt deze beter benut. De toezichthouder is zo in theorie optimaal efficiënt door vooral daar toezicht te houden waar het 't hardst nodig is en waar het ook het meest effectief is.

De WRR (2013) plaatst ook kanttekeningen bij het risicogerichte toezicht. De raad stelt dat een te sterke focus op risicovolle organisaties kan leiden tot een papieren exercitie en bovendien de toezichthouder het zicht op het veld beneemt. Ook zouden gemiddeld presterende instellingen het risicogerichte toezicht als een vorm van verwaarlozing kunnen beschouwen en zich minder geprikkeld kunnen voelen zich verder te verbeteren⁵. Tenslotte zou het ten koste kunnen gaan van de gepercipieerde pakkans, waardoor de graad van naleving daalt. De wens van de Tweede Kamer dat de inspectie periodiek alle scholen in ieder geval onderzoekt en de behoefte van de inspectie zelf om ten behoeve van een andere, wettelijke taak (het beschrijven van de staat van het onderwijs in Nederland) zicht te houden op "gemiddelde" organisaties, gaat enerzijds ten koste van de maximale efficiency die mogelijk zou zijn, maar komt wel tegemoet aan deze vrees.

Als de toezichthouder naar een *school of instelling onderzoek doet met een toezichtkader*, zal het toezicht volgens het model kwalitatief beter en effectiever zijn. Een toezichtkader komt tot stand in overleg tussen de toezichthouder en het onderwijsveld en leidt tot (of is gebaseerd op) gedeelde normen over en wetenschappelijk onderzoek naar het onderwijs. Een toezichtkader omvat alle relevante aspecten van het te beoordelen onderwijs en waarborgt een faire beoordeling van de school. Het bestuur zal in zijn externe verantwoording rekening houden met het toezichtkader, waardoor de bruikbaarheid van de verantwoording voor de toezichthouder toeneemt. Daarnaast zal de kwaliteit van het inspectieoordeel groter zijn indien de toezichthouder een school beoordeelt met een toezichtkader, aldus de theoretische aannames onder het model.

⁵ Dit is een aanname die ook ten grondslag ligt aan het programma Toezicht 2020 van de Inspectie van het Onderwijs. Het is voor toekomstig onderzoek interessant om na te gaan in hoeverre deze veronderstelling ondersteund wordt.

3.1 Efficiency van het toezicht

Heeft het risicogerichte toezicht geleid tot veranderingen in de efficiency van het toezicht en zo ja, welke? Is de risicoanalyse betrouwbaar en valide, komen potentiële risicoscholen hiermee tijdig in beeld en is de detectiesnelheid van (zeer) zwakke scholen toegenomen? Is de toezichtlast voor het veld kleiner geworden? En hoe heeft de inzet van tijd en middelen van de inspectie zich ontwikkeld in relatie tot haar taken?

3.1.1 De risicoanalyse: betrouwbaarheid, validiteit en detectiesnelheid

Jaarlijks onderzoekt de inspectie op een aselechte steekproef van scholen de kwaliteit van het onderwijs. Dit doet zij om een representatief beeld te krijgen van de staat van het onderwijs en om haar methoden van risicodetectie te toetsen (en zonodig aan te scherpen). Om de betrouwbaarheid van de risicoanalyse te toetsen, is eerder de risico-inschatting vergeleken met de feitelijke beoordeling door inspecteurs. Hierbij is het ingeschatte risiconiveau vergeleken met het arrangement zoals vastgesteld door inspecteurs tijdens een schoolbezoek (zie, voor deze methodiek, ook De Wolf, 2009). Alleen wanneer sprake is van een sterke relatie is tussen de opbrengstendetectie en het arrangement, kunnen we spreken van een bruikbare risicoanalyse.

Regelmatig worden de beslisregels van de primaire detectie geëvalueerd. Als de evaluatie daartoe aanleiding geeft, worden de beslisregels aangepast. Deze aanpassingen zijn gericht op het verder verkleinen van het aantal false negatives en false positives.

Basisonderwijs

In Tabel 2 staan de resultaten van het eerdere onderzoek. Hierbij zijn tegenover elkaar gezet de uitkomsten van de risicoanalyse (primaire detectie) en het uiteindelijke arrangement dat de inspectie zou moeten toekennen in de onderzoeken over 2007 en 2008, in totaal 640 basisscholen. Het gaat hier dus nadrukkelijk om een aselechte steekproef van scholen die allemaal door een inspecteur onderzocht en beoordeeld zijn.

		Toegekend arrangement			Totaal	
		Basis	zwak	zeer zwak	abs.	rel.
Primaire detectie	geen risico	98,3%	1,4%	0,3%	286	44,7%
	risico	93,4%	5,6%	1,0%	303	47,3%
	ernstig risico	51,0%	41,2%	7,8%	51	8,0%
Totaal abs.		590	42	8	640	100%
Totaal rel.		92,2%	6,6%	1,3%		

Tabel 2 Risicoanalyse versus feitelijk arrangement 2007 en 2008 (basisonderwijs)⁶

⁶ Bron: Inspectie van het Onderwijs

Verdeling arrangementen na Primaire Detectie 2013		
PD13	Arr na 1-4-2013 en ouder	Totaal
Enstig Risico	Aangepast Zeer Zwak	12
	Aangepast Zwak	73
	Basis	24
	Basis met attendering	10
	Basis met waarschuwing	47
	ouder arrangement: Aangepast Zwak	8
	ouder arrangement: Basis	4
	ouder arrangement: Basis met waarschuwing	4
Totaal Enstig Risico		182
Risico	Aangepast Zeer Zwak	2
	Aangepast Zwak	22
	Basis	39
	Basis met attendering	20
	Basis met waarschuwing	242
	geen arrangement	3
	ouder arrangement: Aangepast Zwak	2
	ouder arrangement: Basis	8
	ouder arrangement: Basis met waarschuwing	1
Totaal Risico		339
Enig Risico	Aangepast Zwak	9
	Basis	110
	Basis met attendering	1049
	geen arrangement	2
	ouder arrangement: Aangepast Zwak	1
	ouder arrangement: Basis	25
	ouder arrangement: Basis met attendering	1
	ouder arrangement: Basis met waarschuwing	1
Totaal Enig Risico		1198
Geen risico	Aangepast Zwak	18
	Basis	4456
	Basis met attendering	12
	Basis met waarschuwing	3
	geen arrangement	11
	ouder arrangement: Aangepast Zwak	10
	ouder arrangement: Basis	70
	ouder arrangement: Basis met waarschuwing	5
Totaal Geen risico		4585
Onbekend risico	Aangepast Zeer Zwak	8
	Aangepast Zwak	23
	Basis	371
	Basis met attendering	190
	Basis met waarschuwing	23
	geen arrangement	23
	ouder arrangement: Aangepast Zeer Zwak	1
	ouder arrangement: Aangepast Zwak	7
	ouder arrangement: Basis	33
Totaal Onbekend risico		679
Foute Verantwoording	Aangepast Zwak	1
	Basis	4
	geen arrangement	2
	ouder arrangement: Basis	1
Totaal Foute Verantwoording		8
Eindtotaal		6991

Tabel 3 Verdeling risico's en arrangementen 2013

Uit dit alles blijkt dat de risicoanalyse licht verbeterd is, in die zin dat het percentage false positives en false negatives wat is afgenomen. Blijft dat de risicoanalyse in het

- De stijging van basis arrangementen bij zwak volgens de primaire detectie kan het gevolg zijn van het gaan gebruiken van de septemberlevering van de examengegevens. Scholen die volgens de PD (gebaseerd op de examens van jaar t) zwak scoren, maar waarvan uit de septemberlevering (met alle examendata van jaar t+1) blijkt dat de onvoldoende opbrengsten bij CE en/of verschil SE-CE mogelijk zijn opgeheven, krijgen vooruitlopend op de volgende editie van de PD al het arrangement basis.

3.1.2 Ontwikkeling toezichtlast voor scholen

Het risicogerichte toezicht beoogt efficiënter te zijn dan andere vormen van toezicht. Zijn er indicaties dat dit voor de Nederlandse situatie het geval is? Wat betekent de nieuwe werkwijze voor de scholen, de besturen en de inspectie zelf?

Berekende toezichtlast scholen

De invoering van het risicogerichte toezicht, in combinatie met een taakstelling voor de inspectie zelf, heeft op verschillende manieren gevolgen gehad. Allereerst was van belang de toezichtbelasting die scholen of instellingen ervaren als gevolg van het toezicht. In de oorspronkelijke inrichting van het toezicht volgens de WOT (dus voor het risicogerichte toezicht) was sprake van een bezoek per jaar, gecombineerd met tenminste één uitgebreid onderzoek in de vier jaar. In de nieuwe inrichting is het jaarlijkse bezoek vervangen door de risicoanalyse. Als scholen dan risico's of tekortkomingen vertonen, krijgen zij vervolgens een veel zwaardere belasting dan in het verleden. Scholen waar niets aan de hand is, krijgen eenmaal in de vier jaar een lichte vorm van onderzoek. In totaliteit is er sprake van een reductie.

De onderstaande tabel laat zien tot welke reducties van toezichtlasten bij de scholen voor funderend onderwijs het risicogerichte toezicht heeft geleid.

	WOT 2001	RGT	Reductie
Jaarlijks/vierjaarlijks onderzoek	6325395	2644223	-3681172
Kwaliteitsonderzoek	1244341	1244341	
Incidenteel	77771	77771	
Totaal basisonderwijs	7647507	3966335	-3681172
Totaal voortgezet onderwijs	2227906	1295242	-932664
Totaal toezichtlast	9875413	5261577	-4612836
Procentueel	100%	53,3%	-46,7%

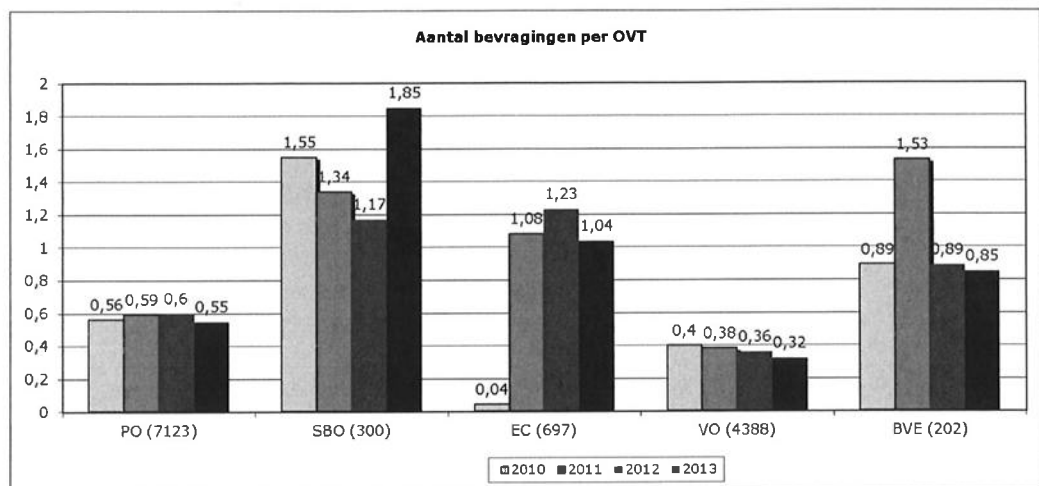
Tabel 6 Ontwikkeling toezichtlasten (in euro's, prijspeil 2009⁸) basis en voortgezet onderwijs⁹

Er is niet enkel sprake van een vermindering van de toezichtlast, maar ook van een verschuiving. Toezicht op (zeer) zwakke scholen is effectiever dan toezicht op goede scholen. De risicogerichte werkwijze heeft ertoe geleid dat de focus verlegd is naar de zwakkere scholen, waar dus meer toezicht aan wordt besteed.

Dit is inzichtelijk te maken door te bekijken hoeveel toezichtlast de scholen ervaren. Deze toezichtlast geeft immers weer in welke mate scholen direct gevolgen

⁸ De aantallen uren zijn nadien niet veranderd. De prijzen uiteraard wel. Op basis van het algemeen prijsindexcijfer (CBS-Statline) zouden deze bedragen met 8,85 procent moeten worden verhoogd.

⁹ Ministerie van OCW, Herberekening reductiemaatregelen 1/3/2007-1/7/2009, Eindrapportage, augustus 2009



Figuur 2 Ontwikkeling van de bevestigingen per sector

Afgezien van enkele pieken laat de bevestigingslast over het algemeen een stabiele of lichtelijk dalende tendens zien. De enorme piek in het SBO in 2013 is een gevolg van een eenmalige actie bij scholen zonder risico's over de afgelopen jaren, bedoeld ter evaluatie van het risicomodel en voorbereiding van inspectiebezoek.

3.1.3 *Ontwikkeling tijd, middelen en taken van de inspectie*

De ontwikkeling van de inzet van tijd en middelen van de inspectie is voortdurend aan veranderingen onderhevig door uitbreiding van taken met nieuwe middelen etc. Voor deze evaluatie is evenwel relevant de ontwikkeling van de formatie voor de kernactiviteiten van de sectoren po, ec en vo. In tabel 7 is die ontwikkeling te zien in aantallen fte per jaar per sector. De formatie is geleidelijk teruggelopen als resultaat van de taakstellingen met betrekking tot de uitgaven van de rijksoverheid, maar zijn overigens constant gebleven aangezien het scholenveld niet ingrijpend is veranderd.

	po	ec	Vo	totaal
2010	133	27,95	46,8	207,75
2011	131	27,95	46,8	205,75
2012	131,1	27,55	46,1	204,75
2013	131,1	26,75	44,6	202,45
2014	126,8	26,75	44,6	198,15

Tabel 7 Ontwikkeling formatie per sector

Uit het verloop van de beschikbare formatie kan worden geconcludeerd dat er in fysieke zin geen sprake kan zijn van toename van de toezichtlast, aangezien daarvoor minder inspecteurs beschikbaar zijn. Met dat geringere aantal inspecteurs is de inspectie er wel in geslaagd de in 2008 ingezette vermindering van het aantal zwakke en zeer zwakke scholen door te zetten.

3.2 **Draagvlak en gedragsverandering in het veld**

3.2.1 *Ervaren effectiviteit*

In het kader van het eerder vermelde onderzoek naar de toezichtlasten, is de scholen ook gevraagd of de investering die zij zelf in het inspectieonderzoek hebben

vijf de onderzochte indicatoren¹ als voldoende beoordeeld. In het speciaal basisonderwijs zijn zicht hebben op de kwaliteit, evalueren/analyseren en verantwoording afleggen de zwakke punten. Het speciaal basisonderwijs heeft minder duidelijk afgebakende landelijke benchmarks voor de opbrengsten dan het basisonderwijs; dit kan een reden zijn voor de verschillen met de besturen van basisscholen. Slechts een kwart van de besturen in het speciaal basisonderwijs is op alle vijf indicatoren als voldoende beoordeeld.

In het voortgezet onderwijs is verantwoording afleggen het zwakste punt, waar een derde van de onderzochte besturen op tekortschiet. Slechts iets meer dan de helft van de besturen is op alle vijf indicatoren als voldoende beoordeeld. In het (voortgezet) speciaal onderwijs zijn evaluatie en analyse, naast borging en verantwoording afleggen, de zwakkere punten. Ook deze sector heeft relatief weinig landelijke benchmarks, waardoor besturen meer moeten doen om de kwaliteit van hun scholen te vergelijken met die van andere scholen.

Analyse van de aandachtspunten die per indicator zijn gescoord maakt duidelijk dat de kwaliteitsbewaking bij veel besturen op een tamelijk globaal niveau verloopt. Een deel van de besturen is verder gevorderd en laat meer specifieke activiteiten zien, zoals meer contacten met scholen over de onderwijskwaliteit en meer aandacht voor analyse van resultaten. Bij veel van de onderzochte besturen is het nog wel de vraag waar nu precies aan gewerkt wordt: slechts de helft stelt duidelijke doelen.

De besturen die er recent in slaagden hun zwakke en zeer zwakke scholen te verbeteren, halen in het basisonderwijs, het speciaal basisonderwijs en het voortgezet onderwijs relatief de hoogste beoordeling op de tien indicatoren van bestuurlijk handelen. Ze worden dus zelfs iets gunstiger beoordeeld dan de besturen die de afgelopen jaren in het geheel geen zwakke of zeer zwakke scholen hebben gehad. De besturen die zwakke en zeer zwakke scholen moesten verbeteren, zijn aanvankelijk niet in staat gebleken om te voorkomen dat deze situatie ontstond maar misschien is het zo dat ze, toen ze eenmaal aan de slag moesten met de kwaliteitsverbetering van hun school of scholen, ook hun eigen functioneren hebben aangepakt.

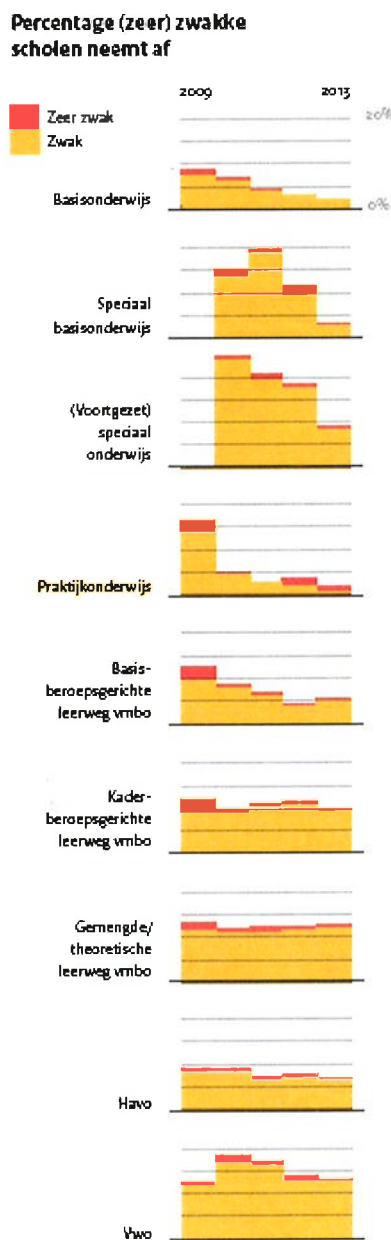
Besturen beschrijven dit fenomeen zelf soms in duidelijke bewoordingen. Na ervaringen met zwakke of zeer zwakke scholen zijn ze meer opbrengstgericht gaan werken, nemen ze signalen van problemen eerder serieus, hebben ze meer contact met de scholen, geven ze scholen feedback en zorgen ze ervoor dat ze betere informatie krijgen. Ze zijn ambitieuzer geworden en stellen hogere eisen aan de scholen, maar ook aan zichzelf (bijvoorbeeld door een hoger opleidingsniveau van bestuurders te vragen). Omdat er geen eerdere gegevens over de besturen zijn van het moment dat de scholen zwak of zeer zwak werden, is niet met zekerheid te zeggen of dit type processen bij veel besturen heeft gespeeld.

Een deel van de besturen schiet tekort in de verantwoording naar belanghebbenden. De informatie in documenten is dan bijvoorbeeld van een te hoog aggregatieniveau. Scholen (of afdelingen of vestigingen) zijn dan niet herkenbaar, terwijl dat voor leerlingen, ouders en anderen nu juist de relevante eenheden zijn. Ook kunnen bestuurders vaak meer werk maken van het bespreken van kwaliteitsgegevens met stakeholders in en om de school, om zo vorm te geven aan een horizontale dialoog.

1 | Deze betreffen: Het bestuur heeft zicht op de kwaliteit van de scholen, evalueert en analyseert de kwaliteit, werkt planmatig aan verbetering, borgt de kwaliteit en verantwoordt zich over de kwaliteit.

3.3 Verbetering onderwijskwaliteit

3.3.1 (Zeer) zwakke scholen



FIGUUR 10 Bron: Inspectie van het Onderwijs, 2015

Een van de manieren om naar de effectiviteit van het toezicht te kijken, is de ontwikkeling in het aantal zeer zwakke scholen te bestuderen. In de afbeelding hiernaast^{1 2} staat de ontwikkeling van het percentage (zeer) zwakke scholen tussen 2009 en 2013. In bijlage 4 is de cijfermatige onderbouwing van deze grafieken opgenomen.

In 2008 is het risicogerichte toezicht daadwerkelijk van start gegaan. Bij sommige scholen, met name in het speciaal onderwijs, is dat een jaar later gebeurd. In de evaluatie van 2010 is duidelijk gemaakt dat het risicogerichte toezicht in eerste instantie leidde tot een toename van het aantal zwakke en zeer zwakke scholen. Dit was een gevolg van de verbetering van de detectie, waardoor scholen eerder werdenesignaleerd. Vanaf dat moment is er evenwel in vrijwel alle onderwijssoorten sprake van een gestage afname van het aantal zwakke en zeer zwakke scholen. Voldeed in 2009 nog bijna negen procent van de regulier basisonderwijs scholen niet aan de basiskwaliteit is dit aan het begin van het school 2013-2014 gedaald naar iets meer als dan twee procent.

Daar zijn twee oorzaken voor. Door de gerichte aanpak en de vroegere signalering van mogelijke problemen, worden minder scholen zeer zwak. En als scholen zeer zwak worden, blijven ze dat door de actieve interventiestrategie, gedurende kortere tijd. Er zijn dan dus op elk moment minder scholen zeer zwak.

Zoals de afbeelding laat zien, zitten er wel verschillen tussen de schoolsoorten. Bij het voortgezet onderwijs is de teruggang van het aantal (zeer) zwakke scholen geringer dan in het basisonderwijs. Daar is het het

praktijkonderwijs dat er duidelijk uitspringt met een zeer substantiële daling.

Op grond van de ontwikkeling van het aantal zwakke en zeer zwakke scholen zou kunnen worden geconcludeerd dat het risicogerichte toezicht zeer effectief is geweest in het waarborgen van de basiskwaliteit.

^{1 2} Afbeelding is ontleend aan het Onderwijsverslag 2012-2013.

In de onderstaande tabel wordt zichtbaar hoe lang scholen er uiteindelijk over hebben gedaan om van zeer zwak naar basistoezicht te gaan. Merk hierbij op dat de N kan afwijken van de N in de vorige tabel, aangezien de optelling van de tussenstappen niet in hetzelfde jaar leidt tot het eindresultaat. Gezien de kleine aantallen worden geen percentage gepresenteerd.

	schooljaar				
	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014
	N=35	N=58	N=40	N=28	N=18
<1 jaar	1	1	0	0	3
1-2 jaar	11	17	10	14	9
>2 jaar	23	40	30	14	6
gemiddelde doorlooptijd	2,4	2,48	2,67	2,24	1,82

Tabel 10 Doorlooptijd van zeer zwakke scholen naar basis^{1 5}

De doorlooptijd van zwakke scholen lijkt een piek te hebben gehad in 2011-2012 maak over een aantal jaren gezien is de doorlooptijd redelijk gelijk gebleven. De doorlooptijd van de zeer zwakke scholen is terug gelopen over de jaren heen. Het lijkt er zelfs op dat een behoorlijk deel van de zeer zwakke scholen zich de laatste jaren binnen een jaar in ieder geval verbetert naar een zwakke school. Hieruit zou men kunnen concluderen dat de inspectie effectief is gebleken om vooral zeer zwakke scholen sneller een verbeterslag te laten maken. Dat kan ertoe hebben bijgedragen dat het totaal aantal is afgenomen.

3.3.3 *Kwaliteit onderwijsproces*

Worden scholen beter beoordeeld? Is de kwaliteit van het onderwijsproces op scholen toegenomen?

^{1 5} Het gaat hier om een traject van zeer zwak naar basis (ongeacht of dit gebeurt via een tussenstap zwak of niet)

Hieruit zou kunnen worden afgeleid dat het toezicht als het gaat om het onderwijsleerproces niet effectief is geweest.

Onderwijsproces VO

TABEL 2.2A

Percentage afdelingen waar de kwaliteit van het onderwijsproces en de kwaliteitszorg voldoende is in 2012/2013 (n=123)

	Percentage voldoende
Schoolklimaat	
De school kent een op ondersteuning en begeleiding gerichte cultuur.	99
Didactisch handelen	
De onderwijsactiviteit heeft een doelgerichte opbouw.	96
De leerlingen krijgen een begrijpelijke uitleg.	99
De leerlingen zijn actief betrokken.	82
Afstemming	
De leraren gebruiken bij de vormgeving van hun onderwijs de analyse van de prestaties van de leerlingen.	25
De leerlingen krijgen effectieve feedback op hun leerproces.	53
De leraar stemt de instructie af op verschillen tussen leerlingen.	30
De leraar stemt de verwerking af op verschillen tussen leerlingen.	11
Ondersteuning en begeleiding	
De school volgt systematisch de vorderingen van de leerlingen aan de hand van genormeerde toetsen.	68
De school bepaalt wat de onderwijs- en ondersteuningsbehoefte is van individuele of groepen leerlingen.	92
De school voert de ondersteuning planmatig uit.	82
De school heeft doelen gesteld die erop gericht zijn om achterstanden te bestrijden.	51
De leraren volgen en analyseren systematisch de voortgang in de ontwikkeling van de leerlingen.	86
De school evalueert regelmatig de effecten van de zorg.	61
Kwaliteitszorg	
De school evalueert systematisch de opbrengsten.	93
De school werkt doelgericht aan de kwaliteit van de opbrengsten.	79
De school evalueert systematisch het onderwijsproces.	72
De school werkt doelgericht aan de verbetering van het onderwijsproces.	88
De school borgt de kwaliteit van het onderwijsproces.	48

Onderwijsproces nauwelijks verbeterd

De inspectie constateert dat de didactische basis van het onderwijzend personeel in ongeveer 80 procent van de lessen op orde is: opbouw, uitleg en taakgerichtheid zijn voldoende. Leraren weten echter in de lessen vaak nog onvoldoende te differentiëren naar de leerbehoeften van de leerlingen. Ze houden te weinig rekening met wat bij de leerlingen qua kennis en vaardigheden al bekend is. Alle leerlingen in de klas krijgen dezelfde instructie en moeten dezelfde opdrachten maken. Dat maakt sommige onderdelen van het onderwijsproces voor een deel van de leerlingen weinig effectief. De mate waarin de leraren didactisch vaardig zijn, verschilt per school.

3.3.4

Leerprestaties

Leerlingen in het basisonderwijs presteerden in de meeste vakken gelijk of iets minder goed dan in 2012. Bij begrijpend lezen en woordenschat haalden ze het hetzelfde niveau. In spelling en rekenen zijn ze iets achteruitgegaan. De daling van het rekenniveau geldt voor alle onderdelen van het vak rekenen. Dit blijkt uit

veel bemoeienis van buiten hun eigen kwaliteit zouden kunnen verbeteren. De inspectie was er vooral voor de borging van de basis.

Die verbetering aan de bovenkant blijkt zich nauwelijks voor te doen. Er bestaat een grote groep scholen en opleidingen waarvan de kwaliteit van het onderwijsleerproces gemiddeld is. Deze scholen verschillen op dit gebied nauwelijks van elkaar. Ook zijn er weinig verschillen in leerprestaties en in leerwinst (de groei in prestaties die leerlingen doormaken). Op deze scholen waren in de afgelopen jaren bijna geen verbeteringen zichtbaar in het onderwijsproces. Ook toetsscores, rendementen, examenresultaten en slaagpercentages zijn nauwelijks omhoog gegaan. Deze groep scholen en opleidingen laat dus niet dezelfde kwaliteitsverbetering zien als de (zeer) zwakke scholen en opleidingen.

Op stelselniveau stagneert hierdoor de kwaliteitsontwikkeling. De verbetering van de kwaliteit van het onderwijsproces van de afgelopen jaren heeft zich vooral aan de onderkant, bij de zwakke en zeer zwakke scholen voorgedaan. Nu dat aantal sterk is gereduceerd, kan verbetering van het onderwijs als geheel alleen komen van de scholen boven de streep en die ontwikkeling zet niet verder door. Op enkele onderdelen van de onderwijskwaliteit is er zelfs sprake van achteruitgang.

Bij deze grote groep scholen en opleidingen met een gemiddelde kwaliteit stelt de inspectie vrijwel altijd één of enkele tekortkomingen vast. Het is voor scholen en opleidingen niet vanzelfsprekend dat de belangrijkste kwaliteitsindicatoren allemaal als voldoende worden beoordeeld. In het basisonderwijs zijn op 46 procent van de scholen alle normindicatoren van het onderwijsproces voldoende. In het voortgezet onderwijs gaat het in de meeste leerwegen om zo'n 20 procent. De tekortkomingen maken dat het onderwijs op al deze onderdelen niet of nauwelijks verbetert. Dit ondanks de kwaliteitsinitiatieven in de meeste sectoren en het stimuleren van opbrengstgericht werken.

In alle onderwijssectoren zijn voorbeelden van scholen en opleidingen die aan alle normindicatoren voldoen: zij hebben geen tekortkomingen. In het basisonderwijs gaat het om iets minder dan de helft van de scholen, in het voortgezet onderwijs om een op de vijf afdelingen. Dit aandeel is de afgelopen jaren niet gestegen en staat in het basisonderwijs onder druk. Het aandeel scholen met hoge opbrengsten is in het basisonderwijs en in het voortgezet onderwijs nog kleiner. Slechts 9 procent van de basisscholen heeft een gemiddelde score op de Eindtoets Basisonderwijs boven de 540. Dit percentage daalt. In het voortgezet onderwijs heeft ook slechts een paar procent van de scholen een hoog gemiddeld examencijfer: 7,5 of hoger. Dit percentage daalde de afgelopen jaren, maar is in 2012 weer gestegen.

3.3.6

Maatschappelijk vertrouwen

De inspectie doet niet zelf onderzoek naar het maatschappelijk vertrouwen in het onderwijs. Het PISA-onderzoek, dat elke drie jaar wordt gehouden, is ook geen onderzoek naar het maatschappelijk vertrouwen in het onderwijs, maar fungeert de laatste jaren wel als een belangrijk ankerpunt voor berichtgeving die dat vertrouwen beïnvloedt. In de pers wordt het gebruikt als een meting van de voor- of achteruitgang van Nederland in internationaal opzicht, en als zodanig beïnvloedt het het maatschappelijk vertrouwen.

Dat beeld is er het afgelopen decennium een geweest van een hoge positie op de wereldranglijst die geleidelijk erodeerde. De opkomst van Aziatische landen en de hoge positie van Finland op de ranglijsten, waren daarbij belangrijke elementen. De

4 Conclusie

4.1 De theorie achter het vernieuwde toezicht

Het vernieuwde toezicht past in een bredere beleidstheorie van governance waarin de kwaliteit van het onderwijs geborgd, respectievelijk verbeterd moet worden. Deze beleidstheorie is beschreven in paragraaf 2.1. Centraal daarin staat het bevoegd gezag dat verantwoordelijk is voor die kwaliteit en daarover verantwoording aflegt aan de samenleving. Die samenleving is in het model onderverdeeld naar de directe belanghebbenden bij het onderwijs en naar de inspectie als representant van de overheid.

De eerste laag van borging is gelegen bij het bevoegd gezag zelf. Het bevoegd gezag onderhoudt daartoe zelf allereerst een systeem van kwaliteitszorg waarin de relevante prestaties van de scholen of instellingen worden gemeten en worden getoetst aan doelstellingen die daartoe zijn geformuleerd. Als afwijkingen worden geconstateerd, dient dat te leiden tot verbeteracties. Naast dit systeem van kwaliteitszorg opereert het intern toezicht op het functioneren van het bevoegd gezag. De belanghebbenden vormen de tweede laag van het model. Zij zijn degenen die gebruik maken van het onderwijs of te maken krijgen met de resultaten ervan. In het model van governance vervullen zij een essentiële rol aangezien zij, beter dan een centrale overheid, in staat worden geacht het resultaat van het functioneren van het bevoegd gezag te beoordelen en te beïnvloeden. Zij hebben er immers direct belang bij en kunnen zelf schade ondervinden als de kwaliteit tekortschiet. Zij bieden bovendien een nabije referentie voor het bevoegd gezag waardoor een variëteit mogelijk is die via centrale richtlijnen niet te realiseren is. De derde laag, of het sluitstuk, van de borging wordt gevormd door de centrale overheid in de vorm van algemeen geldende voorwaarden waarop door de inspectie toezicht wordt gehouden. De inspectie is het sluitstuk waar zij optreedt als borgingsmechanisme op het moment dat de beide eerste lagen daar kennelijk niet in zijn geslaagd. De inspectie heeft daarnaast een eigenstandige rol waar zij de minister in staat stelt zijn verantwoordelijkheid waar te maken.

De inspectie kan scholen tot kwaliteitsverbetering brengen. Cruciaal hierbij zijn een frequente en gerichte analyse van onafhankelijke bronnen (al dan niet in combinatie met de verantwoording van besturen), selectieve en proportionele bezoeken met een waarderingskader en interventies en openbare oordelen. In dit conclusiehoofdstuk worden conclusies getrokken over bestaand wetenschappelijk en nieuw praktijkbewijs voor deze werkwijze.

De inspectie kan daarnaast op meer indirecte wijze bijdrage aan kwaliteitsverbetering van scholen. Zo kan de inspectie stimuleren dat het bestuur tot meer en effectiever eigenstandig handelen komt, waar het kwaliteit van het onderwijs betreft (al dan niet door het interne toezicht hiertoe aangezet). De bestuursinstek in het vernieuwde toezicht beoogt de rol van de besturen dan ook nadrukkelijk te versterken en hiermee op een indirecte wijze tot kwaliteitsverbetering te komen. Ook kan de inspectie de rol van belanghebbenden als ouders en leerlingen te versterken. Versterking van hun rol stimuleert de besturen tot kwaliteitsverbetering, aldus het model. In deze conclusie wordt ook ingegaan op de mate waarin wetenschap en praktijk op dit punt.

signalen en andere info) in combinatie met de eigen verantwoording van het bestuur.

De keuze van de (risico)indicatoren is cruciaal in de risicoanalyse. Wil selectiviteit leiden tot een grotere efficiency, dan moet gebruik gemaakt worden van een beperkt aantal, relatief eenvoudig te verkrijgen indicatoren die voldoende betrouwbaar en informatief zijn over de relevante prestaties van het toezichtobject. De toetsing van de indicatoren aan de waargenomen werkelijkheid, ook bij andere dan risicovolle objecten, laat zien dat de keuze voor de risicoindicatoren goed lijkt. Tegelijkertijd zijn veel indicatoren vooral indicaties van onderpresteren. Hiermee loopt de toezichthouder wel de kans dat risico's relatief laat in beeld komen. Van vroegtijdige signalering is nog niet altijd sprake. Waar mogelijk zal de risicoanalyse met indicatoren moeten worden aangevuld waarmee ook vroegtijdige signalering mogelijk wordt. De toezichthouder zal daarnaast alert moeten blijven op strategisch gedrag omtrent deze risicoindicatoren, omdat de wetenschap ons leert dat een deel van de scholen de neiging heeft ze in positieve zin te beïnvloeden.

4.4 Proportioneel en selectief toezicht

Het proportionele en selectieve toezicht heeft geleid tot een belangrijke efficiencyslag voor de inspectie en op scholen, zo mogen we op basis van deze evaluatie concluderen. Op de scholen is sprake van een reductie van de toezichtlast, en ook de inspectie werkt efficiënter. Tegelijkertijd zien we een verschuiving van de lasten op scholen. Gerichter toezicht leidt tot minder inspecties op goede scholen en tot een halvering van de toezichtlast op deze scholen. Toezicht op de (zeer) zwakke scholen is geïntensiveerd. Dit betekent meer last voor deze scholen (in het PO met 40%).

De vraag is of het proportionele en selectieve toezicht, naast efficiënter, ook effectiever is. In theorie is toezicht het meest effectief op scholen die niet aan de norm voldoen. Onderzoek toont aan dat toezicht het meeste effect heeft op de kwaliteit van het onderwijs waar die kwaliteit tekortschiet. Een vorm van selectiviteit die kijkt naar indicaties van tekortschietende kwaliteit, leidt er dus toe dat de toezichthouder zijn inspanningen vooral op die plaatsen richt waar leerlingen de grootste risico's lopen, maar waar ook de gevolgen van die inspanningen het grootst zijn. Bij gelijke inzet, is risicogericht toezicht dus effectiever dan andere vormen van toezicht. Tegelijkertijd is er in Nederland een taakstelling van 20% voor de inspectie. Het netto-effect van deze twee verschijnselen hebben we nog niet in kaart kunnen brengen. Deze evaluatie levert geen indicaties op voor afname van de effectiviteit. We mogen dan ook concluderen dat, met een afname van de totale toezichtinzet (=kosten), het resultaat (= opbrengsten) gelijk is gebleven. Dat is evident een verbetering van de efficiency.

Concluderend mogen we dus stellen dat de huidige Nederlandse praktijk duidelijk laat zien dat risicogericht toezicht inderdaad efficiënter is; de toezichtlast voor de scholen is verminderd en de inspectie werkt efficiënter. De ontwikkeling van het aantal zwakke en zeer zwakke scholen laat zien dat het, naast efficiënter, ook effectiever is geworden. Die effectiviteit blijft evenwel beperkt tot juist die categorie van scholen. Scholen die voldoen aan de norm van basiskwaliteit, blijken zich niet verder te verbeteren.

4.5 Meer en sterkere interventies

Met de introductie van het risicogerichte toezicht, is ook het interventierepertoire uitgebreid. Traditioneel waren het openbaar oordeel en het opvoeren van de

stimulerende werking van het toezicht voor het risicogerichte toezicht, heeft dat niet kunnen verhinderen.

Met de ontwikkeling van Toezicht 2020 streeft de inspectie ernaar de komende jaren ook de groep boven de norm concreter aan te spreken. Het blijft een interessante vraag of toezicht ertoe kan bijdragen dat onder toezicht gestelden die aan de norm voldoen, zich toch verder gaan verbeteren.

Strategisch gedrag

Een probleem bij het risicogerichte toezicht is nog wel dat de keuze voor risico-indicatoren leidt tot gedragsbeïnvloeding bij de onder toezicht gestelden. Die kunnen onder meer proberen de resultaten op te poetsen of zich, met verwaarlozing van andere, exclusief te richten op die indicatoren. Het is de vraag of de toezichthouder voldoende in staat is middels het naast elkaar leggen van verschillende indicatoren dit te ondervangen? Dit zal continue aandacht moeten hebben. Daarnaast lijkt het verstandig te streven naar verbetering van de risicoanalyse, met name uitbreiding met indicatoren van tussentijdse leerwinst en andere meer vroegtijdige indicaties van onderpresteren.

Reductie maatschappelijke risico's

De wijze waarop de inspectie het risicogerichte toezicht heeft ingevuld, wordt in de literatuur 'objectgerichte risicoanalyse' genoemd (De Wolf, 2007). Hiermee kunnen scholen geprioriteerd worden, afhankelijk van hun risiconiveau. Dit betekent evenwel niet dat hiermee ook automatische de maatschappelijke risico's in het onderwijs verminderd worden. Zo leidt het risicogerichte toezicht niet automatisch tot minder voortijdig schoolverlaters. Sparrow (2000) raadt daarom aan ook na te gaan waar de belangrijkste maatschappelijke risico's liggen, en pleit ervoor de risicoanalyse zo in te vullen dat deze ook bijdraagt aan risicoreductie op dit terrein. Hoewel anderen betwijfelen of dit kan (Power, 2006) is het raadzaam de relatie tussen beide soorten risico's eens nader te verkennen en te zoeken naar een vorm van risicogericht toezicht waarmee ook de maatschappelijke risico's gereduceerd kunnen worden. De inspectie heeft voor dit laatste een nieuw traject gestart, namelijk vernieuwing van het stelseltoezicht. Het is nog te vroeg om hier resultaten van te zien.

Methodiek evaluatie onderzoek

Met dit onderzoek hebben we ook laten zien hoe toezicht geëvalueerd kan worden en hoe waardevol een dergelijke evaluatie van het toezicht is. Door een reconstructie van de theorie achter het toezicht en een wetenschappelijke en praktijkevaluatie hiervan, hebben we sterke en zwakke punten in het onderwijstoezicht in kaart gebracht. Opvallend hierbij is de overeenstemming tussen het wetenschappelijke bewijs en de resultaten van de praktijkevaluaties. De evaluatieresultaten van beide onderdelen lijken sterk op elkaar. Tegelijkertijd zijn er laten beide onderdelen andere accenten zien. Een combinatie van wetenschappelijke en praktijkevaluatie heeft dan ook onze voorkeur.

Tot slot is het nog de vraag hoe generaliseerbaar deze resultaten zijn. Zijn ze ook bruikbaar voor andere sectoren en andere toezichthouders? Dit is ons inziens het geval, vooral omdat ze inzicht geven in de mechanismen achter de werkwijze. De wetenschappelijke inzichten hierover zijn per definitie breder geldig dan alleen het onderwijstoezicht, terwijl de inzichten uit de praktijk wellicht anders zullen uitpakken in andere (toezicht)velden.

- Inspectie van het Onderwijs (2014a) Resultaten tevredenheidsonderzoek Inspectie van het Onderwijs 2012/2013
- Inspectie van het Onderwijs (2014b) Tevredenheidsonderzoek Primair Onderwijs schooljaar 2012/2013
- Inspectie van het Onderwijs (2014c) Tevredenheidsonderzoek Expertisecentra schooljaar 2012/2013
- Inspectie van het Onderwijs (2014d) Tevredenheidsonderzoek Voortgezet Onderwijs schooljaar 2012/2013
- Inspectie van het Onderwijs (2014e) Bevragingslast: stand van zaken 2013
- Inspectie van het Onderwijs (2012) Tevredenheidsonderzoek besturen primair onderwijs schooljaar 2011/2012
- Inspectie van het Onderwijs (2013) De kwaliteit van het bestuurlijk handelen in het funderend onderwijs
- Internetspiegel (2009a) Rapportage evaluatie-onderzoek Inspectie van het Onderwijs (primair onderwijs). Onderzoek naar de beleving van kwaliteit, effectiviteit, tevredenheid, vertrouwen en klachten.
- Internetspiegel (2009b) Rapportage evaluatie-onderzoek Inspectie van het Onderwijs (voortgezet onderwijs). Onderzoek naar de beleving van kwaliteit, effectiviteit, tevredenheid, vertrouwen en klachten.
- Internetspiegel (2008) Inspectie van het Onderwijs tevredenheidsonderzoek besturen Primair Onderwijs
- Internetspiegel (2008) Inspectie van het Onderwijs tevredenheidsonderzoek besturen Voortgezet Onderwijs.
- J/M voor ouders, 2009, Onderzoek vertrouwen in onderwijs onder ouders
- J/M voor ouders, 2012, Onderzoek vertrouwen in onderwijs onder ouders
- Jacob, B.A. & Levitt, S.D. (2003) Rotten apples, an investigation of the prevalence and predictors of teacher cheating. In: *The Quarterly Journal of Economics*, 118(3), 843 tm 877.
- Janssens, F.J.G. & De Wolf, I.F. (2009). Analysing the assumption of a policy programme: an ex-ante evaluation of Educational Governance in the Netherlands. In: *American Journal of Evaluation*, 30(3), 411-425.
- Janssens & De Wolf (2007) Effectmeting voor de Inspectie van het Onderwijs. Keynote-lezing op de Onderwijsresearchdagen.
- Janssens, F.J.G. (2006). *De inspectie komt op bezoek*. Alphen a/d Rijn: Kluwer.
- Jos de Jonge (2008) [artikel in NTOR]
- Kordes, J. Bolsinova, M. Limpens, G. Stolwijk, R. (2013) *Resultaten Pisa 2012. Praktische kennis en vaardigheden van 15-jarigen*.
- Leeuw, F.L. (2008) *Gedragmechanismen achter overheidsinterventies en rechtsregels*. Inaugurele rede, Universiteit Maastricht.
- Leeuw, F. & Willemsen, F. (2006). Toezicht en inspectie: trends, kosten en baten. *NJB*, 37, 2108-2116.
- Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, 24, 5-20.

- Webbink, De Wolf, Woessmann, van Elk, Minne & vd Steeg (2009) Wat is bekend over de effecten van kenmerken van onderwijsstelsels; een literatuurstudie. CPB-rapport.
- Weiss, C.H. (2000) Which links in which theories shall we evaluate? *New Directions for Evaluation*, 87, 35-45.
- Wetenschappelijke Raad voor het Regeringsbeleid. (2013). *Toezen op publieke belangen. Naar een verruimd perspectief op rijkstoezicht*. (No. 89). Den Haag/Amsterdam: Amsterdam University Press.
- Wetenschappelijke Raad voor het Regeringsbeleid. (2014). *Van tweeluik naar driehoeken versterking van interne checks and balances bij semipublieke organisaties*. (). Den Haag/Amsterdam: Amsterdam University Press.
- Willemsem, F. & Leeuw, F. (2007) Maatschappelijke effecten van toezicht en inspectie. Effectevaluatie uit de praktijk. In: Leeuw, Kerseboom en Elte (eds) *Turven, tellen, toetsen; over toezicht, inspectie, handhaving en evaluatie in hun maatschappelijke betekenis in Nederland*. BJU.
- WODC (2010) 'Toezicht en effectmeting, het kan' Researchsynthese van onderzoek naar de effectiviteit van toezichtinterventies.
- Wolf, I.F. de & Janssens, F.J.G. (2007). Effects and side effects of school inspections and accountability in education: a review of empirical studies. In: *Oxford Review of Education*. 33, 3, 379-396.
- Wolf, I.F. de (2007) Risicoanalyse en risicomanagement bij toezichthouders. In: Leeuw, Kerseboom en Elte (eds) *Turven, tellen, toetsen; over toezicht, inspectie, handhaving en evaluatie in hun maatschappelijke betekenis in Nederland*. BJU.

- Als de externe toezichthouder intervenieert leidt dat ertoe dat het bestuur verbeteringen aanbrengt in de vastgestelde tekortkomingen
- Als belanghebbenden kritiek uiten, leidt dat ertoe dat het bestuur verbeteringen aanbrengt in de aangegeven tekortkomingen
- Als belanghebbenden stemmen met de voeten, leidt dat ertoe dat het bestuur verbeteringen aanbrengt in de aangegeven tekortkomingen

Belanghebbenden

Wensen/klachten/ervaringen

- Als belanghebbenden wensen hebben, uiten ze deze aan de instelling
- Als belanghebbenden klachten hebben, uiten ze deze aan de instelling
- Als belanghebbenden klachten hebben, uiten ze deze aan de toezichthouder

Voice/stemmen met de voeten

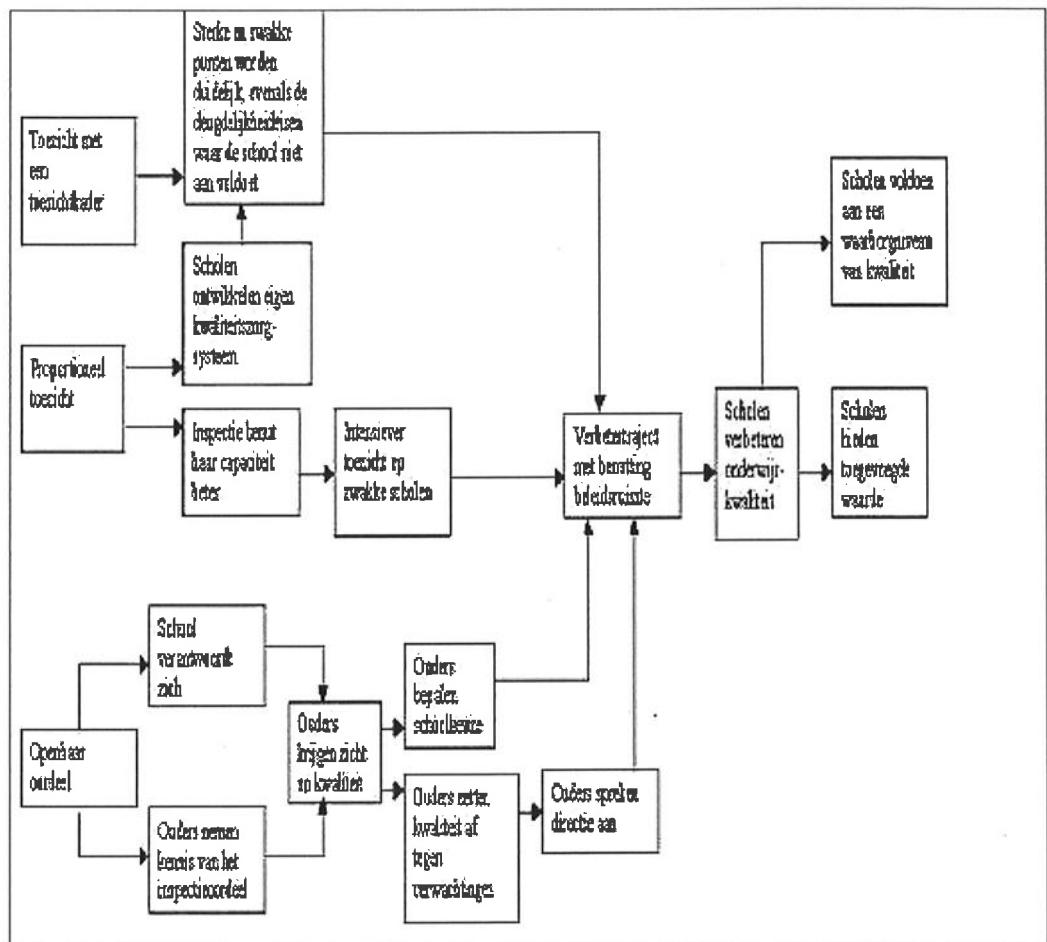
- Als aan behoeften van belanghebbenden niet wordt voldaan, melden zij zich niet aan.
- Als aan behoeften van belanghebbenden niet wordt voldaan, spreken zij het bestuur aan of vertrekken
- Als uit de externe verantwoording tekortkomingen blijken, spreken de belanghebbenden het bestuur aan of vertrekken
- Als uit een openbaar oordeel/interventie van de externe toezichthouder tekortkomingen blijken, spreken de belanghebbenden het bestuur aan of vertrekken

Overheid/toezichthouder

Proportioneel/selectief toezicht

- Algemeen: Als de toezichthouder proportioneel/selectief toezicht wil houden, zal hij moeten beschikken over een adequaat instrument van risicoanalyse.
- Als de toezichthouder proportioneel/selectief toezicht houdt, zal hij efficiënter en effectiever zijn.
- Als de toezichthouder proportioneel/selectief toezicht wil houden, maakt hij maximaal gebruik van de externe verantwoording van het bestuur
- Als de toezichthouder proportioneel/selectief toezicht wil houden, zal het instrument de risico's moeten kunnen vaststellen
- Als de toezichthouder proportioneel/selectief toezicht wil houden, zullen de vastgestelde risico's ook relevant moeten zijn voor onderwijskwaliteit of naleving.
- Als de toezichthouder proportioneel/selectief toezicht wil houden, zal hij het instrument van de risicoanalyse regelmatig moeten ijken
- Als de toezichthouder de externe verantwoording van het bestuur op waarde wil kunnen schatten, gebruikt hij daarvoor onafhankelijke bronnen
 - Bronnen zijn onafhankelijk als ze gegevens opleveren die niet gecontroleerd worden door het te beoordelen object
 - Signalen van belanghebbenden en de omgeving maken deel uit van de onafhankelijke bronnen
- Als de toezichthouder proportioneel/selectief toezicht houdt, gebruikt hij onafhankelijke bronnen als detectiemiddel
- Als de toezichthouder proportioneel/selectief toezicht houdt, benut hij zijn capaciteit beter
 - Toezicht wordt vooral gericht op scholen met tekortkomingen
 - Toezicht op scholen met tekortkomingen is effectiever dan op scholen zonder tekortkomingen

BIJLAGE 2. THEORIE WOT2001 door Ehren et.al (2006)



Bron: Ehren, Leeuw & Scheerens (2006)

BIJLAGE 3: CONTEXT VAN (DE WET OP) HET ONDERWIJSTOEZICHT

Voor 2001 was het werk van de inspectie van het onderwijs geregeld in onderdelen van verschillende, zogenaamde sectorwetten (wetten die betrekking hebben op een onderwijssector). Met het totstandkomen van de Wet op het Onderwijstoezicht (WOT) in 2001 zijn al die verspreide voorschriften gebundeld in een wet. Deze is in 2012 herzien.

De inspectie van het onderwijs houdt toezicht op de kwaliteit van het onderwijs en de naleving van wet- en regelgeving door scholen. In de herziene WOT ziet in artikel 3 het takenpakket er uit als volgt:

Artikel 3. Taken

1. Het toezicht op het onderwijs is opgedragen aan de inspectie.
2. Het toezicht omvat de volgende taken:
 - a. het beoordelen en bevorderen van de kwaliteit van het onderwijs, daaronder begrepen de kwaliteit van het onderwijspersoneel, aan instellingen, als bedoeld in de onderwijswetten met uitzondering van de Wet op het hoger onderwijs en wetenschappelijk onderzoek, en het beoordelen en bevorderen van de kwaliteit van de uitoefening van de taken van de regionale expertisecentra en de kenniscentra beroepsonderwijs bedrijfsleven,
 - b. het beoordelen en bevorderen van de naleving van de bij of krachtens een onderwijswet gegeven voorschriften,
 - c. het beoordelen en bevorderen van de kwaliteit van het stelsel voor hoger onderwijs, met inbegrip van het stelsel van accreditatie, bedoeld in artikel 1.1, onderdelen s en t, van de Wet op het hoger onderwijs en wetenschappelijk onderzoek,
 - d. het beoordelen en bevorderen van de financiële rechtmatigheid door in ieder geval het verrichten van onderzoek naar de rechtmatige verkrijging van de bekostiging, naar de controlerapporten van de door het bestuur aangewezen accountant, naar de rechtmatigheid van de bestedingen en naar de rechtmatigheid van het financieel beheer van de bekostigde instellingen,
 - e. het beoordelen en bevorderen van de rechtmatigheid en de doeltreffendheid van de uitoefening van de taken, opgedragen aan het college van burgemeester en wethouders bij of krachtens de hoofdstukken 3 en 6 van de Wet kinderopvang,
 - f. het rapporteren over de ontwikkeling van het onderwijs en over de uitoefening van de taken door de instellingen, de regionale expertisecentra en de kenniscentra beroepsonderwijs bedrijfsleven, in het bijzonder over de kwaliteit daarvan, en
 - g. het verrichten van andere bij of krachtens de wet aan de inspectie opgedragen taken.
3. Bij algemene maatregel van bestuur kan worden bepaald dat Onze Minister de inspectie mandaat kan verlenen om bestuurlijke sancties op te leggen en andere besluiten te nemen.

De activiteiten die onderzocht worden in dit rapport, zijn die in lid 2 onder a en b. Het onderwijsveld waar de inspectie toezicht op houdt, kent vijf sectoren. Dit zijn

BIJLAGE 4: ONTWIKKELING VAN HET AANTAL ZWAKKE EN ZEER ZWAKKE SCHOLEN

Basisonderwijs

Tabel 4.1 Toezichtarrangementen basisonderwijs op 1 januari 2009 en op 1 september 2010, 2011, 2012 en 2013 (in percentages, n 2013=7023)

	1 januari 2009	1 september 2010	1 september 2011	1 september 2012	1 september 2013
n	7199	7173	7160	7107	7023
Basis	91,1	93,0	95,6	96,9	97,8
Zwak	7,4	6,0	3,8	2,8	2,0
Zeër zwak	1,5	1,0	0,6	0,2	0,2
Totaal	100	100	100	100	100

Bron: Inspectie van het Onderwijs, 2013

Voortgezet onderwijs

Tabel 4.2 Toezichtarrangementen voortgezet onderwijs op 1 september 2009, 2010, 2011, 2012 en 2013 (in percentages, n 2013 = 2730)

	1 september 2009	1 september 2010	1 september 2011	1 september 2012	1 september 2013
n	2777	2598	2721	2715	2730
Basis	87,9	88,2	88,9	89,7	89,9
Zwak	10,3	10,7	10,2	9,4	9,5
Zeër zwak	1,8	1,1	0,9	0,9	0,6
Totaal	100	100	100	100	100

Bron: Inspectie van het Onderwijs, 2013

Voortgezet onderwijs naar onderwijssoort

Tabel 4.3 Toezichtarrangementen **basisberoepsgerichte leerweg vmbo** op 1 september 2009-2013

	1 september 2009		1 september 2010		1 september 2011		1 september 2012		1 september 2013	
	n=496		n=444		n=480		n=471		n=470	
	n	%	n	%	n	%	n	%	n	%
Basis	434	87,5	406	91,4	448	93,3	451	95,8	443	94,3
Zwak	48	9,7	34	7,7	28	5,8	17	3,6	24	5,1
Zeër zwak	14	2,8	4	0,9	4	0,8	3	0,6	3	0,6
Totaal	496	100	444	100	480	100	471	100	470	100,0

Bron: Inspectie van het Onderwijs, 2013