

# DRAAGT ONDERWIJSTOEZICHT BIJ AAN SCHOOLVERBETERING?

EFFECTEN VAN ONDERWIJSTOEZICHT IN HET PRIMAIR ONDERWIJS<sup>1</sup>

Frans J.G. Janssens  
Universiteit Twente  
Inge de Wolf  
Universiteit van Amsterdam

## 1. INLEIDING

Het overheidstoezicht voert inspecties uit en verzamelt hierbij informatie over de vraag of een handeling of zaak voldoet aan de daaraan gestelde wettelijke eisen, vormt zich hierover een oordeel en intervenueert als niet aan de wettelijke eisen is voldaan (Janssens, 2005). Wat de politiek, het beleid, de maatschappij en ook de inspecties vervolgens willen weten, is hoe succesvol deze inspectieactiviteiten zijn geweest. Dat een inspectie zich alleen verantwoordt over de aantallen gerealiseerde toezichtcontacten en daaruit voortvloeiende inspectieproducten, wordt veelal als ontoereikend ervaren (Leeuw & Willemsen, 2006). Relevanter is de vraag wat het effect is geweest op de handhaving van wet- en regelgeving en de kwaliteit van de dienstverlening waarop toezicht wordt gehouden. Dit laatste is recent onderstreept door de Algemene Rekenkamer (2005), die in haar rapport 'Handhaven en Gedogen' stelt dat inspecties de effectiviteit van hun (handhavings)beleid veel meer zouden moeten meten.

Ook voor de Onderwijsinspectie is kennis van de effectiviteit van het onderwijstoezicht van belang voor de verbetering van het toezicht en voor het afleggen van verantwoording over de besteding van middelen. Maar wat dient onder effect van toezicht te worden verstaan? De Leidraad Effectmeting bij inspecties definieert 'effect' als: *Een waarneembare verandering in hetgeen met de toezichts- en handhavingsactiviteiten van de inspectie wordt beoogd als gevolg van de activiteiten van die inspectie* (Vossen, 2005, p. 2).

De activiteiten die de onderwijsinspectie uitvoert zijn gekoppeld aan haar wettelijke taken, waarvan de belangrijkste zijn: de beoordeling en bevordering van de kwaliteit van scholen en de rapportage over de staat van het onderwijs. Het belangrijkste middel om deze taken uit te voeren is het doen van onderzoek bij scholen op basis van een waarderingskader dat de kwaliteitsaspecten en wettelijke voorschriften bevat die de inspectie in haar onderzoek betreft. Naarmate het onderwijs niet voldoet aan de indicatoren waarmee deze kwaliteitsaspecten en deugdelijkheidseisen worden beoordeeld, volgen interventies die erop gericht zijn het onderwijs in overeenstemming te brengen met de bedoelingen (Janssens, 2006).

De 'waarneembare veranderingen' uit bovenstaande definitie, de effecten van onderwijstoezicht, hebben in het geval van de onderwijsinspectie betrekking op verbeteringen in het gedrag van schoolbesturen, directies en leraren en op verbetering van de kwaliteit van

---

<sup>1</sup> Keynote ten behoeve van het thema Beleid & Organisatie, door de eerste auteur gepresenteerd op de ORD 2007 te Groningen

de onderwijsprocessen en de leerprestaties als gevolg van toezichtactiviteiten (De Wolf & Janssens, 2005; Algemene Rekenkamer, 2005 en Ehren, 2006).

Effecten van onderwijstoezicht kunnen dus worden opgevat als verbeteringen van de kwaliteit van de school die tot uitdrukking komen in de kwaliteit van het onderwijs, met als afhankelijke variabelen: onderwijsprocessen en de leerprestaties (Ehren, 2006, p. 15). Het is echter de vraag of het toezicht voldoende impulsen kan geven aan schoolverbetering en ervoor kan zorgen dat scholen effectiever worden. Vooral in Engeland is onderzoek gedaan naar de effecten van toezicht. De Wolf en Janssens (2005) omschrijven echter een aantal methodologische problemen die aan de meeste van deze onderzoeken ten grondslag liggen. De onderzoeksmethoden die zijn gebruikt om relaties tussen toezicht en effecten vast te stellen maken veelal geen gebruik van experimentele of quasi-experimentele designs, met uitzondering van Rosenthal (2004), voor de Engelse, en Luginbuhl, Webbink & De Wolf (2007) voor de Nederlandse situatie. De causaliteit van vastgestelde relaties tussen toezicht en effecten is dus niet onomstreden. Niet alle onderzoeksresultaten bieden bovendien inzicht in effecten zoals hier geformuleerd, namelijk verbetering van onderwijsprocessen en van de leerresultaten van leerlingen. Veelal zijn de resultaten beperkt tot verbeteractiviteiten en gedragsverandering van leerkrachten en schoolleiders, meestal gerapporteerd door de betrokkenen zelf (zie bijv. De Wolf & Janssens, 2005; Ehren, 2006)

Schoolbezoeken van de Engelse inspectie, waarvan de werkwijze veel lijkt op die van de Nederlandse inspectie, lijken invloed te hebben op het gedrag van leerkrachten en directies. Leerkrachten rapporteren over wijzigingen in de manier van lesgeven als gevolg van een schoolbezoek, en directeuren zeggen maatregelen te nemen om de kwaliteit van het onderwijs of de organisatie van de school te verbeteren. Het aantal maatregelen varieert echter sterk, evenals de inhoud van de maatregelen. De inhoud varieert van het vergroten van de betrokkenheid van leerkrachten bij de school, het opstellen van gedragsregels voor leerlingen, het nemen maatregelen om examenresultaten te verhogen, het verbeteren van het pedagogisch-didactisch handelen, het wijzigen van de wijze van toetsing en evaluatie van leerlingen tot het verbeteren van het leerstofaanbod en de differentiatie van het onderwijs (Ehren, 2006; Matthews & Sammons, 2004 en McCrone, Rudd, Blekinsop & Wade, 2006).

In dit paper wordt geprobeerd een bijdrage te leveren aan de kennis over de werking van het Nederlandse onderwijstoezicht op basis van zowel gegevens van de onderwijsinspectie als de uitkomsten van wetenschappelijk onderzoek. Omdat er naar het effect van toezicht op scholen voor primair onderwijs het meeste onderzoek is gedaan, beperk ik me in dit verband tot deze sector. De periode waarover wordt gerapporteerd betreft vooral de jaren 1998 tot en met 2006. In deze periode heeft de onderwijsinspectie vanaf 2002 het toezicht uitgeoefend op basis van de Wet op het Onderwijstoezicht (WOT) die in 2007 wordt geëvalueerd. Daarom is het een voor de hand liggend moment om de effecten van het onderwijstoezicht te evalueren.

In de WOT zijn de taken van de onderwijsinspectie beschreven, de inhoud van het toezicht, de werkwijze van de inspectie en de daarbij behorende onderzoeksvormen. De inspectie kreeg een uitgebreidere taak. De inspectie houdt niet alleen meer toezicht op naleving van wettelijke voorschriften door scholen, maar beoordeelt ook andere aspecten van de kwaliteit van het onderwijs: de onderwijsprocessen, zoals het pedagogisch klimaat en de leerlingenzorg, en de leerresultaten. De stimulerende taak van de inspectie werd specifiekier gedefinieerd. Deze zou zich in het bijzonder moeten richten op de bevordering van de kwaliteitszorg, waaronder de zelfevaluatie. Voorts diende het toezicht proportioneel van karakter te zijn, dat wil zeggen

extensief op scholen waar de kwaliteit tenminste voldoende is en intensief waar de onderwijskwaliteit lager is.

Omdat bij de start van de WOT nog geen sprake was van een systematisch opgezet programma voor de effectmeting van het onderwijstoezicht, is het paper gebaseerd op de analyse van empirische gegevens zoals die binnen en buiten de inspectie over de bedoelde periode beschikbaar waren. Het feit dat thans alleen gebruik gemaakt kon worden van deze empirische gegevens, bepaalt in de hoge mate de reikwijdte van de effectmeting en de daarop gebaseerde conclusies. Dat betekent dat lang niet alle eventuele effecten van het onderwijstoezicht kunnen worden geanalyseerd omdat er geen (voldoende) gegevens beschikbaar zijn.

De beantwoording van de vraag of onderwijstoezicht effectief is als volgt aangepakt. Om de effecten van toezicht te kunnen verklaren is allereerst een theoretisch model ontwikkeld. Dit model wordt in paragraaf 2 gepresenteerd en is gericht op de relatie tussen onderwijstoezicht en de kwaliteitsverbetering van scholen. Effecten van het toezicht op het onderwijsbeleid zijn buiten beschouwing gelaten. Op basis van dit model is in paragraaf 3 een aantal onderzoeksvragen geformuleerd en is gezocht naar empirische gegevens om deze te kunnen beantwoorden. Paragraaf 4 bevat de resultaten en gaat achtereenvolgens in op de vraag in welke mate PO-scholen effecten van het toezicht ervaren en of dit resulteert in kwaliteitsverbetering. Tot slot wordt in paragraaf 5 de vraagstelling beantwoord en van commentaar voorzien.

## **2. EEN THEORETISCH MODEL VAN HET EFFECT VAN ONDERWIJSTOEZICHT OP SCHOLEN**

Om te kunnen vaststellen in welke mate sprake is van effecten van toezicht is een theoretisch kader nodig op basis waarvan deze effecten aan het toezicht kunnen worden toegeschreven, inclusief de mechanismen waardoor en de context waarbinnen deze tot stand komen. Aanknopingspunten voor zo'n theoretisch kader zijn te vinden in Ehren, Scheerens & Leeuw (2005; zie ook Ehren, 2006) die de beleidstheorie hebben gereconstrueerd die ten grondslag ligt aan de WOT.

Op basis van deze reconstructie komt Ehren (2006) in haar dissertatie tot de conclusie dat de kwaliteitsverbetering van het onderwijs het beoogde effect van onderwijstoezicht is. De onderzoeken die de onderwijsinspectie op scholen verricht, richten zich immers op de vaststelling van de feitelijke kwaliteit van de school, inclusief de handhaving van wettelijke voorschriften, waarbij in ieder geval wordt beoordeeld of de instelling voldoet aan een waarborgniveau. Onder 'kwaliteit' dient volgens de WOT te worden verstaan, de leerprestaties maar ook de volgende proceskenmerken: aanbod, de leertijd, het schoolklimaat, het pedagogisch klimaat, het didactisch handelen van leraren, de leerlingzorg en de kwaliteit van toetsing en afsluiting.

Indien de kwaliteit van het onderwijs onder dat waarborgniveau zit, is er sprake van een 'zeer zwakke school'. Deze kwaliteitsbepaling wordt aan de school gerapporteerd maar ook via het internet in de vorm van inspectierapporten en kwaliteitskaarten openbaar gemaakt. De achterliggende gedachtegang is dat scholen worden geacht zich op basis van de beoordeling van de inspectie in te spannen om de kwaliteit te behouden respectievelijk te verbeteren. De

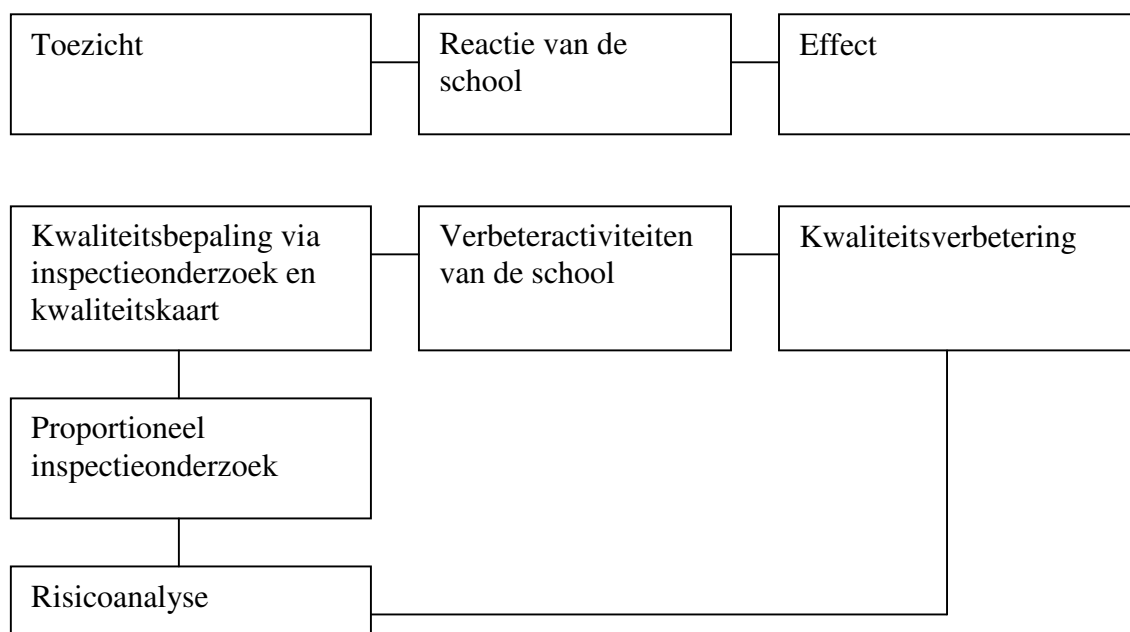
veronderstelling is dat de scholen de beoogde kwaliteitsverbetering kunnen realiseren op basis van de feedback van de inspectie naar aanleiding van het inspectieonderzoek, de rapportage daarover en/of de daarop gebaseerde kwaliteitskaart (Ehren, 2006; Janssens, 2005; De Wolf & Janssens, 2005). Scholen worden geacht op basis van deze feedback gerichte acties te ondernemen om sterke punten te behouden en zwakke punten te verbeteren (Visscher, 2002; Visscher & Coe, 2003). In de meeste gevallen kunnen scholen daarin eigen keuzes maken en accenten leggen. Dat geldt in veel mindere mate voor scholen die als zeer zwak worden beoordeeld en in een intensief toezichttraject belanden. In dat geval is het oordeel van de inspectie leidend voor de in gang te zetten verbeteractiviteiten (zie Ehren, 2006). Het onderwijstoezicht beoogt ook effect te hebben op het onderwijsbeleid op landelijk niveau, met name als gevolg van rapportages over de staat van het onderwijs. Dit effect is echter op deze plaats buiten beschouwing gebleven.

Om de beoogde effecten te bereiken past de onderwijsinspectie drie middelen toe.

- Het *inspectie-onderzoek*, bestaande uit verschillende (al dan niet onaangekondigde) onderzoeksvormen: Jaarlijks onderzoek (JO), Periodiek Kwaliteitsonderzoek (PKO), Nader Onderzoek (NO), Onderzoek naar de Kwaliteitsverbetering (OKV) en Incidenteel Onderzoek. Deze onderzoeksvormen worden proportioneel ingezet, dat wil zeggen dat enerzijds de inhoud van het toezicht wordt afgestemd op de (verantwoordings)gegevens van eventueel zelfonderzoek van de instellingen en anderzijds dat het toezicht intensief is waar de kwaliteit van de instellingen dat noodzakelijk maakt.
- *Risicomanagement* dat erop is gericht via risico-analyses te bewaken dat op basis van ontwikkelingen in de kwaliteit van instellingen de juiste onderzoeksvormen op het juiste moment worden ingezet.
- *Openbaar maken* van haar bevindingen over de kwaliteit van instellingen via publicatie van instellingsrapporten en bijbehorende kwaliteitskaarten.

Uit het de combinatie van de beoogde effecten en de bijbehorende middelen valt een procesmodel van de effecten van onderwijstoezicht af te leiden, dat in figuur 1 is weergegeven. Dit model bestaat uit drie onderdelen (toezicht, reactie instelling en effect), die elk door een serie factoren worden beïnvloed (Ehren & Visscher, 2006).

*Figuur 1: Procesmodel effecten van toezicht*



### ***Reikwijdte***

Onderzoek naar effecten en neveneffecten van toezicht is een complex onderzoeksgebied dat nog tamelijk onontgonnen is (zie De Wolf & Janssens, 2005; Ehren, 2006 en Leeuw & Willemsen, 2006). Dit wordt enerzijds veroorzaakt door het gebrek aan onderzoek op basis van robuuste en wetenschappelijk gevalideerde modellen van gedragsbeïnvloeding als gevolg van toezicht, zoals bijvoorbeeld te vinden in figuur 1. Anderzijds speelt ook de beperkte beschikbaarheid van empirische gegevens om dergelijke modellen te toetsen, een rol (Van Dijk, 2003). Dat geldt ook voor dit paper. De empirische gegevens die gebruikt zijn om de effecten van toezicht te analyseren zijn niet verzameld met het oog op toetsing van het eerder beschreven theoretische model, maar betreffen vooral de uitkomsten van enquêtes onder scholen naar de tevredenheid met het toezicht en de oordelen van inspecteurs over de kwaliteit van scholen over meerdere jaren. Daarnaast is gebruik gemaakt van twee wetenschappelijke publicaties over onderzoek naar het onderwijstoezicht, te weten een quasi-experimentele studie van het CPB (2007) naar het effect van het toezicht op de eindopbrengsten van PO-scholen en de dissertatie van Ehren (2006), een onderzoek naar de relatie tussen toezicht en schoolverbetering op PO-scholen op basis van tien case studies.

Bij de vraag of het toezicht effect heeft op de kwaliteitsverbetering van scholen spelen verschillende factoren een rol (zie Ehren & Visscher, 2006). Dit betreft in de eerste plaats eigenschappen van het inspectietoezicht zelf, zoals de rol van de inspecteur, de inrichting van het onderzoek en de kwaliteit van de feedback. In de tweede plaats betreft dit eigenschappen van de onderzochte scholen, zoals de verandercapaciteit, de interne kwaliteitszorg en de rol van bestuur en de schoolleiding bij verandertrajecten. Daarnaast hebben ook contextfactoren invloed op eventuele verbeterprocessen van de school, zoals de ondersteuning vanuit de omgeving van de school en de onderwijsbegeleiding. Cumulatie van deze factoren veroorzaakt dat een school zich op basis van de uitkomsten van inspectieonderzoek wel of niet zal verbeteren. Over de eigenschappen van de scholen en de rol van de context in relatie tot het toezicht, zijn geen voldoende gegevens beschikbaar. Ehren (2006) vormt hierop een uitzondering, maar in deze studie is geen verband gelegd tussen toezicht en feitelijke kwaliteitsverbetering, maar wel tussen toezicht en acties van scholen om tot verbetering te komen.

Een ander probleem bij het identificeren van effecten van toezicht is dat de inspectieonderzoeken niet random worden uitgevoerd, maar primair op basis van de door de inspectie eerder beoordeelde kwaliteit, eventuele risico's op kwaliteitsverlies en op basis van signalen zoals klachten van ouders en deelnemers. Het eventuele effect van het toezicht is daarom niet eenduidig toe te schrijven aan de resultaten van het inspectieonderzoek, maar mogelijk ook aan andere onbekende factoren (zie Luginbuhl, Webbink & De Wolf, 2007). Omdat alle scholen jaarlijks worden onderzocht is een gerandomiseerd experiment niet mogelijk. De gegevens waarvan in deze rapportage gebruik is gemaakt zijn ook niet verzameld of geanalyseerd binnen een quasi-experimenteel design. Een uitzondering hierop vormt de studie van Luginbuhl, Webbink en De Wolf (2007), maar hierin blijven de acties van de scholen om de kwaliteit van het onderwijs te verbeteren impliciet.

Het bovenstaande leidt ertoe dat er gekozen is voor een exploratief onderzoek op basis van het eerder beschreven theoretisch model, waarmee geen causale verbanden gelegd worden tussen toezicht en schoolverbetering. Er is derhalve eerder sprake van plausibiliteit dan van causaliteit.

### 3. VRAAGSTELLING EN METHODE

De vraagstelling die ten grondslag ligt aan het paper is de volgende: *Leiden de activiteiten van de Inspectie van het Onderwijs tot de verbetering van de kwaliteit van PO-scholen, gemeten aan zowel proceskenmerken als de leerprestaties?*

Onder activiteiten van de inspectie dienen in dit verband te worden verstaan: het onderzoeken van de kwaliteit van scholen, de mondelinge terugkoppeling tijdens een schoolbezoek en de rapportages via inspectierapporten en kwaliteitskaarten.

Bovenstaande onderzoeksvraag wordt aan de hand van de volgende deelvragen beantwoord.

1. Is er sprake van kwaliteitsverbetering op PO-scholen die tenminste twee keer door de inspectie zijn onderzocht?
2. Leidt het toezicht tot acties op scholen om de kwaliteit te behouden of te verbeteren?
3. Leidt het toezicht tot kwaliteitsverbetering van scholen?

De eerste deelvraag wordt beantwoord op basis van inspectiegegevens over de kwaliteitsontwikkeling van scholen die door de inspectie de afgelopen jaren zijn onderzocht. Het gaat hier om de oordelen van inspecteurs over de bezochte scholen voor PO in de periode 2002-2006 die zijn opgeslagen in de databestanden van de onderwijsinspectie

Om de tweede deelvraag te beantwoorden zijn gegevens uit twee bronnen gebruikt. In de eerste plaats zijn dat de oordelen van besturen, schoolleiders en leraren over het beoogde effect van het onderzoek op scholen en de feedback die men daarover heeft ontvangen. De empirische gegevens die hiervoor zijn gebruikt, zijn in de eerste plaats gegevens over de waardering en het nut van schoolbezoeken die plaatsvonden in de periode 2002-2006. Deels is hierbij gebruik gemaakt van onderzoeken die door derden zijn gerapporteerd en deels van onderzoeken die de onderwijsinspectie zelf in het kader van haar eigen kwaliteitszorg uitvoert naar de tevredenheid van scholen over de uitgevoerde schoolbezoeken. In de tweede plaats is gebruik gemaakt van de uitkomsten van de dissertatie van Ehren (2006) die onder andere heeft onderzocht tot welke verbeteracties het toezicht op PO-scholen leidt. In de derde plaats is gebruik gemaakt van oordelen van ouders, deelnemers en scholen over het beoogde effect van openbare inspectierapporten en kwaliteitskaarten op schoolverbetering en schoolkeuzes van ouders en deelnemers. Hierbij is in belangrijke mate gebruik gemaakt van de bevindingen van onderzoek dat door derden is uitgevoerd.

De derde deelvraag wordt beantwoord op basis van enerzijds onderzoek uitgevoerd door het CPB (Luginbuhl, Webbink & De Wolf, 2007) waarin de schoolprestaties van PO-scholen worden vergeleken die in 1999-2003 wel en niet door de inspectie zijn onderzocht. Anderzijds is gebruik gemaakt van inspectiegegevens.

### 4. RESULTATEN

In paragraaf 4.1 wordt eerst een overzicht gegeven van de ontwikkeling die het toezicht de laatste tien jaar heeft doorgemaakt wat betreft de onderzoeksvormen en de aantallen scholen die onderzocht zijn. In de daaropvolgende paragrafen wordt nader ingegaan op het effect van toezicht op de onderwijsverbetering.

Paragraaf 4.2 behandelt de vraag of de kwaliteit van PO-scholen is verbeterd die tenminste twee keer door de inspectie zijn onderzocht. Paragraaf 4.3 gaat in op de mate waarin volgens

het scholenveld inspectieonderzoek nuttig en bruikbaar is voor schoolverbetering. Paragraaf 4.4 beantwoordt de vraag of kwaliteitskaarten effect hebben op schoolverbetering. En in paragraaf 4.5 komt de relatie aan de orde tussen toezicht en kwaliteitsverbetering van scholen.

#### **4.1. ONTWIKKELINGEN IN HET TOEZICHT**

De kern van het inspectiewerk bestaat uit het onderzoeken van scholen, zowel in de periode vóór als na de invoering van de WOT. Voor de inwerkingtreding van de WOT paste de inspectie in het IPO twee onderzoeksvormen toe: het Integraal Schooltoezicht (IST) en het Reguliere Schooltoezicht (RST). Het RST was erop gericht op hoofdlijnen en aan de hand van beperkt aantal kernindicatoren de kwaliteit van een school vast te stellen. Gedurende één schooldag werden lessen geobserveerd en vonden er interviews plaats met de schoolleiding en leraren die belast waren met zorg en begeleiding van leerlingen. Indien uit dit onderzoek bleek dat er mogelijk problemen waren met de kwaliteit van het onderwijs werd dit RST op korte termijn gevolgd door een IST, waarin ook andere indicatoren aan de orde werden gesteld. Een IST was dus enerzijds breder van karakter en indicatoren werden meer diepgaand onderzocht. Het onderzoek nam dan ook meerder dagen in beslag.

Het IST werd in die periode ook gebruikt om bij een jaarlijkse steekproef van scholen de kwaliteit in volle omvang te beoordelen met het oog op het beschrijven van de kwaliteit van het basisonderwijs ten behoeve van het jaarlijkse onderwijsverslag.

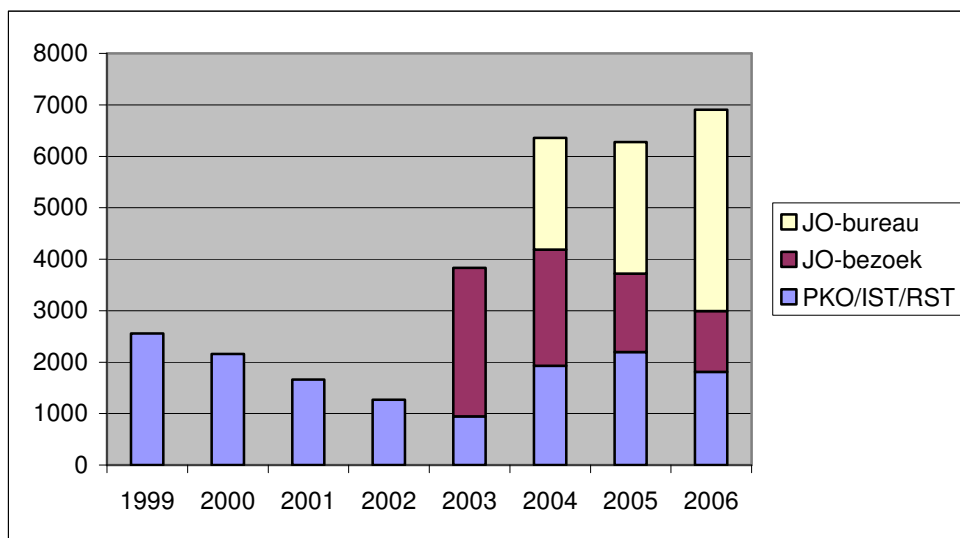
Na de invoering van de WOT wordt het PKO en het Jaarlijkse Onderzoek (JO) ingevoerd. Het PKO is nog het meest te vergelijken met het RST omdat deze onderzoeksvorm zich ook richt op een beperkt aantal kernindicatoren. Het JO is een nieuwe onderzoeksvorm waarin standaard alleen de opbrengsten worden beoordeeld en de kwaliteitszorg. Deze onderzoeksvorm is uitdrukkelijk op verzoek van de Tweede Kamer als onderzoeksvorm aan de WOT toegevoegd, omdat de toenmalige volksvertegenwoordiging erop stond dat de onderwijsinspectie jaarlijks een toezichtcontact met alle scholen had (TK 27783, nr. 22, p.1). Uitgebreide onderzoeken die het karakter en de omvang hadden van het toenmalige IST zijn nog steeds mogelijk in de vorm van een Nader Onderzoek, maar worden in de praktijk zeer weinig toegepast.

Figuur 2 geeft het aantal onderzoeken in het basisonderwijs weer dat vanaf 1999 is uitgevoerd, uitgesplitst naar kwaliteitsonderzoeken (PKO, IST, RST) en jaarlijkse onderzoeken (JO). Een jaarlijks onderzoek start als een bureauonderzoek (JO-bureau) op basis van bij de inspectie beschikbare gegevens en gegevens die PO-scholen aanleveren via een jaarlijkse vragenlijst. De kern van deze gegevens beperkt zich tot de opbrengsten en de zorgleerlingen. Als uit de analyse van deze gegevens blijkt dat er voor enig jaar geen risico's zijn voor de kwaliteitsontwikkeling van een school, dan beperkt het jaarlijkse onderzoek zich tot dit een JO-bureau.

Uit figuur 2 blijkt in de eerste plaats dat het totaal aantal onderzoeken sinds de invoering van de WOT in 2002 sterk gestegen is, met name in de jaren 2003 en 2004. Dat is niet verwonderlijk omdat sinds de invoering van de WOT scholen jaarlijks worden onderzocht en tenminste ééns in de vier jaar een kwaliteitsonderzoek krijgen en eens per vier jaar een Jo-bezoek. Jaarlijks wordt nagenoeg de hele populatie onderzocht. In de periode voorafgaande aan de WOT lag de onderzoeksfrequentie veel lager; gemiddeld genomen werden jaarlijks ca 2000 scholen onderzocht met een RST of IST.

Kijken we alleen naar de onderzoeksvormen waarin de onderwijskwaliteit zo breed mogelijk wordt beoordeeld (PKO, IST en RST) dan valt op dat de aantallen over de jaren heen geen grote schommelingen laten zien: gemiddeld genomen werd zowel voor als na de invoering van de WOT ongeveer een kwart van de PO-scholen met een kwaliteitsonderzoek bezocht.

Een tweede conclusie is dat vanaf 2004 de verhouding tussen JO-bezoek en JO-bureau binnen de jaarlijkse onderzoeken aan het verschuiven is. In 2003 werd bij alle jaarlijkse onderzoeken de school ook daadwerkelijk bezocht, in 2004 was dit nog maar bij ruim de helft het geval en in 2006 nog slechts bij een kwart. In de overige driekwart van de gevallen beperkt het jaarlijkse onderzoek zich tot een JO-bureau. De invoering van risicogericht toezicht is één van de redenen waarom het percentage JO-bezoek is afgenomen.

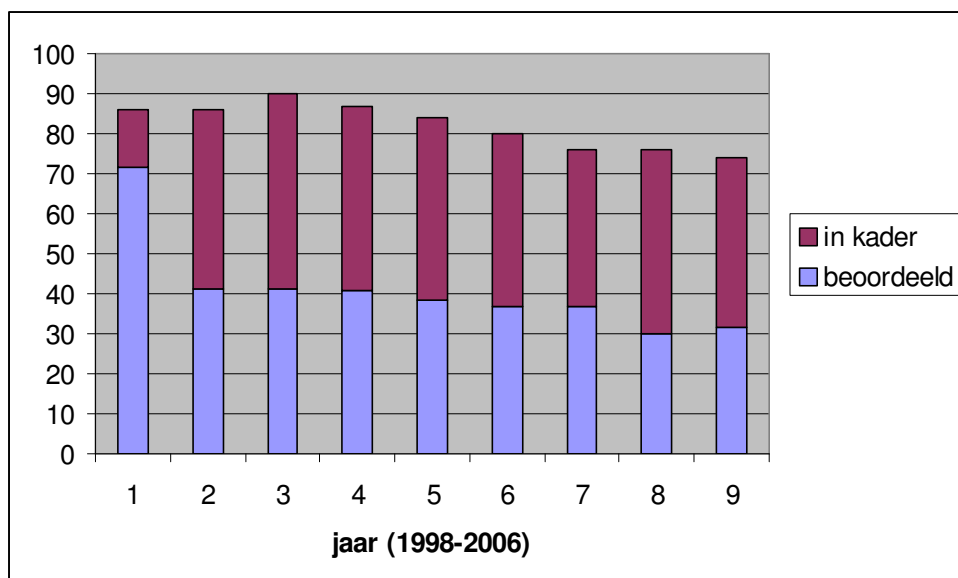


*Figuur2 Aantallen reguliere kwaliteitsonderzoeken in het basisonderwijs, 1999-2006*

Samenvattend kan gesteld worden dat sinds de invoering van de WOT jaarlijks alle PO-scholen worden onderzocht en het aantal feitelijke schoolbezoeken is verdubbeld ten opzichte van de periode daarvoor. Wel worden sinds de invoering van de WOT door de inspectie steeds minder indicatoren tijdens een onderzoek beoordeeld (zie figuur 3).

Het gemiddelde aantal indicatoren dat in de loop der jaren in de inspectieonderzoeken (RST, IST en PKO) feitelijk is beoordeeld, is te vinden in figuur 3. Wat opvalt is dat zowel voor als na de invoering van de WOT sprake is van een proportionele toepassing van de waarderingskaders. Gemiddeld genomen blijkt dat vanaf 1999 steeds iets minder dan de helft van het totaal aantal indicatoren uit een waarderingskader tijdens een inspectieonderzoek beoordeeld wordt. Daarnaast laat figuur 3 zien dat ook het aantal indicatoren in het waarderingskader over de jaren heen beduidend minder is geworden.





*Figuur 3 Gemiddeld aantal beoordeelde indicatoren per kwaliteitsonderzoek in het PO (1998-2006)*

In 2006 is bij slechts 7 van de 1114 PKO's het volledige kader afgenomen. Bij bijna tweederde deel van de scholen werd tussen de 30 en 40 indicatoren beoordeeld en bij het andere deel veelal tussen de 20 en 30 indicatoren.

## 4.2. KWALITEITSVERBETERING VAN SCHOLEN

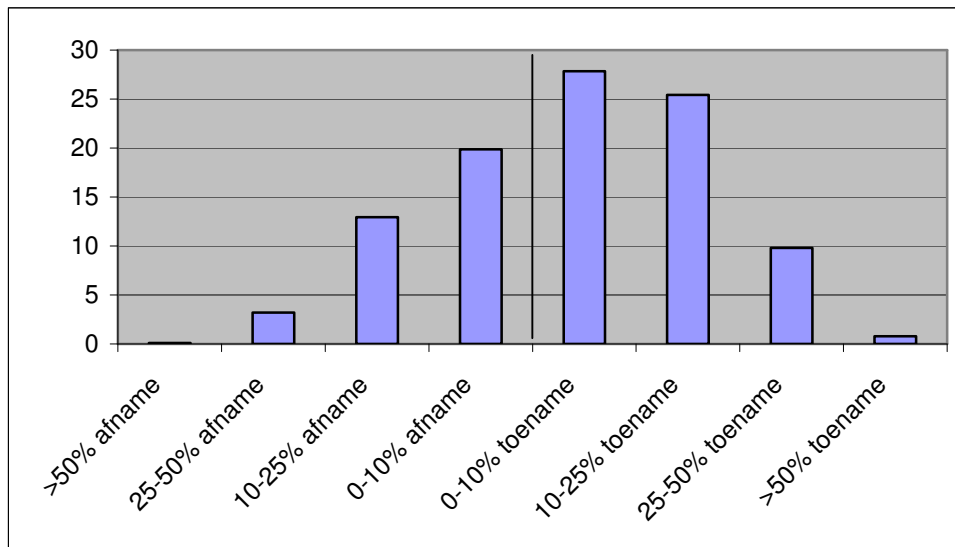
In deze paragraaf wordt een meerjarenoverzicht gegeven van de kwaliteitsverandering in drie groepen basisscholen, door de uitkomsten van het eerste inspectieonderzoek te vergelijken met die van een vergelijkbaar tweede inspectieonderzoek. Deze groepen betreffen in de eerste plaats de gehele populatie basisscholen die in de periode 1999-2006 tenminste twee maal zijn onderzocht. Gezien de looptijd is het tweede onderzoek in veruit de meeste gevallen niet door dezelfde inspecteurs verricht.

In de tweede plaats komt de kwaliteitsverandering aan de orde op scholen die eerder een inspectieonderzoek hebben gehad dan volgens de vierjaars-cyclus geïndiceerd was omdat hun kwaliteit een risico vormt. Ten slotte komt de groep scholen aan de orde die als zeer zwak zijn gekwalificeerd en derhalve onder de meest intensieve vorm van toezicht staan of hebben gestaan. Nagegaan wordt in welke mate zich tussen het eerste en tweede bezoek veranderingen hebben voorgedaan in de aantallen indicatoren die als voldoende of onvoldoende zijn beoordeeld. Hiermee kunnen geen causale verbanden worden gelegd tussen toezichtinterventies en de kwaliteit van scholen, maar kan wel een indicatie worden verkregen van de eventuele bijdrage van het toezicht aan schoolverbetering.

### *Kwaliteitsontwikkelingen in de populatie*

Als scholen, zoals blijkt uit de vorige paragraaf, op basis van de uitkomsten van het toezicht verbeteractiviteiten in gang zetten, dan mag worden verwacht dat toezichtinterventies, een rol spelen in de kwaliteitsverbetering van scholen. Dit zou betekenen dat het percentage scholen dat aan de indicatoren uit het waarderingskader voldoet, over de jaren heen toe moet nemen. Omdat het overgrote deel van de basisscholen in de periode 1998-2006 tenminste tweemaal

met een kwaliteitsonderzoek door de inspectie is onderzocht, kan worden nagegaan of er verschillen zitten tussen het eerste en het tweede onderzoek<sup>2</sup> (zie figuur 4).



Figuur 4 Percentage voldoende's tussen eerste en laatste schoolbezoek in de periode 1998 - 2006 (N=5095<sup>3</sup>)

Indien per school het eerste en het laatste onderzoek met elkaar worden vergeleken, blijkt dat het aantal indicatoren dat als 'voldoende' is beoordeeld met 5% is toegenomen. Over de gehele lijn hebben de bezochte basisscholen zich dus verbeterd. Dit geldt overigens niet voor alle scholen, zoals te zien is in figuur 4. Links van de verticale lijn staan de basisscholen (36%) waar sprake is van een afname van het aantal voldoende's. Het tweede onderzoek leidde in deze gevallen tot negatievere oordelen dan het eerste. Rechts van de verticale lijn staan de basisscholen die zich hebben verbeterd (63%) of in kwaliteit gelijk zijn gebleven (1%).

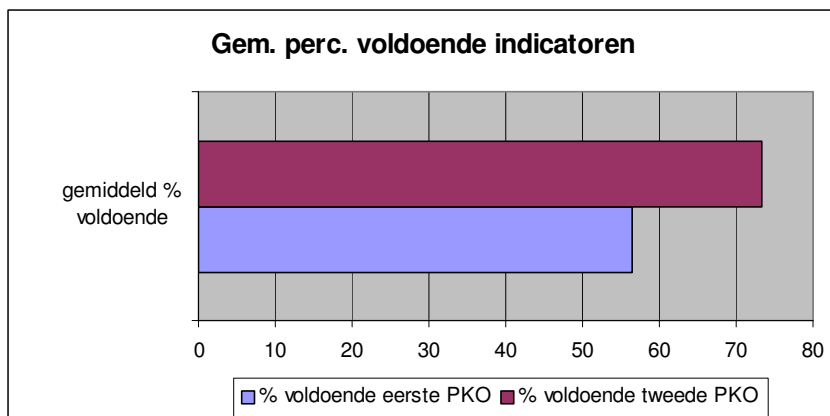
#### *Kwaliteitsontwikkelingen op risicoscholen*

In beginsel krijgen volgens de WOT PO-scholen eens in de vier jaar een PKO, tenminste als de inspectie zich geen zorgen maakt over de kwaliteit. Bij scholen waar dit wel het geval is wordt vaker een PKO uitgevoerd. Sinds 2003 zijn 209 PO-scholen eerder met een PKO bezocht dan de gebruikelijke vierjaarlijkse termijn omdat ze als risicoschool gedefinieerd werden. In bijna alle gevallen zijn deze scholen binnen twee jaar opnieuw met een PKO bezocht.

Bij het tweede PKO op risicoscholen blijkt dat het gemiddelde aantal indicatoren dat door de inspectie als voldoende is beoordeeld met 15% is gestegen (figuur 5).

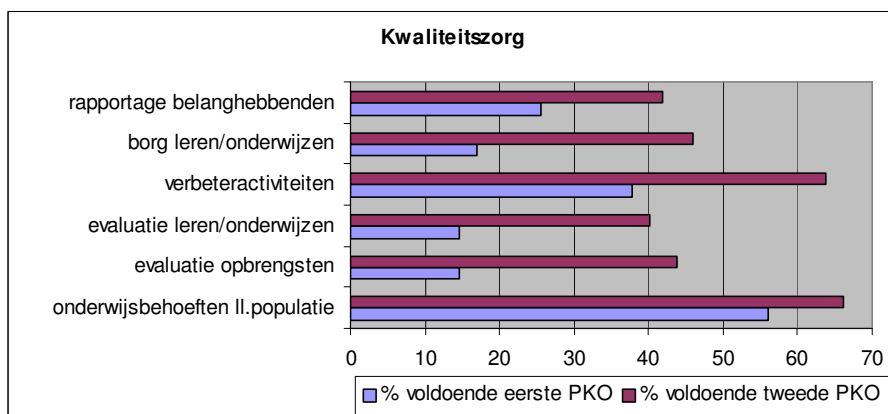
<sup>2</sup> Het is niet eenvoudig te bepalen op welke indicatoren scholen zich door de jaren heen verbeterd hebben. De reden hiervoor is dat het waarderingskader over de jaren heen is gewijzigd. Waar onderlinge vergelijking mogelijk was, is nagegaan of het percentage scholen dat aan de indicatoren voldoet door de jaren heen toeneemt.

<sup>3</sup> Nog niet alle basisscholen zijn voor de tweede maal bezocht.

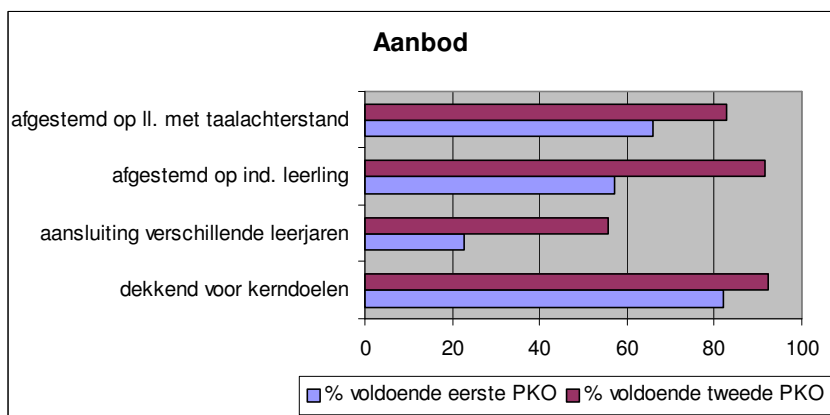


*Figuur 5 Gemiddelde percentage voldoende en onvoldoende 1<sup>ste</sup> en 2<sup>e</sup> PKO op risicoscholen, 2003 – 2006 (N=209)*

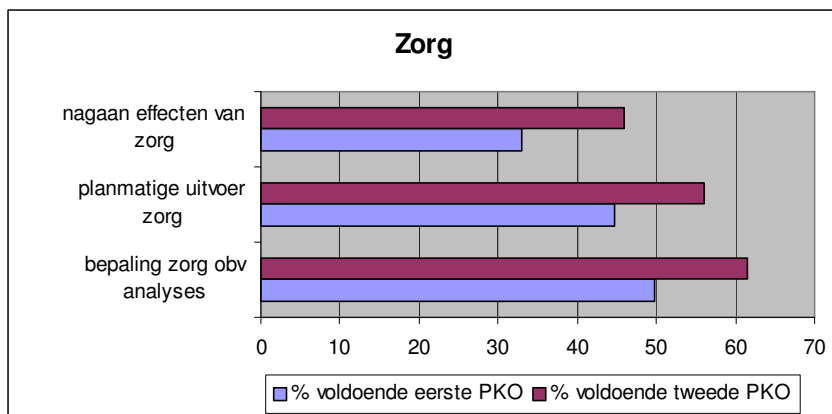
De sterkste verbetering valt te constateren bij indicatoren van de kwaliteitszorg (figuur 6), de het leerstofaanbod (figuur 7), zorg en begeleiding (figuur 8 en 9) en de leerprestaties (figuur 10). Op nogal wat indicatoren heeft 20 tot 40% van de scholen zich bij een tweede PKO verbeterd. Indicatoren die niet of nauwelijks verbeteren, zijn de leertijd (hier is zelfs sprake van een afname in het percentage voldoende), de sfeer op school en de actieve rol van de leerlingen tijdens de lessen.



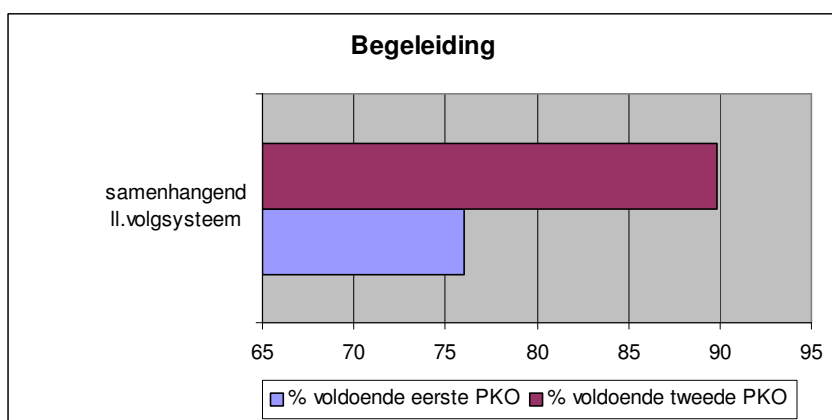
*Figuur 6 Verschillen in beoordeling tussen 1<sup>ste</sup> en 2<sup>e</sup> PKO 2003 – 2006 voor het onderdeel Kwaliteitszorg (N=209)*



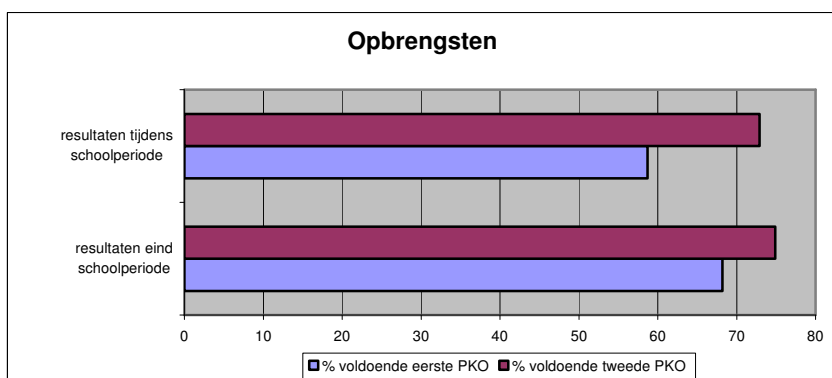
*Figuur 7 Verschillen in beoordeling tussen 1<sup>ste</sup> en 2<sup>e</sup> PKO 2003 – 2006 voor het onderdeel Leerstofaanbod (N=209)*



*Figuur 8 Verschillen in beoordeling tussen 1<sup>ste</sup> en 2<sup>e</sup> PKO 2003 – 2006 voor het onderdeel Zorg (N=209)*



*Figuur 9 Verschillen in beoordeling tussen 1<sup>ste</sup> en 2<sup>e</sup> PKO 2003 – 2006 voor het onderdeel Begeleiding (N=209)*



*Figuur 10 Verschillen in beoordeling tussen 1<sup>ste</sup> en 2<sup>e</sup> PKO 2003 – 2006 voor het onderdeel Opbrengsten (N=209)*

#### *Kwaliteitsontwikkelingen op zeer zwakke scholen*

De groep die het meest intensief door de inspectie is onderzocht zijn de zogenaamde ‘zeer zwakke scholen’. Dit zijn scholen waar de leeropbrengsten en het onderwijsleerproces beide als ‘onvoldoende’ zijn beoordeeld. Op deze scholen is het toezicht het meest intensief, omdat deze scholen een verbeterplan aan de inspectie moeten voorleggen dat intensief door de inspectie wordt gemonitord. Na twee of drie jaar onderzoekt de inspectie of deze scholen zich al dan hebben verbeterd.

In de periode voorafgaande aan de WOT (1998-2002) zijn 284 basisscholen als zeer zwak beoordeeld. Dit is 4% van de onderzochte scholen (N=7140). Van deze groep scholen is nagegaan welke ontwikkeling ze tot op heden hebben doorgemaakt (zie tabel 2).

Ontwikkeling van 284 zeer zwakke scholen (1998 - 2006)	N
- opgeheven of gefuseerd	10
- Verbeterd voor 2002	76
- Verbeterd na 2002	162
- Nog steeds zeer zwak	30
- Wederom zeer zwak	6

*Tabel 2 Ontwikkeling van zeer zwakke scholen in de periode 1998 – 2006*

De overgrote meerderheid (87%) heeft zich direct of na enige jaren verbeterd. Zo bracht 84% van de scholen hun opbrengsten op het juiste niveau. Het leerstofaanbod op de scholen die zich verbeterden, is bij het laatste schoolbezoek aanzienlijk vaker dekkend voor de kerndoelen, het sluit beter aan bij verschillen tussen leerlingen en het wordt vaker aan alle leerlingen volledig aangeboden. Het plannen en het efficiënt gebruik van de leertijd is een fractie verbeterd, maar dit was niet het grootste probleem van de zeer zwakke scholen. Het pedagogisch klimaat is bij het laatste schoolbezoek vaker stimulerend voor de leerlingen en het schoolklimaat is ook veiliger geworden. De instructie is nu duidelijker, beter georganiseerd en de leerlingen vertonen ook een grotere betrokkenheid. De prestaties van de leerlingen worden veel vaker met landelijk genormeerde toetsen gepeild. Dit gebeurt op deze scholen nu zelfs vaker dan landelijk. Dit is uiteraard een belangrijk startpunt om tot verbeteringen te komen. Op deze scholen worden ook vaker maatregelen getroffen voor zwakke leerlingen. Ondanks deze soms vrij sterke verbeteringen zijn deze scholen echter nog niet geheel op het niveau van de modale basisschool.

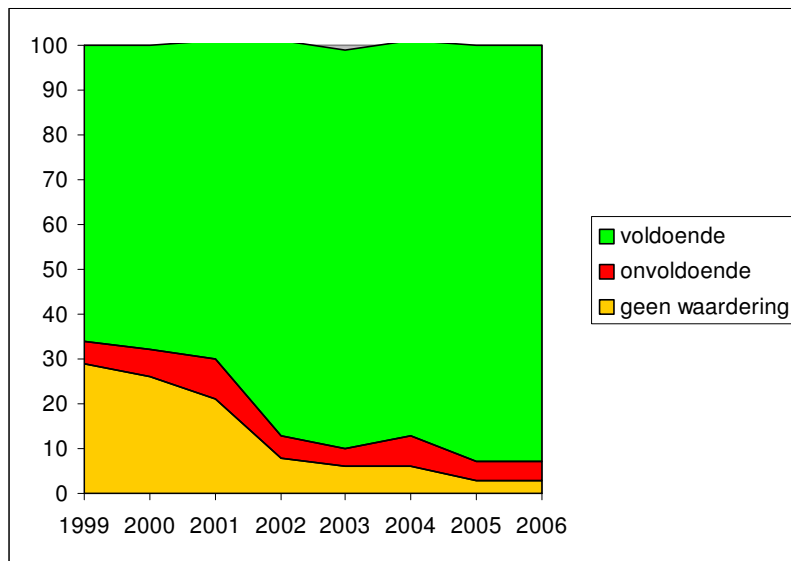
Wat de organisatie betreft valt op dat het functioneren van het bestuur en het onderwijskundig leiderschap vooruitgegaan zijn en op het gemiddelde niveau is gekomen. Ook de kwaliteitszorg is op de scholen die eerder zeer zwakke onderwijsresultaten hadden, duidelijk beter geworden. Deze scholen hebben concrete maatregelen getroffen om de kwaliteit van het onderwijs te verbeteren en te behouden.

Een klein deel van de scholen (13%) is nog zeer zwak of heeft nog onvoldoende of onbekende leerresultaten. We mogen dus concluderen dat de meeste zeer zwakke basisscholen zich verbeterd hebben. Overigens zijn er na 2002 ook weer nieuwe zeer zwakke scholen bijgekomen. Het gaat hier echter om een kleiner aantal dan in de jaren voor de WOT. In de periode daarna (2003-heden) zijn 147 scholen als zeer zwak beoordeeld. Dat is 2,1% van de in die periode onderzochte scholen (N=6964). Van deze 147 scholen zijn er vijf die na twee jaar opnieuw als zeer zwak zijn beoordeeld. Het totale percentage zeer zwakke basisscholen neemt dus af.

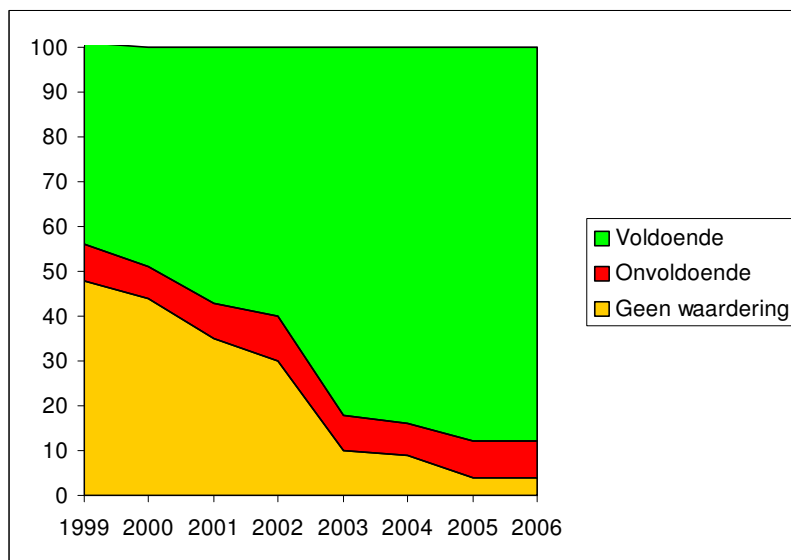
#### *Ontwikkelingen in opbrengstmetingen*

Sinds 1999 beoordeelt de inspectie de leerprestaties van basisscholen. Het gaat hierbij om de eindopbrengsten, meestal beoordeeld op grond van de prestaties op de Eindtoets van het Cito, en om de zogenaamde tussenopbrengsten, waarbij toetsen uit leerlingvolgsystemen worden gebruikt om de taal- en rekenen/wiskunde prestaties in de leerjaren drie tot en zeven te beoordelen. Scholen krijgen het oordeel ‘voldoende’ als de leerprestaties op of boven het niveau liggen dat op grond van de leerlingpopulatie mag worden verwacht en het oordeel ‘onvoldoende’ als de leerprestaties hier gedurende meerdere jaren onder liggen. In het geval

de scholen geen of onvoldoende genormeerde toetsen gebruiken worden de opbrengsten door de inspectie niet beoordeeld.



*Figuur 11 Beoordeling eindopbrengsten op basisscholen in de periode 1999-2006*



*Figuur 12 Beoordeling tussenopbrengsten op basisscholen in de periode 1999-2006*

In figuren 11 en 12 staat weergegeven hoe de beoordeling van de eindopbrengsten en de tussenopbrengsten zich sinds 1999 hebben ontwikkeld. Hieruit kunnen twee conclusies worden getrokken. De eerste is dat in de loop der jaren het percentage scholen waarvan de eind- en tussenopbrengsten niet beoordeeld kunnen worden, sterk is gedaald. Meer scholen kunnen zich in de loop der jaren over hun opbrengsten verantwoorden, omdat ze in de loop der jaren steeds vaker toetsen zijn gaan gebruiken op basis waarvan de inspectie de opbrengsten kan beoordelen. In 2006 bleek dat nog steeds niet mogelijk op ongeveer 5% van de scholen.

Een tweede conclusie is dat het relatieve percentage basisscholen met opbrengsten die als 'voldoende' zijn beoordeeld, is gestegen. Het percentage scholen met onvoldoende opbrengsten is gedurende deze periode ongeveer gelijk gebleven.

### *Conclusie*

Afgaande op de oordelen van de inspectie kan worden geconcludeerd dat tweederde deel van de basisscholen die de tussen 1998 en 2006 tenminste twee maal door de inspectie zijn bezocht, zich, gemiddeld genomen, hebben verbeterd. De kwaliteitsverbetering op risicoscholen is nog groter. Ook voor de scholen die onder het meest intensieve toezicht staan, geldt dat de overgrote meerderheid zich in de loop der jaren weet te verbeteren. Van de zeer zwakke scholen is bijna 90% na enige jaren niet meer als 'zeer zwak' te beschouwen en ook het aantal risicoscholen dat aan bijna alle indicatoren voldoet, is fors toegenomen. Vooral het didactisch handelen, de kwaliteitszorg, het leerstofaanbod en de opbrengsten zijn verbeterd. Voor wat betreft de opbrengsten, valt te constateren dat scholen zich hierover steeds vaker kunnen verantwoorden, omdat meer scholen landelijk genormeerde toetsen zijn gaan gebruiken. Bovendien is het relatieve percentage scholen met goede opbrengsten over de jaren heen toegenomen.

Hiermee is niet gezegd dat er een oorzakelijk verband is tussen toezicht en schoolverbetering, maar wel dat het basisonderwijs zich in de afgelopen heeft verbeterd. Wel mag worden aangenomen dat de toename van het aantal scholen dat zich in de afgelopen zeven jaren over hun opbrengsten kunnen verantwoorden een gevolg is van het feit dat de inspectie sinds 1999 ook de leerprestaties is gaan beoordelen.

De vraag is nu of en in welke mate deze positieve ontwikkelingen in de kwaliteit van de PO-scholen aan het onderwijstoezicht zijn toe te schrijven.

### **4.3. REACTIES VAN SCHOLEN OP INSPECTIEONDERZOEK**

Het eerste deel van de eerste onderzoeksvraag heeft betrekking op de relatie tussen inspectieonderzoeken vanaf de inwerkingtreding van de WOT en de verbeteracties die scholen ondernemen op basis van de uitkomsten daarvan.



Over de reacties van scholen op het toezicht zijn twee soorten bronnen beschikbaar. In de eerste plaats onderzoek naar de waardering van het onderwijstoezicht door het onderwijsveld. Het gaat hier om zowel extern onderzoek (Emmelot, et al., 2004; Hofman et al., 2005 en Q\*primair, 2006) als eigen onderzoek van de inspectie dat sinds 2002 is uitgevoerd naar de ervaringen van scholen met inspectieonderzoek. Deze onderzoeken hebben in alle gevallen betrekking op zelfrapportages van scholen en andere betrokkenen.

In de tweede plaats is gebruik gemaakt van de dissertatie van Ehren (2006) die rapporteert over 10 case studies van PO-scholen die in 2004 en 2005 door de inspectie zijn bezocht.

#### *Zelfrapportages van scholen*

Vrijwel alle studies naar de tevredenheid van scholen met het toezicht tonen aan dat ruim 80% van de direct betrokkenen (docenten, schoolleiders) tevreden is over het onderwijstoezicht en van mening is dat inspectiebezoeken nuttig voor de school zijn. Op 50 tot 70% van de scholen blijkt, volgens de ondervraagden, onderwijstoezicht te leiden tot positieve gedragsbeïnvloeding van leraren en schoolleiders en tot veranderingen in het schoolbeleid (zie AVS, 2004; De Wolf & Janssens, 2005; Janssens, 2005, Emmelot, et al., 2004 en Hofman et al., 2005). Dat het onderwijsveld tamelijk positief is over het onderwijstoezicht geldt

overigens ook de periode voorafgaand aan de WOT (zie bijv. Braster, 1987 en Van Esch & Laemers, 1990).

Over de inspectiebezoeken die sinds de invoering van de WOT hebben plaatsgevonden zijn de schooldirecties in het basisonderwijs behoorlijk tevreden. Het speciaal onderwijs is minder tevreden, omdat directies vinden dat de Inspectie onvoldoende begrip heeft voor de specifieke situatie in hun sector. Het oordeel van directeuren in het basisonderwijs over het waarderingskader is gemiddeld genomen gematigd positief. De meeste schoolleiders vinden niet dat het waarderingskader hen afremt in hun schoolontwikkeling en het maken van eigen keuzes. Van de schoolbesturen vindt echter 35% dat het toezichtkader te weinig rekening houdt met de omstandigheden van de school. Traditionele vernieuwingsscholen en scholen van kleine richtingen zijn het meest ontevreden over het waarderingskader. Het waarderingskader is volgens critici uit deze hoek te veel afgestemd op de gemiddelde school en sluit niet aan op de grote variëteit die het onderwijsveld in Nederland kent. Er zijn echter ook scholen die sterk lijken te steunen op het inspectiekader en het bezoek van de inspecteur. Soms kan het inspectiebezoek gebruikt worden om door de schoolleiding gewenste veranderingen door te voeren (zie Onderwijsraad, 2006).

Q-Primair (Hofman & De Boom, 2006) heeft met een monitor in kaart gebracht hoe het staat met het kwaliteitszorgbeleid van basisscholen (N=233) en besturen (N=156), mede in relatie tot het onderwijstoezicht. Bijna alle scholen (94%) vinden dat het inspectietoezicht de school stimuleert tot schoolverbetering en 87% vindt dat het de kwaliteit van de school bevordert. In bijna een derde deel van de scholen zijn, volgens de ondervraagden, de meeste verbeteringen geïndiceerd door het inspectietoezicht. Een vergelijkbare groep vindt wel dat in het inspectietoezicht te sterk het perspectief van schoolontwikkeling domineert. Ruim 80% van de bovenschoolse besturen rapporteert het toezichtkader te gebruiken als basis voor schoolzelfevaluatie en vindt dat het voldoende ruimte laat voor een eigen bovenschoolse beleid. Verder meldt een groot deel van de besturen (81%) de uitkomsten van inspectieonderzoeken te gebruiken voor vaststelling van haar eigen kwaliteitsbeleid, al hebben de meeste besturen ook eigen opvattingen over kwaliteit (74%).

In het evaluatie-onderzoek van de WOT dat door het SCO-Kohnstamminstituut is uitgevoerd over de sectoren PO, VO en BVE (Emmelot, et al., 2004) is specifiek gevraagd naar de stimuleringsfunctie van de inspectie. Het feit dat de Inspectie expliciet vraagt naar zelfevaluatiegegevens heeft volgens de scholen invloed op de inspanningen die scholen op dit terrein leveren. Eigen beleidsbeslissingen en afspraken met besturen en /of bovenschools management worden echter ook belangrijk gevonden. De scholen vinden het inspectierapport nuttig als impuls voor kwaliteitsverbetering, maar niet in sterke mate. Het nut van een inspectierapport wordt vooral gezocht in de mogelijkheden die het biedt voor reflectie op de eigen kwaliteitszorg en voor bijstelling van het schoolplan.

Kritische geluiden zijn er ook (zie Emmelot et al., 2004). Zo blijkt uit de panelgesprekken met directeuren van basisscholen dat het waarderingskader van de inspectie een (te) hoge invloed kan hebben op de wijze waarop scholen hun zelfevaluaties richten. Voorts was het onderwijsveld van mening dat de inspectie in het toezicht (te) veel nadruk legt op de opbrengsten van scholen.

Ook de inspectie zelf doet onderzoek naar de mate waarin PO-scholen die bezocht zijn voor een inspectie-onderzoek dit als effectief hebben ervaren. Scholen wordt na afloop van het onderzoek de mogelijkheid geboden via een anonieme vragenlijst te rapporteren in welke



mate ze tevreden zijn met de voorbereiding en uitvoering van het onderzoek en in welke mate ze dit als nuttig hebben ervaren.

De mate waarin scholen PKO's over de afgelopen vier jaren als effectief hebben ervaren staat weergegeven in tabel 1. Hierbij is gebruik gemaakt van items uit de vragenlijsten van de inspectie over schoolbezoeken die betrekking hebben op het door de scholen gepercipieerde effect en geformuleerd zijn in de vorm van een stelling. Per stelling is vermeld hoeveel procent van de basisscholen het er 'enigszins' of 'zeer mee eens' is. Hoe hoger het percentage, des te groter het door de scholen verwachte effect van het PKO.

Tabel 3 Percentage basisscholen dat aangeeft het met de stelling eens te zijn (2003-2006) (N=2368)

Evaluatie PKO PO	2003 n=680	2004 n=606	2005 n=583	2006 n=499
De manier waarop de inspectie het onderzoek heeft verricht, was voor ons stimulerend (of: versterkt het leerproces van onze school).	86	89	90	93
De rapportage bevat voldoende gegevens waar wij in de toekomst wat mee kunnen.	98	99	99	98
Het inspectieonderzoek stimuleert ons in onze kwaliteitsverbetering.	93	96	95	95
De tijd die wij in het onderzoek hebben geïnvesteerd staat in verhouding tot de opbrengst van het onderzoek voor ons.	86	87	87	83

Uit tabel 3 is af te lezen dat over de jaren heen vrijwel alle PO-scholen rapporteren positieve effecten van de schoolbezoeken te ervaren. Bijna alle basisscholen (> 85%) zijn van mening dat het inspectierapport voldoende gegevens bevat die in de toekomst bruikbaar zijn voor schoolontwikkeling en dat het toezicht stimuleert tot kwaliteitsverbetering. Iets minder positief zijn basisscholen over de verhouding van de tijd die het in het inspectie-onderzoek kost en de baten van het onderzoek, al vindt nog altijd 83 tot 87% de tijdsinvestering acceptabel. Het beeld is de afgelopen jaren redelijk constant gebleven, ongeacht het soort onderzoek waarbij de school betrokken is geweest.

Uit zowel het externe als het eigen onderzoek van de inspectie naar de opvattingen van scholen over het toezicht kan worden geconcludeerd dat scholen er in hoge mate tevreden over zijn en het nuttig vinden voor schoolverbetering. Dit type onderzoek betreft in alle gevallen zelfrapportages van scholen. Er is echter één studie waarin is nagegaan wat er feitelijk op scholen gebeurt voor, tijdens en na een inspectieonderzoek.

#### *Case studies*

Ehren (2006) onderzocht op tien PO-scholen wat het effect is van een inspectieonderzoek (PKO) waarbij de school door een inspecteur werd bezocht. In alle gevallen werden de directeur, de interne begeleider en een leerkracht voorafgaande aan en vier weken na het schoolbezoek ondervraagd en werd het bezoek geobserveerd. Drie maanden na het bezoek werd de directeur nogmaals bevraagd, na een half jaar gevolgd door een vragenlijst. Alle scholen bleken als gevolg van het bezoek verbeteringen door te voeren. Zij ontwikkelen bijvoorbeeld extra lesstof, of schakelen begeleiding in voor het verbeteren van het didactisch handelen. Ongeveer de helft van de verbeterprocessen in de scholen is gericht op condities van schooleffectiviteit op schoolniveau zoals de kwaliteit van het curriculum, de andere helft is gericht op condities op groepsniveau zoals gestructureerde instructie. De hoeveelheid verbeterprocessen varieerden van enkele kleine wijzigingen, zoals het afschaffen van een toets, tot grote veranderprocessen, zoals het invoeren van leerlijnen in de onderbouw en het aanpassen van het didactisch handelen van leerkrachten. De hoeveelheid verbeterprocessen

lijkt samen te hangen met de feedback die tijdens het bezoek door de inspecteur wordt gegeven, en de mate waarin deze samengaat met onvoldoendes in de beoordeling en afspraken over verbetering.

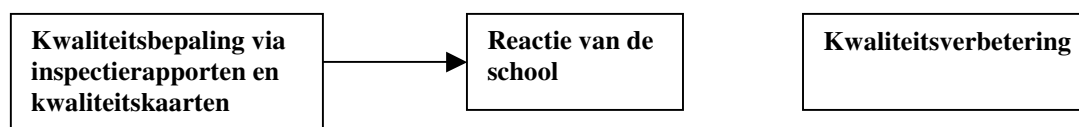
De verbeteringen hingen niet samen met de verandercapaciteit van de school, de steun vanuit de omgeving of met kenmerken van de inspecteur. Wanneer inspecteurs over zwakke punten feedback en verbeter suggesties gaven en met de school hierover afspraken maakten, ontstond een krachtige impuls voor verbetering.

### *Conclusie*

Uit de onderzoeken die gebaseerd zijn op zelfrapportages van scholen valt af te leiden dat de scholen zeer tevreden zijn met het toezicht, zoals dat onder de WOT is uitgevoerd. Een ruime meerderheid van de scholen rapporteert dat het inspectieonderzoek stimuleert tot kwaliteitsverbetering. Het onderzoek van Ehren (2006) ondersteunt deze conclusie en laat zien dat inspectieonderzoeken ook daadwerkelijk leiden tot activiteiten die gericht zijn op schoolverbetering.

## **4.4. EFFECTEN VAN DE KWALITEITSKAART**

Het eerste deel van de eerste onderzoeksvraag heeft ook betrekking op de relatie tussen openbare informatie over de kwaliteit van scholen (= inspectierapporten en kwaliteitskaarten) en de verbeteracties die scholen ondernemen op basis van de uitkomsten daarvan.



Mede onder invloed van de WOT maakt de inspectie de uitkomsten van inspectieonderzoeken ook openbaar via kwaliteitskaarten die primair bedoeld zijn voor ouders en deelnemers om zich een oordeel te vormen over de kwaliteit van scholen. De veronderstelling is dat deze openbare informatie ouders en deelnemers in staat stelt de schoolkeuze en de discussie met scholen over de onderwijskwaliteit te ondersteunen met objectieve gegevens (Janssens & Visscher, 2004; Janssens, 2005). Openbare informatie over de kwaliteit van scholen wordt geacht een vorm van publieke controle mogelijk te maken ('Vreemde ogen dwingen') waardoor de kwaliteit door de omgeving kan worden beïnvloed, hetzij doordat ouders zich met de kwaliteit van het onderwijs gaan bemoeien, dan wel 'stemmen met de voeten'. Ook zou van de openbaarheid een effect uitgaan omdat scholen hierdoor sterker met elkaar concurreren en er hierdoor een extra prikkel krijgen om goed te scoren op de kwaliteitskaart (Meijer, 2004; De Wolf & Janssens, 2005 en Janssens, 2005). Er is echter weinig onderzoek beschikbaar over het effect van openbare inspectierapportages en kwaliteitskaarten.

Dat ouders en deelnemers de kwaliteitskaarten niet primair raadplegen om in de discussie met scholen te gaan of de schoolkeuze te heroverwegen (Meijer, 2004; De Wolf & Janssens, 2005; Janssens, 2005 en Kenniscentrum voor Ordeningsvraagstukken, 2007), betekent niet dat deze kaarten geen effecten hebben op de kwaliteit van scholen. Volgens onderzoek naar de effecten van kwaliteitskaarten (Meijer, 2004), kunnen deze kaarten wel de reputatie van scholen beïnvloeden, alleen al omdat scholen elkaars kwaliteitskaarten onderling vergelijken. Dit heeft tot gevolg dat scholen zich proberen te verbeteren op punten waarin zij slechter scoren dan concurrerende scholen. Meijer laat zien dat scholen de openbaarheid van de inspectiegegevens ervaren als een stimulans om hun functioneren te verbeteren, o.a. door de gegevens te gebruiken als benchmark (Meijer, 2004). Leraren en directeuren blijken zelfs de

grootste gebruikers te zijn van kwaliteitskaarten (zie o.a. Meijer, 2004, p. 54; Inspectie van het Onderwijs, 2005). Ook uit het onderzoek van Q\*primair (Hofman & De Boom, 2006) blijkt dat de helft van de schoolleiders van mening is dat de internet-publicaties tot schoolverbetering leidt.

Sinds mei 2006 publiceert de inspectie maandelijks een dynamische lijst van zeer zwakke scholen voor (speciaal) basisonderwijs, expertisecentra en voortgezet onderwijs. Die lijst is primair bedoeld om ouders van deze scholen te informeren. Uit onderzoek blijkt dat deze lijst in ieder geval door de scholen niet altijd wordt geapprecieerd (Beek & Van den Heuvel, 2007). Het gaat daarbij niet om het feit dat de school als zeer zwak is bestempeld, want dat was de scholen bekend. Meerdere besturen gaven aan dat 'de rode kaart' van de inspectie is gewaardeerd: hiermee wordt duidelijk dat er echt iets moet gebeuren en dat wordt ook (sterker) opgepakt. De voornaamste kritiek had betrekking op het feit dat de school op een lijst stond en dat dit de aandacht van de ouders trok die niet op de hoogte waren van de status van de school. Blijkbaar leidt zo'n lijst wel tot (uitgestelde) reacties van de pers en van ouders. In ieder geval verstoort deze reacties, volgens de scholen, wel het inmiddels in gang gezette verbetertraject.

### *Conclusie*

Op basis van het beschikbare onderzoek naar de impact van openbare inspectierapporten en kwaliteitskaarten kan worden geconcludeerd dat – volgens de scholen – ook deze rapportages stimuleren tot verbeteracties omdat ze als benchmark worden gebruikt.

## **4.5. EFFECTEN VAN TOEZICHT OP DE KWALITEITSVERBETERING**

Het tweede deel van de eerste onderzoeksvraag betreft de relatie tussen toezichtactiviteiten en de kwaliteitsverbetering van scholen. Het ultieme effect van toezicht zou moeten zijn dat, vanwege de interventies van het toezicht, de kwaliteit van het onderwijs verbetert, gemeten aan de leerprestaties.



Er is echter voor de Nederlandse situatie maar één onderzoek bekend waarin de relatie tussen toezicht en de leerprestaties is onderzocht. Luginbuhl, Webbink en De Wolf (2007) vergeleken de veranderingen in de scores op de Eindtoets Basisonderwijs bij scholen die in enig jaar door de inspectie zijn bezocht met die van scholen die in dat jaar niet door de inspectie zijn bezocht in de periode 1999-2003 (N=6230). Om eventuele selectie-effecten bij het onderzoeken van scholen uit te sluiten is een eventueel effect op twee manieren onderzocht. In de eerste plaats via een 'standard fixed effect model' waarin gecontroleerd is voor schoolkenmerken die tijdsafhankelijk zijn. Hierdoor bleek het mogelijk dat de correlaties tussen inspectieonderzoeken en factoren als de fysieke infrastructuur van de school, de effectmaten niet beïnvloeden, maar mogelijk alleen de veranderingen in de kwaliteit van de scholen. Om voor deze vorm van bias te controleren is in de tweede plaats een random steekproef van scholen in de analyse betrokken. De inspectie trok in de bedoelde periode deze steekproef jaarlijks om scholen te selecteren die bezocht werden om gegevens te verzamelen voor de Onderwijsverslagen over de betrokken jaren (1999 N=370; 2000 N= 181;

2001 N= 199). Door gebruik te maken van gegevens over meerdere jaren konden de onderzoekers ook het effect schatten van inspectieonderzoeken over een periode van vier jaren.

In het onderzoek zijn twee typen inspectieonderzoeken vergeleken die van elkaar verschillen in de aantallen indicatoren die worden beoordeeld. De ene variant bevat alle indicatoren uit de toenmalige waarderingskaders (IST) en de andere slechts een selectie daaruit (RST). De scholen uit de steekroeven zijn door de inspectie steeds op alle indicatoren beoordeeld. In het CPB-onderzoek is over dezelfde periode de ontwikkeling van de toetscores van scholen na een bepaald type inspectieonderzoek vergeleken met de ontwikkeling in scores van scholen die zo'n inspectieonderzoek niet hebben gekregen.

Het onderzoek laat zien dat inspectieonderzoeken bijdragen aan de verbetering van leerprestaties. Twee jaar na een inspectieonderzoek bleken de toetsprestaties met twee tot drie procent van een standaarddeviatie gestegen te zijn. Voor rekenen/wiskunde is deze het sterkst en zelfs na vier jaren nog significant. Inspectieonderzoeken waar alle indicatoren zijn beoordeeld leiden tot hogere toetsprestaties, dan het andere type inspectieonderzoek. De analyse van de random steekproef laat weliswaar kleinere positieve effecten zien, maar deze zijn wel consistent met de andere bevindingen. Ook voor rekenen/wiskunde geldt in dit geval dat er een positief en significant effect is.

## 5. DISCUSSIE

De invoering van de WOT in 2002 heeft ertoe geleid dat vanaf dat moment alle PO-scholen jaarlijks aan enig inspectieonderzoek worden onderworpen. Gemiddeld genomen krijgt ongeveer een kwart van de scholen een kwaliteitsonderzoek. De rest van de scholen krijgt een zogenaamd jaarlijks onderzoek waarin een beperkt aantal indicatoren wordt beoordeeld, meestal de opbrengsten en de kwaliteitszorg. Het aantal jaarlijkse inspectieonderzoeken is dus sinds de invoering van de WOT weliswaar fors toegenomen, maar dit wordt vooral veroorzaakt door de invoering van het JO dat meestal een bureauonderzoek inhoudt. Het aantal scholen dat voor en na de WOT een kwaliteitsonderzoek krijgt is, gemiddeld genomen, ongeveer gelijk gebleven.

De onderzoeksvraagstelling luidde: Leiden de activiteiten van de Inspectie van het Onderwijs tot de verbetering van de kwaliteit van PO-scholen, gemeten aan zowel proceskenmerken als de leerprestaties? Op basis van een theoretisch model is verondersteld dat de volgende activiteiten onderdeel zijn van het 'werkzame bestanddeel' waarmee het effect van toezicht kan worden verklaard: de mondelinge terugkoppeling tijdens een schoolbezoek en de rapportages via inspectierapporten en kwaliteitskaarten naar aanleiding van een inspectieonderzoek. Vanwege het exploratieve karakter is de onderzoeksvraag uitgesplitst in de volgende deelvragen: is er sprake van kwaliteitsverbetering van PO-scholen, leidt het toezicht tot acties op scholen om de kwaliteit te behouden of te verbeteren en leiden deze op hun beurt ook tot daadwerkelijke kwaliteitsverbetering?

Op basis van de geanalyseerde gegevens kan worden geconcludeerd:

1. De kwaliteit van tweederde van de PO-scholen die de afgelopen tien jaar door de inspectie tenminste tweemaal zijn bezocht, is op belangrijke onderdelen gestegen, zoals de kwaliteitszorg, de zorg & begeleiding en de opbrengsten. Dat geldt de hele populatie,

maar nog meer voor de scholen waar het toezicht is geïntensiveerd omdat de inspectie risico's ziet of omdat de scholen als 'zeer zwak' zijn beoordeeld.

2. Dat inspectieonderzoeken bijdragen aan kwaliteitsverbetering blijkt uit een quasi-experimenteel onderzoek van het CPB. In dit onderzoek is het effect van twee typen inspectieonderzoeken onderzocht op de scores van de Eindtoets basisonderwijs van het Cito bij scholen die in enig jaar wel en niet zijn bezocht. Twee jaar na een inspectieonderzoek bleken de toetsprestaties significant te zijn gestegen, voor rekenen/wiskunde was het effect zelfs na vier jaren nog significant.
3. Tussen de 80 en 90% van de PO-scholen rapporteert dat inspectieonderzoeken stimuleren tot schoolverbetering. Ondanks het feit dat het hier gaat om gegevens die de onderwijsinspectie zelf heeft verzameld, worden de bevindingen bevestigd door extern uitgevoerd onderzoek. Ook onderzoek naar kwaliteitskaarten laat zien dat deze vorm van openbare terugkoppeling scholen stimuleert tot kwaliteitsverbetering.
4. Uit case studies blijkt dat scholen die door de inspectie met een PKO zijn onderzocht een scala aan verbeteractiviteiten entameren. Er zijn dus indicaties dat inspectieonderzoeken waarbij scholen worden bezocht, daadwerkelijk leiden tot concrete verbeteractiviteiten.

Op basis van het bovenstaande lijkt de conclusie gerechtvaardigd dat er een positieve samenhang is tussen toezicht en de kwaliteitsverbetering van scholen, maar het is niet duidelijk waardoor het effect van het toezicht precies wordt veroorzaakt. Wel is aannemelijk dat de mate waarin sprake is van kwaliteitsverbetering samen lijkt te hangen met de feedback op basis van een inspectieonderzoek waarin zowel onderwijsprocessen als de leerresultaten worden beoordeeld. Hiervoor gelden de volgende argumenten.

Alle scholen die twee maal met een kwaliteitsonderzoek zijn bezocht hebben zich verbeterd, maar de scholen die risicovol of zeer zwak zijn, het meest. Deze laatste scholen worden door de inspectie veelal jaarlijks bezocht met een PKO of een JO-bezoek om de voortgang te monitoren. In veruit de meeste gevallen vindt op deze scholen het PKO eerder plaats dan de reguliere termijn van vier jaar. Daarmee is niet gezegd dat er een causaal verband is tussen toezicht en kwaliteitsverbetering, maar wel dat deze scholen vaker en intensievere feedback van de inspectie op hun prestaties en functioneren krijgen, dan scholen die in een vierjaarscyclus zitten.

Het onderzoek van Ehren (2006) laat niet alleen zien dat elk inspectieonderzoek leidt tot een serie verbeteractiviteiten, maar ook dat de omvang en reikwijdte van deze activiteiten lijken samen te hangen met de aard van de feedback van de inspecteur, zoals het geven van oordelen over sterke en zwakke punten van het onderwijs, het geven van verbeter suggesties en het maken van afspraken over de verbeteractiviteiten.

De CPB-studie (Luginbuhl, Webbink & De Wolf, 2007) suggereert dat het wat uitmaakt welke type onderzoek de inspectie inzet. De door hen geconstateerde effecten zijn groter bij een inspectieonderzoek waarin alle indicatoren aan de orde kwamen en dat twee schoolbezoekdagen duurde (IST), dan bij een onderzoeksvorm waarbij alleen de belangrijkste kernindicatoren worden beoordeeld en dat binnen één werkdag is afgerond (RST). De twee typen onderzoeken verschillen ook nog op andere punten van elkaar. Bij IST werden meer lessen geobserveerd, meer personeelsleden (vaker) geïnterviewd en werden ook ouders en leerlingen bevraagd. De terugkoppeling bij een IST was gericht op het hele schoolteam en bij RST alleen op de personeelsleden die bij het onderzoek waren betrokken. Dat een IST meer effect heeft dan een RST zou dus kunnen worden verklaard door de aard en de omvang van de terugkoppeling. Omdat bij een IST alle kwaliteitsaspecten werden beoordeeld en bij het RST alleen de kernindicatoren, ligt het voor de hand dat een IST meer aanknopingspunten biedt voor verbetering dan een RST.

Er zijn dus indicaties dat het effect van toezicht op schoolverbetering samenhangt met de aard, de omvang en de frequentie van de feedback op zowel proces- als opbrengstkenmerken (zie ook Matthews & Sammons, 2004 en Ehren & Visscher, 2006). Deze vorm van feedback vormt de kern van de stimulerende taak van de inspectie (zie Onderwijsraad, 2006). De mechanismen die in het toezicht een rol spelen en tot verbeterprocessen en kwaliteitsverbetering van scholen leiden, zijn nog maar beperkt in kaart gebracht. Empirische gegevens over de relatie tussen toezicht en schoolverbetering geven aanleiding te veronderstellen dat uitgebreide inspectiefeedback meer effect heeft dan feedback op een beperkt aantal kwaliteitsindicatoren en dat het toezicht het meeste effect heeft op risico- en zeer zwakke scholen. Er is dus veel voor te zeggen deze scholen met een zo uitgebreid mogelijk instrumentarium te onderzoeken waarin zowel proceskenmerken als opbrengsten een rol spelen, daarop uitgebreide mondelinge en schriftelijke feedback te geven en de voortgang op de scholen nauwgezet te volgen. Nader onderzoek zal echter moeten uitwijzen welke vorm van feedback en monitoring het meest effectief is, ook voor de scholen die minder intensief worden onderzocht dan de risico- en zeer zwakke scholen.

## 6. LITERATUUR

### 7.

- AVS (2004). *Scholenpanel Kwaliteitszorg en inspectie*. AVS Kadernieuws, februari 2004 ([www.avs.nl](http://www.avs.nl))
- Algemene Rekenkamer (2005). *Handhaven en Gedogen: TK 2004-2005*, 30 050 nrs.1-2, pag. 5 en pag. 20.
- Beek, S & Van den Heuvel, P. (2007). *Het stoplicht sprong op rood! Verslag van onderzoek naar bestuurlijke interventies naar zeer zwakke scholen*. Den Bosch: KPC groep.
- Braster, J.F.A. (1987). *De inspecteur en de school*. Rotterdam: Erasmus Universiteit.
- Dijk, van, B. (2003). Kennisbeperkingen in de handhavingspraktijk: weten of wedden? *Overheidsmanagement*, 16(4),7-10
- Ehren, M.C.M. (2006). *Toezicht en schoolverbetering*. Delft: Eburon (diss.)
- Ehren, M.C.M, Leeuw, F.L. & Scheerens, J. (2005) On the Impact of the Dutch Educational Supervision Act: Analyzing Assumptions Concerning the Inspection of Primary Education. *American Journal of Evaluation*. (26) 1, 60-7
- Ehren, M.C.M. & Visscher, A.J. (2006). Towards a theory on the impact of school inspections. *British Journal of Educational Studies*. (54)1, 51-72
- Emmelot, Y., Karsten, S., Ledoux, G, Vermeulen, A. (2004). *Ervaringen met het vernieuwde toezicht*. Amsterdam: SCO-Kohnstamm Instituut.
- Esch, W. van, & Laemers, M. (1990). *De onderwijsinspectie en de wet: een onderzoek naar toepassing en interpretatie van wettelijke regelingen door de onderwijsinspectie*. Nijmegen: ITS
- Hofman, R.H., Dijkstra, N.J., De Boom, J. & Hofman, W.H.A. (2005). *Kwaliteitszorg in het primair onderwijs: eindproject*. Groningen: GION (BOPO-project 412-02-010)
- Hofman, R.H. & De Boom, J. (2006) *Q\*Primair monitor kwaliteitszorg ,Eindrapportage* Groningen: GION Rijksuniversiteit Groningen, J. de Boom, RISBO Erasmus Universiteit Rotterdam
- Janssens, F.J.G. (2005). *Toezicht in discussie: over onderwijstoezicht en Educational Governance*. Enschede: Universiteit Twente (Oratie)
- Janssens, F.J.G. (2006). *De inspectie komt op bezoek*. Alphen a/d Rijn: Kluwer.
- Janssens, F.J.G., & Visscher, A.J. (2004). Naar een kwaliteitskaart voor het primair onderwijs. In: *Pedagogische Studiën*. (81)371-383.

- Kenniscentrum voor Ordeningsvraagstukken (2007). *De psychologie van het kiezen: over consumentengedrag in geliberaliseerde markten*. Den Haag: Ministerie van Economische zaken
- Leeuw, F. & Willemsen, F. (2006). Toezicht en inspectie: trends, kosten en baten. *NJB*, 37, 2108-2116
- Luginbuhl, R., Webbink, D. & De Wolf, I. (2007). *Measuring the effect of school inspections on primary school performance: a study based on Cito test scores*. Den Haag: CPB. CPB discussion paper
- Matthews, P & Sammons, P. (2004). *Improvement through inspection: an evaluation of the impact of Ofsted's work*. London: Institute of Education, University of London/Ofsted
- McCrone, T., Rudd P., Blekinsop, S. & Wade, P. (2006). *Impact of section 5 inspections: maintained schools in England*. London: Nfer
- Meijer, A. (2004). *Vreemde ogen dwingen. De betekenis van internet voor maatschappelijke controle in de publieke sector*. Den Haag: Boom Juridische uitgevers
- Onderwijsraad (2006). *Doortastend onderwijstoezicht*. Den Haag: Onderwijsraad.
- Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of Education Review*, 23(2), 143-152.
- Vanhoof, J & Van Petegem, P. (2007). Kwaliteitsvolle zelfevaluaties: resultaat en proces door de ogen van Vlaamse deskundigen. Paper gepresenteerd op de ORD 2007 te Groningen.
- Visscher, A.J. (2002). A framework for studying school performance feedback systems. In: Visscher A.J. & Coe, R. (eds). *School improvement through performance feedback*. Lisse: Swets&Zeitlinger
- Visscher A.J. & Coe, R. (2003). School performance feedback system: conceptualization, analysis and reflection. *School Effectiveness and School Improvement*, 14(3),321-350
- Vossen, R. (vz) (2005). *Een waarneembare verandering in hetgeen met de toezichts- en handhavingsactiviteiten van de inspectie wordt beoogd als gevolg van de activiteiten van die inspectie*. Den Haag: IG-beraad
- Wolf, I.F. de & Janssens, F.J.G. (2005). *Effects and side effects of inspections and accountability in education; an overview of empirical studies*. Amsterdam: Scholar <http://www.ase.uva.nl/scholar/workingpapers.cfm#UUIIDD2B24045ZA1CEZ4535Z90E154CE9D845FB0>