

Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 van het basisonderwijs in 2014

Jaarlijkse Meting Taal en Rekenen 2014 – Jaarlijks peilingsonderzoek naar het onderwijsniveau
Bas Hemker en Jan van Weerden



Jaarlijkse Meting Taal en Rekenen

Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 van het basisonderwijs in 2014

Bas Hemker en Jan van Weerden

© Stichting Cito Instituut voor Toetsontwikkeling Arnhem (2014)

Niets uit dit werk mag zonder voorafgaande schriftelijke toestemming van Stichting Cito Instituut voor Toetsontwikkeling worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotografie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook.

Inhoud

Voorwoord 5

Samenvatting 7

1 Inleiding 9

2 Vraagstelling en aanpak 11

- 2.1 Vraagstelling 11
- 2.2 Methode 11
 - 2.2.1 Kwaliteit van de meetinstrumenten 12
 - 2.2.2 Steekproeftrekking en analyse 12
 - 2.2.3. Achtergrondvariabelen 14
- 2.3 Hoe is er geanalyseerd? 15
 - 2.3.1 Vergelijking tussen en over de jaren 15
 - 2.3.2 Analyses voor de leerlingen in 2014 16
 - 2.3.3 Significantie en effectgrootte 16
 - 2.3.4 De gekozen rapportageschaal 17

3 De resultaten voor jaargroep 8 19

- 3.1 De vergelijking over de jaren 19
 - 3.1.1 Taalvaardigheden 20
 - 3.1.2 Rekenen 22
- 3.2 Verschillen voor achtergrondvariabelen 23
 - 3.2.1 Effecten op leerlingniveau 25
 - 3.2.2 Effecten op schoolniveau 38

4 Conclusies 43

- 4.1 Algemeen beeld 43
- 4.2 Invloed van achtergrondvariabelen 43
- 4.3 Discussie 44

Literatuur 47

Bijlagen

- 1 Gemiddelden en standaarddeviaties per vaardigheid gecategoriseerd naar achtergrondvariabele in jaargroep 8 in 2014 50
- 2 Effectschattingen voor alle vaardigheden gecategoriseerd naar achtergrondvariabelen in jaargroep 8 in 2014 51
- 3 De effecten per vaardigheid per jaar over de jaren heen: Effectschattingen voor alle vaardigheden gecategoriseerd naar achtergrondvariabelen in jaargroep 8 per jaar voor alle peilingen vanaf 2008 52
- 4 De gemiddelde vaardigheid per vaardigheid per jaar over de jaren heen: Gemiddelden en standaarddeviaties per vaardigheid gecategoriseerd naar achtergrondvariabele in jaargroep 8 per jaar voor alle peilingen vanaf 2008 58
- 5 Verdeling van de achtergrondvariabelen in de onderzoekspopulatie in 2014 die een eindtoets gemaakt heeft 64

Voorwoord

Dit rapport vormt de neerslag van de zevende meting van de taal- en rekenvaardigheden van leerlingen in het basisonderwijs sinds 2008. De meting is van start gegaan naar aanleiding van de kwaliteitsagenda voor het Primair Onderwijs, verschenen in 2007 onder de titel 'Scholen voor morgen' en wordt sindsdien gecontinueerd als landelijke monitor voor een aantal basisvaardigheden.

Bij de vorige metingen werden naast de metingen in groep 8 ook metingen gedaan in groep 4. Dat is dit jaar niet meer het geval. Wat wel hetzelfde gebleven is, is dat bij de uitvoering van dit onderzoek gebruik wordt gemaakt van data uit een toetsstelsel dat al bij de meeste scholen wordt gebruikt voor het in kaart brengen van onderwijsopbrengsten bij rekenen en taal in jaargroep 8, namelijk de Eindtoets. Door deze keuze wordt de extra toetslast voor leerlingen en leerkrachten beperkt.

De uitkomsten van deze jaarlijkse peiling geven een indicatie voor de stand van zaken met betrekking tot een belangrijk deel van de referentieniveaus behorende bij de doorlopende leerlijnen zoals geformuleerd in het advies van de Expertgroep Doorlopende leerlijnen *Over de drempels met taal en rekenen*.

De resultaten in dit onderzoek worden meegenomen in het jaarlijkse verslag *De staat van het onderwijs* van de Inspectie van het onderwijs.

Bij de uitvoering van het project zijn diverse mensen betrokken geweest, waarvan er enkele met name genoemd moeten worden. Sylvia Walraven was bij de ondersteuning zeer belangrijk en Anne-Marie Anthonissen heeft in de coördinatie veel betekend. Ten slotte wil ik graag Loes Hiddink bedanken voor de bureauredactie.

We hopen dat dit rapport en de daarin beschreven resultaten hun weg vinden naar het onderwijsveld en de beleidsmakers.

Dr. B.T. Hemker
Senior onderzoeker

Samenvatting

In dit rapport wordt verslag gedaan van de zevende jaarlijkse peiling van de reken- en taalvaardigheid in jaargroep 8 sinds 2008 (JMTR, voorheen JPON). Hiervoor is gebruikgemaakt van gegevens uit de Eindtoets van 2014 voor Woordenschat, Spelling, Begrijpend lezen en Rekenen-Wiskunde.

Bij een vergelijking van de resultaten van 2014 met die van 2013 blijken er nauwelijks verschillen te zijn. Terwijl er in de periode 2009-2012 sterke wisselingen optraden bij de drie taalvaardigheden, met ups en downs, kunnen we constateren dat er in de laatste drie jaar een stabiel beeld is ontstaan. De uitkomsten voor de verschillende taalonderwerpen hebben zich sinds 2011 op een meer onderling vergelijkbare manier ontwikkeld. De gemiddelde stijging in vaardigheid is hierdoor nu van 2008 tot en met 2014 over alle drie de taalvaardigheden zeer vergelijkbaar. Over de hele periode van zeven jaar bezien, is de gemiddelde groei ongeveer 1% van de standaarddeviatie per jaar. Al is de gemiddelde groei gelijk, er moet wel opgemerkt worden dat deze groei grillig was en niet voor alle vaardigheden ieder jaar op dezelfde wijze verliep.

Voor de rekenonderwerpen zien we ook een stabilisatie op een iets hoger niveau dan in 2008. Voor deze drie onderwerpen geldt echter dat ze sinds het begin van de meting minder grote onderlinge fluctuaties hebben vertoond dan de resultaten bij taal. Wel zien we dat er na een stijging in niveau in respectievelijk 2010 en 2012 er nu een stabilisatie heeft plaatsgevonden op een iets lager niveau, vergelijkbaar met dat van 2010. Ook zien we dat de groei bij de vaardigheid Getallen en bewerkingen iets achterblijft bij die van de vaardigheden Breuken, procenten en verhoudingen en Meten, meetkunde, tijd en geld. De gemiddelde groei is voor ieder van de afzonderlijke rekenschalen, net als bij de taalvaardigheden, over de zeven jaar ook ongeveer 1% van de standaarddeviatie per jaar. De groei was over de rekenvaardigheden heen minder grillig dan bij de taalvaardigheden, maar er was geen sprake van een constante groei: er zijn ook tussentijdse dalingen geweest.

Bij elkaar genomen blijkt dat in 2014 zowel taal als rekenen een stabiel en consistent beeld geven. In beide gevallen is de groei vergelijkbaar met een gemiddelde van iets minder dan 1% van de standaarddeviatie per jaar. Op zich lijkt een dergelijke groei niet opzienbarend, maar over de jaren en vaardigheden heen bezien is dit toch een redelijk positief resultaat.

Dit resultaat is vooral ook positief te noemen als we naar de verschuiving in leeftijd van de leerlingen kijken, vooral ten gevolge van de afname van het aantal vertraagde leerlingen in de periode van 2008 tot en met 2014. Bij de vergelijking van de vaardigheid over de gehele populatie is te zien dat dit – over de jaren heen – niet ten koste is gegaan van het prestatieniveau. Sterker nog: daar is eerder sprake van een stijging. Het zou betekenen dat hier een efficiëntieslag gemaakt is door het onderwijs.

Bij de vergelijking van de vaardigheid van de twee onderscheiden groepen bij de variabele leertijd zien we dat de vertraagde leerlingen aanzienlijk minder goed presteren dan de niet-vertraagde leerlingen (een verschil van meer dan 50% van de standaarddeviatie). Het verschil tussen de twee groepen verandert niet substantieel over de jaren heen. Dergelijke vergelijkingen zijn ook bij andere variabelen onderzocht.

De verschillen in prestaties zijn tussen jongens en meisjes zijn vergelijkbaar met vorige jaren. Wel zien we dat in de loop van de tijd de prestaties van jongens en meisjes dichter bij elkaar komen te liggen. Bij de variabele formatiegewicht valt vooral op dat Woordenschat de enige vaardigheid is waarbij leerlingen met een laag formatiegewicht (0.30) significant hoger scoren dan leerlingen met een hoog formatiegewicht (1.20). Meestal is dat effect andersom.

Door de jaren heen is het niveau van de leerlingen die geadviseerd worden op specifieke niveaus van het voortgezet onderwijs in te stromen niet fundamenteel gewijzigd. Het verschil in vaardigheid verschilt enorm tussen leerlingen met de verschillende doorstroomadviezen. Gemiddeld is het verschil tussen de

vaardigheid van een vmbo-BB-leerling en een vwo-leerling meer dan 4 standaarddeviaties. Het geeft aan hoe groot de vaardigheidsverschillen aan het eind van de basisschool al zijn.

De variabele thuistaal heeft geen toegevoegd effect naast formatiegewicht, stratum, leertijd en geslacht bij Spelling en Rekenen, maar wel bij Begrijpend lezen en vooral bij Woordenschat. Over het algemeen geldt voor deze vaardigheden: hoe meer Nederlands de leerling thuis spreekt, hoe hoger de vaardigheid.

De variabelen regio, verstedelijking, schoolgrootte en denominatie hebben geen toegevoegd effect naast de variabelen formatiegewicht, leertijd, geslacht en stratum.

Er is geen onderzoek gedaan naar mogelijke oorzaken voor de uitkomsten over de jaren heen. Zo hebben we, anders dan bij regulier peilingsonderzoek (PPON), geen zicht op het onderwijsaanbod of veranderingen daarin, laat staan dat dergelijke gegevens direct gekoppeld kunnen worden aan de leerlingprestaties.

Dat maakt het lastig om aan deze uitkomsten een interpretatie te verbinden. We weten dus niet op welke wijze bijvoorbeeld het gebruik van andere lesmethoden of het besteden van meer tijd aan bepaalde onderwerpen te maken heeft met deze resultaten. Om een relatie te leggen tussen de gesignaleerde vooruitgang en wijzigingen in onderwijsbeleid is nader onderzoek nodig naar tussenliggende factoren.

Inleiding

In het kader van de kwaliteitsagenda 'Scholen voor morgen' is in 2008 het Jaarlijks Peilingsonderzoek naar het Onderwijsniveau (JPON) van start gegaan. Dit jaarlijkse onderzoek is gericht op het monitoren van het onderwijsniveau op het gebied van taal- en rekenvaardigheid in het basisonderwijs. In 2008 is verslag gedaan van de eerste jaarlijkse niveaupeiling van taal en rekenen in jaargroep 8 en 4 (Hemker & Van Weerden, 2009). Het project was ondergebracht bij PPO, de periodieke peiling van het onderwijsniveau, maar kent een andere methodiek en gebruikt andere instrumenten. De overeenkomst is dat het ook hier gaat om periodiek onderzoek, zij het met een hogere frequentie, en dat ook hier resultaten op een vaste meetschaal worden gerapporteerd, wat maakt dat de uitkomsten over de jaren heen vergelijkbaar zijn. De belangrijkste verschillen zijn te vinden in de gebruikte instrumenten, de wijze van afnemen en de beschikbare achtergrondgegevens.

In de loop van de tijd is de meting meer bekend geraakt als de Jaarlijkse Meting Taal & Rekenen (JMTR). Het uitgangspunt van het onderzoek is echter onveranderd en sluit aan op de kwaliteitsagenda 'Scholen voor morgen'. Deze kwaliteitsagenda was gericht op duurzame verbetering van het primair onderwijs en legde de prioriteit bij de verhoging van de taal/lees- en rekenopbrengsten. Dit is gecontinueerd in het beleid van daaropvolgende kabinetten en heeft onder andere geleid tot nieuwe wetgeving en voorstellen daartoe, waarin de positie van rekenen en taal en het toezicht daarop verder is verstevigd.

In de aanbevelingen van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen (EGDLL, ofwel de commissie Meijerink, SLO, 2008) werd vastgelegd wat leerlingen moeten kennen en kunnen op dit gebied bij het verlaten van de basisschool: de referentieniveaus voor taal en rekenen. Dit is bij wet geregeld in 2010. Om vast te stellen wat de beginsituatie was en hoe het niveau zich sindsdien heeft ontwikkeld in het basisonderwijs is de Jaarlijkse Meting Taal & Rekenen (JMTR; voorheen JPON) gestart.

Scholen beschikken doorgaans over voldoende toetsgegevens, gebaseerd op leerlingvolgsystemen en eind- of entreetoetsen, als input voor verbeterplannen op groeps-, school- en bestuursniveau. Deze gegevens kunnen echter ook, onder bepaalde condities, worden ingezet voor een landelijke niveaubepaling, zoals dat in het project JMTR wordt gedaan. Wel moet worden vastgesteld dat de bestaande toetsystemen inhoudelijk nog niet zijn afgestemd of ingericht op de referentieniveaus van de commissie Meijerink. De toetsen die nu ingezet kunnen worden zijn nog niet dekkend voor het hele terrein dat met de referentieniveaus wordt beschreven. Met name de productieve vaardigheden bij taal (schrijven en spreken) ontbreken nog, net als taalvaardigheden als begrijpend luisteren of grammaticaliteit (of kunnen omgaan met elementen uit de 'begrippenlijst'). De toetsresultaten geven echter wel een duidelijke indicatie van de stand van zaken op een aantal belangrijke vaardigheden en maken het ook mogelijk daarin de ontwikkeling van jaar tot jaar te laten zien.

In 2008 is deze jaarlijkse peiling voor de eerste keer uitgevoerd. Daarmee is een start gemaakt met een reeks van jaarlijkse rapporten, waarin het niveau van taal en rekenen in het primair onderwijs in beeld wordt gebracht. Daarbij is 2008 steeds het uitgangspunt voor de vergelijking. Door te vergelijken met eerdere metingen kan in beeld worden gebracht hoe de beheersing van de vaardigheid door de tijd verandert. Een van de manieren waarop dat gebeurt, is door te kijken hoeveel procent van de leerlingen een bepaald niveau door de tijd heen halen. Dit kon in 2008 echter nog niet gedefinieerd worden in termen van de referentieniveaus aangezien die nog niet bepaald waren.

Een exacte uitspraak met betrekking tot de behaalde referentieniveau is ook in 2014 nog niet mogelijk aangezien de eindtoetsdata nog niet in het kader van de referentieniveaus verzameld zijn. De koppeling Eindtoets – Referentieniveaus is daardoor niet goed zonder extra aannamen te maken¹.

Het verslag begint met een beschrijving van de opzet van het onderzoek, de gebruikte toetsen en de wijze waarop deze toetsen met elkaar vergeleken kunnen worden. Vervolgens wordt de onderzoekspopulatie beschreven en worden de definities van de achtergrondvariabelen gegeven (hoofdstuk 1). Ook wordt beschreven hoe we de resultaten van de onderzoekspopulatie op de gemeten vaardigheden kunnen vergelijken met voorgaande jaren. Vervolgens beschrijven we de resultaten van de leerlingen in jaargroep 8 (hoofdstuk 2). Hierin wordt de jaarvergelijking met de voorgaande jaren gegeven. Daarna volgt een prognose van het percentage leerlingen dat het niveau 1F en het niveau 1S gehaald heeft voor de vaardigheden Begrijpend lezen en Rekenen (hoofdstuk 3). Deze prognoses worden echter gedaan met een aantal aannamen en geven dus niet met zekerheid deze percentages aan. Daarvoor moeten we de resultaten van 2015 afwachten. In de hoofdstukken 2 en 3 geven we de resultaten voor de onderwerpen voor rekenen en taal, en de relatie van deze uitkomsten met enkele achtergrondvariabelen zoals geslacht en thuistaal. Het laatste hoofdstuk bevat een samenvatting, conclusie en discussie.

¹ Voor de vaardigheden die wel overeenkomen met die van de referentieniveaus, zoals begrijpend lezen en rekenen kan er eventueel onder strikte aannamen wel een indicatie gegeven worden van hoeveel procent van de verschillende groepen van leerlingen de niveaus 1F en 1S behaalt. Doordat deze aannamen nu hier niet te toetsen zijn, worden deze resultaten hier niet gepubliceerd. Het is de verwachting dat dit in 2015 wel mogelijk is.

2 Vraagstelling en aanpak

2.1 Vraagstelling

Het belangrijkste doel van de Jaarlijkse Meting Taal en Rekenen (JMTR) is het nauwkeurig vaststellen van veranderingen in de taal- en rekenvaardigheden van leerlingen in jaargroep 8 van het basisonderwijs over de jaren heen. Daarnaast is het doel na te gaan in hoeverre de prestatieverschillen tussen bepaalde groepen leerlingen gelijk zijn gebleven dan wel groter of kleiner zijn geworden. In JPON werden ook de resultaten van groep 4 in het onderzoek betrokken, maar dat is vanaf 2014 niet meer het geval.

2.2 Methode

Voor deze jaarlijkse peiling maken we gebruik van instrumenten die reeds in een ander kader worden ingezet. Het zijn dus geen nieuwe toetsen die afzonderlijk voor dit project zijn ontwikkeld. We maken hierbij voor jaargroep 8 wel gebruik van de toetsen die vallen onder de Eindtoets Basisonderwijs. Deze toetsen bevatten een aantal onderdelen voor rekenen en taal die we goed voor dit doel kunnen gebruiken, maar die elk jaar volledig worden vernieuwd. Omdat de Eindtoets elk jaar andere items bevat, kunnen de resultaten niet zomaar van jaar tot jaar met elkaar worden vergeleken. Door gebruik te maken van aanvullende databronnen is hier toch een goede oplossing voor gevonden. We beschrijven dit verderop in dit hoofdstuk (par. 2.2.1).

De hier genoemde toetsen zijn in aantallen items en mogelijk te onderscheiden onderwerpen lang niet zo gedifferentieerd als in het reguliere onderzoek van PPON tot nu toe gebruikelijk was. Het aantal rapportage-eenheden is daardoor beperkt tot een zevental onderwerpen in jaargroep 8, daar waar er in de laatste rekenpeiling voor deze groep (Scheltens, Hemker, & Vermeulen, 2013) alleen al onderscheid gemaakt werd tussen meer dan 20 subvaardigheden binnen het reken-wiskundeonderwijs. Daardoor was het in PPON bijvoorbeeld mogelijk de vooruitgang op schattend rekenen waar te nemen, en tegelijkertijd de verminderde vaardigheid in vermenigvuldigen en delen. Ook worden bij PPON meer en andere itemvormen ingezet dan alleen de vaste vierkeuzevraag die in de Eindtoets gebruikt wordt.

Een belangrijk verschil met PPON is verder de afnameconditie. Bij PPON worden alle toetsen en taken afgenomen door een getrainde toetsleider. In het geval van de JMTR is de eigen leerkracht de toetsleider aangezien deze de Eindtoets afneemt. Dat vormt een mogelijke aantasting van de objectiviteit van de afname. Een verschil dat mogelijk een positief effect heeft op de in dit onderzoek gerapporteerde resultaten is het verschil in impact van de verschillende toetssituaties. De Eindtoets kan worden beschouwd als een 'high-stake' toets, een toets waar leerlingen hun uiterste best op zullen doen, niet alleen vanwege de gevolgen van de uitslag voor hun verdere toekomst, maar ook vanwege de grote mate van aandacht en publiciteit eromheen. Voor de afnames voor PPON is dat niet het geval. Deze verschillen in conditie hebben hun weerslag op de hoogte van de toetsscores, zo is inmiddels uitgezocht (zie verder Hemker, 2013).

Voor een vergelijking over de jaren heen zijn deze risico's van minder groot belang, omdat de condities immers in de meeste gevallen niet over de jaren heen zullen variëren. Het is dus geen probleem om een trend in beeld te brengen. Voor een vergelijking van deze uitkomsten met die van andere onderzoeken zijn ze dat echter wel. Zo zullen resultaten op de Eindtoets een ander beeld kunnen geven van de mate waarin de landelijk vastgestelde referentieniveaus zijn gehaald dan de uitkomsten van een meting met het leerlingvolgsysteem of een PPON-meting.

Uit overwegingen van efficiëntie en kosten zijn voor de jaarlijkse niveaupeilingen dus overwegend gegevens gebruikt die toch al door scholen worden verzameld. Voor de meting van de taal- en rekenvaardigheid in jaargroep 8 zijn dat de verzamelde gegevens uit campagne van de Eindtoets


Basisonderwijs 2014. In deze campagne wordt een aantal toetsen afgenomen. Om te beginnen is dat de reguliere eindtoets die door veruit de meeste kinderen wordt gemaakt. Dat is een toets op papier die in een vaste periode van drie dagen in februari gemaakt wordt. In 2014 was dat van 11 tot en met 13 februari.

Daarnaast is gebruikgemaakt van gegevens die verzameld zijn met twee toetsen die in exact dezelfde periode worden afgenomen en een identieke functie hebben: de Eindtoets Niveau en de ankertoets. Om de gegevens voor Spelling van de verschillende jaren te kunnen vergelijken wordt voor de equivalering gebruikgemaakt van de resultaten van de op een beperkt aantal scholen afgenomen toets Basisvaardigheden. De resultaten van de leerlingen op deze toets zelf worden echter niet hier gerapporteerd aangezien deze in een andere periode worden afgenomen. Dat zelfde geldt voor de digitale varianten van de Eindtoets. Ook deze worden niet binnen de vaste periode afgenomen. De resultaten op deze digitale toetsen vallen daardoor ook buiten dit onderzoek.

Voor taal hebben we de onderwerpen Begrijpend lezen (BL), Spelling (Sp) en Woordenschat (Wo) onderzocht. Bij rekenen gaat het om de onderdelen Getallen en bewerkingen (GB), Breuken, procenten en verhoudingen (BPV) en Meten, meetkunde, tijd en geld (MMTG). De drie rekenvaardigheden blijken hoog samen te hangen en de resultaten die bij deze schalen in de afgelopen jaren zijn gevonden lijken ook sterk op elkaar, zodat het bij rekenen ook mogelijk is om over een samengestelde rekenschaal te rapporteren. Dat is niet het geval bij de taalvaardigheden: de samenhang tussen de onderwerpen is lager dan bij rekenen en de resultaten verschillen per subschaal. De resultaten op een samengestelde schaal voor taal zouden sterk afhankelijk zijn van de specifieke samenstelling van deze schaal.

2.2.1 Kwaliteit van de meetinstrumenten

De Eindtoets Basisonderwijs (EB) wordt elk jaar opnieuw samengesteld en bestaat uit opgaven die al in proeftoetsen in voorgaande jaren zijn uitgetoetst. Dat geldt ook voor de Eindtoets Niveau en de ankertoets. De psychometrische kenmerken zijn van tevoren goed in te schatten en blijken telkens van hoog niveau te zijn. De eigenschappen van de Eindtoets worden elk jaar gepubliceerd in een afzonderlijk document en dat is ook voor 2014 het geval (Cito, 2014). We mogen hier stellen dat de psychometrische kwaliteit bij de Eindtoets gewaarborgd is door de gehanteerde procedures. Nadere informatie hierover is te vinden in de uitgebreide verantwoording van de Eindtoets van 2010 (Van Boxtel e.a., 2012; zie ook www.toetswijzer.nl). Deze verantwoording is aangeboden aan de COTAN (Commissie Test Aangelegenheden Nederland) en de toets is op alle relevante criteria als voldoende en goed beoordeeld.

Het is belangrijk hier te wijzen op het feit dat de wijze waarop deze toets en de versies daarvan in dit onderzoek worden ingezet wel een andere is dan waarvoor zij is bedoeld. Het primaire doel van de Eindtoets Basisonderwijs is namelijk het opleveren van een goede prognose voor de kans van slagen in verschillende vormen van voortgezet onderwijs voor individuele leerlingen. In deze studie gaat het echter niet om de totaalscore op alle onderdelen, de standaardscore, maar om de uitkomsten op delen van de toets. In de JMTR worden verschillende onderwerpen van de toets afzonderlijk gebruikt om een beeld te schetsen van de kwaliteit van het onderwijs op systeemniveau. Voor dat doel zijn de psychometrische eisen die aan een toets worden gesteld anders. Met dat fenomeen is rekening gehouden in de analyses, op een vergelijkbare wijze als in voorgaande jaren (zie verder Hemker & Van Weerden, 2009; Hemker, Kuhlemeier & Van Weerden, 2010; Hemker, Kordes en Van Weerden, 2011; Hemker en Van Weerden, 2012; Van Weerden, Hemker, Straat & Mulder, 2013; Van Weerden, Hemker, Straat & Mulder, 2014 

2.2.2 Steekproeftrekking en analyse

In jaargroep 8 is er sprake van een steekproefprocedure, zodat de resultaten als representatief mogen worden gezien voor het niveau in Nederland op het betreffende meetmoment. Voordat deze steekproef getrokken werd, is eerst bepaald welke leerlingen buiten de onderzoekspopulatie vallen. Achterwege gelaten zijn leerlingen van scholen in het buitenland, het speciaal onderwijs en het voortgezet onderwijs. In de dataset zijn deze onder andere te herkennen aan het ontbreken van een BRIN-code (buitenland), het

ontbreken van gegevens over de verdeling van de formatie binnen een school (kenmerkend voor speciaal onderwijs en voortgezet onderwijs), of aan het feit dat in de database al specifiek is aangegeven dat het geen school uit het basisonderwijs betreft. Tevens zijn de leerlingen die de digitale versie van de Eindtoets hebben gemaakt buiten beschouwing gelaten, omdat deze leerlingen de toets op een ander moment gemaakt hebben. Daarnaast zijn de scholen afgezet tegen de lijst met scholen voor primair onderwijs van DUO. Als wel de hoofdvestiging (00) in de lijst staat, maar niet de specifieke vestiging, is de school alsnog opgenomen (dit betreft een zeer beperkt aantal scholen). Dit waren ook in eerdere onderzoeken van JPON/JMTR de exclusiecriteria. Van de 161.192 leerlingen die een van de eindtoetsen gemaakt heeft vallen hierdoor in totaal 9679 leerlingen uit de dataset omdat zij niet tot de onderzoekspopulatie behoren. Als we alleen de scholen willen opnemen die ook in 2013 in de data voorkwamen dan vallen er nog eens 3199 extra leerlingen af. We willen ons vooral richten op een vergelijkbare dataset, om een goede jaarvergelijking te maken. Daarom streven we naar zoveel mogelijk overlap met de dataset van 2013. De basis voor de analyses is dientengevolge een dataset met 148.314 leerlingen. Ter controle zijn overigens ook analyses uitgevoerd met de iets grotere dataset met 151.513 leerlingen waarin de 3199 leerlingen ook weer zijn opgenomen. Het percentage uitval van leerlingen is bij beide datasets vergelijkbaar met de percentages in 2013. De uitval bij de kleinere dataset was 1% groter en bij de grotere dataset 1% kleiner dan in 2013.

Uit het databestand met 148.314 leerlingen zijn twee typen steekproeven getrokken: één ten behoeve van de schaalconstructie en één voor de analyse met achtergrondvariabelen. Het eerste bestand is de "kalibratiesteekproef" die gebruikt wordt om de opgaven van verschillende jaren op dezelfde schaal te kunnen plaatsen en de toetsresultaten daarmee over de jaren heen vergelijkbaar te maken. Voor deze schaalanalyses is gebruikgemaakt van de resultaten van alle leerlingen die beschikbaar zijn voor de optimale koppeling tussen jaargangen. Bij het schalen van de opgaven is de representativiteit van de steekproef van minder groot belang, vanwege populatieonafhankelijkheid van de itemparameterschattingen bij de gebruikte analysemethode (item response theorie, zie verder paragraaf 2.3). In deze steekproef hebben we alle leerlingen opgenomen die de ankertoets (EA) of de toets Basisvaardigheden (TBV), samen met de Eindtoets, hebben gemaakt. Daarnaast zijn er steekproeven genomen van leerlingen die de reguliere eindtoets (EB) gemaakt hebben (10%) en die de Eindtoets Niveau (EN) gemaakt hebben (50%). De OPLM-steekproef is zodoende als volgt opgebouwd: 12.399 leerlingen die alleen de EB gemaakt en 461 die daarbij ook de TBV gemaakt hebben; 3.831 leerlingen die de EA en 11.617 leerlingen die de EN hebben gemaakt. De EN-leerlingen werden ingesloten omdat zij een andere Eindtoets hebben gemaakt. Door nu de groepen EN en EB ongeveer even groot te maken is het aantal observaties per opgave optimaal. De ankertoets en de TBV werden ingesloten om de toetsen op onderdelen te kunnen koppelen.

De kwaliteit van de gebruikte schalen is niet anders dan die in voorgaande jaren. Met de gebruikte aantallen kan iedere kleine afwijking in veel gevallen als significant gekwalificeerd worden, maar deze afwijking heeft geen effect heeft op de resultaten (zie voor nadere toelichting de vorige JPON-rapportages).

Na de schaalanalyse is met de geconstrueerde schalen, die in principe dezelfde psychometrische kenmerken bezaten als in voorgaande jaren, de analyse gedaan op achtergrondvariabelen, met in de eerste plaats de jaarvergelijking. Bij deze analyse is de representativiteit wel van groot belang. Daartoe is een nieuwe steekproefprocedure uitgevoerd, zodanig dat van ieder van de vier onderscheiden groepen (EB, EA, EN en EB+TBV) 20% van de leerlingen in de steekproef van ongeveer 30.000 leerlingen terecht komt. Deze procedure is vijf keer, zonder teruglegging, herhaald. De resultaten in dit rapport zijn gebaseerd op de meest representatieve steekproef van de vijf. Om de robuustheid van deze steekproef te bepalen zijn SAUL-analyses op drie verschillende steekproeven uitgevoerd. Er bleek weinig steekproeffluctuatie te zijn. De kenmerken van deze groepen op achtergrondvariabelen zijn volkomen vergelijkbaar met de totale groep en die zijn terug te vinden in de zogenaamde *Terugblik bij de Eindtoets* (Cito, 2014). We mogen dus stellen dat deze steekproeven een representatief beeld geven van de Nederlandse populatie in jaargroep 8. De verdeling van de achtergrondvariabelen is gegeven in **bijlage 4.**



De analyses van de effecten van de achtergrondvariabelen zijn gedaan op de vaardigheidsscores, zodat de gerapporteerde resultaten vergelijkbaar zijn met die van de eerdere rapportages. Naast de geschatte gemiddelden (en andere verdelingseigenschappen) van de latente variabelen is er bij de analyse gebruikgemaakt van het programma SAUL (*Structural Analysis of a Univariate Latent Variable*; Verhelst en Verstralen, 2002). Hiermee zijn de marginale effecten en de gecorrigeerde effecten berekend.

2.2.3 Achtergrondvariabelen

Behalve de genoemde taal- en rekenvaardigheden zijn ook een aantal achtergrondkenmerken in het onderzoek betrokken. Deze zijn te onderscheiden in leerlingkenmerken en schoolkenmerken.

Leerlingkenmerken:

- Geslacht: jongens-meisjes;
- Leertijd: leerlingen die al of niet eens hebben gedoubleerd, resp. regulier en vertraagd;
- Formatiegewicht: gewicht van de leerling voor de formatieregeling op grond van opleiding en herkomst (alleen bij de oude regeling) van de ouders; verdeeld in drie categorieën: geen gewicht (0.00), laag gewicht (0.30) en hoog gewicht (1.20);
- Thuis taal algemeen: drie niveaus (1) Leerling spreekt thuis alleen Nederlands, (2) Leerling spreekt naast Nederlands ook een andere taal, of (3) Leerling spreekt thuis geen Nederlands, maar alleen een andere taal;
- Thuis taal detail: alleen Nederlands gesproken versus verschillende talen (al dan niet in combinatie met Nederlands);
- Advies VO: ingeschat niveau vervolgonderwijs door leerkracht (vijf hoofdniveaus: vmbo-basisberoeps, vmbo-kaderberoeps, vmbo-theoretisch (of gemengd), havo, en vwo; en de vier tussenliggende combinatieniveaus);
- Wel of niet een IJK-code: de code I staat voor (Allochtone) leerlingen die aan het begin van jaargroep 8 vier jaar of korter in Nederland zijn en die het Nederlands onvoldoende beheersen om de opgaven in de Eindtoets goed te kunnen lezen; code J staat voor leerlingen die naar verwachting naar het (voortgezet) speciaal onderwijs of naar het praktijkonderwijs (pro) gaan; en code K staat voor leerlingen die naar verwachting in aanmerking komen voor het leerwegondersteunend onderwijs (lwoo).
- Gemaakte toets: Eindtoets Basis of Eindtoets Niveau (voorheen Niveautoets).

Schoolkenmerken:

- Stratum: schoolindeling op basis van het percentage leerlingen met een formatiegewicht, verdeeld in drie categorieën; 1 = weinig (<10%), 2 = matig (10-25%) en 3 = veel (>25%);
- Regio van de locatie van de school: noord, oost, west, zuid (indeling CBS);
- Urbanisatiegraad van de locatie van de school (vijf groepen; van zeer sterk stedelijk tot en met niet stedelijk);
- Schoolgrootte (200 of minder leerlingen versus meer dan 200 leerlingen).

In de loop van de tijd, tussen 2008 en 2014, zijn de definities van een aantal variabelen veranderd, zoals de definitie van het formatiegewicht en van stratum. De definitie die tegenwoordig wordt gehanteerd kan gevonden worden in de 'Brochure nieuwe gewichtenregeling basisonderwijs' uit april 2008, van het CFI (later DUO; zie ook www.duo.nl). Deze definities zijn stapsgewijs toegepast in het onderwijs.

Vereenvoudigd weergegeven ziet de regeling er als volgt uit. Bij de opleidingen van de ouders wordt onderscheid gemaakt tussen een drietal categorieën:

Categorie 1: > de ouder heeft maximaal (speciaal) basisonderwijs gehad.

Categorie 2: > de ouder heeft maximaal lbo/vbo, praktijkonderwijs, vmbo basis- of kaderberoepsgerichte leerweg gedaan, of de ouder heeft maximaal twee leerjaren van een andere opleiding in het VO gedaan.

Categorie 3: > de ouder heeft meer dan twee jaar mavo, havo, vwo dan wel vmbo gemengde of theoretische leerweg gevolgd.

Op basis van deze categorieën worden de volgende formatiegewichten toegekend:

- 0.00 (geen gewicht) voor leerlingen met (minstens) één ouder uit categorie 3;
- 0.30 voor leerlingen die niet gewicht 0.00 en niet gewicht 1.20 hebben;
- 1.20 voor leerlingen die geen ouder hebben in categorie 3 en (minstens) één uit categorie 1.

Opgemerkt moet worden dat de wijziging in definitie van formatiegewichten in de loop van de in dit onderzoek bestreken periode ook een rol speelt bij de vergelijking over de jaren. Tevens heeft deze wijziging gevolgen gehad voor de indeling in de strata van de scholen. Hoe daarmee is omgegaan is uitvoerig toegelicht in de rapportage van het peilingsjaar 2010 (Hemker, Kordes & Van Weerden, 2011). Het is ook een van de redenen waarom gecorrigeerde gegevens een verkeerd beeld zouden kunnen geven.

De achtergrondvariabelen voor de leerlingen worden bij de Eindtoets verzameld via het antwoordblad dat bij deze toets hoort en waarop leerlingen en leerkrachten een aantal gegevens invoeren. In de jaren 2012 en 2013 is de achtergrondvariabele thuistaal niet opgevraagd bij de afname van de Eindtoets. Dat was in 2014 wel weer het geval. De variabele IJK is in de jaren 2012 en 2013 meegenomen in de analyses als alternatief voor de variabele thuistaal. In 2014 is thuistaal wel weer opgevraagd, maar voor de vergelijkbaarheid over de afgelopen jaren is ook de variabele IJK weer meegenomen in de analyses.

2.3 Hoe is er geanalyseerd?

Hieronder geven we beknopt weer hoe we de resultaten van 2014 hebben vergeleken met die van 2013 en de jaren daarvoor. Voor een uitgebreide technische verantwoording van de steekproef, de gebruikte toetsen, de statistische analyse en de rapportagemethodiek wordt verwezen naar de technische rapportages van eerdere jaren (Hemker & Van Weerden, 2009; Hemker, Kuhlemeier en Van Weerden, 2010; Hemker, Kordes & Van Weerden, 2011; Van Weerden, Hemker, Straat & Mulder, 2012). We gaan hier vooral in op de afwijkingen in methodiek die er dit peilingsjaar zijn toegepast.

2.3.1 Vergelijking tussen en over de jaren

Een voorwaarde om vaardigheden van de verschillende jaren direct met elkaar te kunnen vergelijken, is dat de vaardigheden op dezelfde schaal gemeten zijn. Aan deze voorwaarde is voldaan als de leerlingen in de steekproef van 2014 precies dezelfde toetsen hebben gemaakt als de leerlingen in eerdere steekproeven. De Eindtoets Basisonderwijs, die hier wordt gebruikt voor de meting van jaargroep 8, wordt echter ieder jaar volledig vernieuwd. Dat zou een directe vergelijking onmogelijk maken. Wij hebben de vergelijkbaarheid echter kunnen realiseren doordat een aantal achtstegroepers telkens een aantal extra opgaven krijgen voorgelegd via zogenaamde ankertoetsen, die wel over de jaren heen ongewijzigd zijn gebleven. Ter aanvulling op die gegevens kon ook gebruik worden gemaakt van een overlap van items uit de Niveautoets met items uit de toets Basisvaardigheden. Deze laatste toets werd voor het eerst in 2011 op een groot aantal scholen in jaargroep 8 afgenomen en is in 2012, 2013 en 2014 opnieuw gebruikt. Wel is er een duidelijke afname in het aantal scholen dat deze toets gebruikt.

Met behulp van een speciale analysetechniek met bijbehorend programma: het One Parameter Logistic Model (OPLM; Verhelst, 1993; Verhelst en Glas, 1995), een variant van het item response model, zijn schalen geconstrueerd voor alle afzonderlijke onderwerpen. Dit is reeds in 2008 gedaan, waarbij de toetsen allen op dezelfde schaal werden gezet, met een gemiddelde van 250 punten en een standaardafwijking van 50. Dit is vergelijkbaar met de aanpak van PPON. Op deze wijze kunnen de prestaties op de verschillende toetsen over de jaren heen met elkaar worden vergeleken (vgl. Hemker, Kordes & Van Weerden, 2011).

Voor het maken van de schalen is het geen probleem, maar voor het vergelijken van prestaties in jaargroep 8 over de jaren heen moet ook rekening gehouden worden met de samenstelling van de responsgroep. Als de samenstelling gelijk is, dan is een directe vergelijking mogelijk, maar als er verschillen in samenstelling bestaan, moet onderzocht worden waar die wijziging vandaan zou kunnen komen.

Veranderingen in de samenstelling van een responsgroep kunnen het gevolg zijn van zogenaamde steekproeffluctuaties, maar ook van 'echte' veranderingen in de samenstelling van de populatie.

Het probleem van steekproeffluctuaties zal zich voor jaargroep 8 niet zo gauw voordoen. Er is namelijk in elk jaar op vergelijkbare wijze een zeer grote aselechte steekproef uit een bestand van vele honderdduizenden leerlingen getrokken (telkens ongeveer 85% van de populatie). Wel kan er zich een wijziging in de samenstelling van de populatie hebben voorgedaan. Het gevolg zou kunnen zijn dat op basis van de analyses die daar geen rekening mee houden er onjuiste conclusies getrokken zouden kunnen worden. Aan de andere kant zou het onterecht corrigeren van de gegevens ook tot onjuiste conclusies kunnen leiden, bijvoorbeeld als de eigenschappen van de groepen door de tijd veranderen zonder dat het label van die groep verandert. Een ander voorbeeld is als er relatief veel meer vwo-leerlingen zouden zijn, en daarvoor gecorrigeerd zou worden: de vooruitgang zou dan juist zitten in de toename van de vwo-leerlingen, terwijl de toename in vaardigheid dan "weg gecorrigeerd" zou worden. Om die reden is ervoor gekozen om voor de jaarvergelijkingen de ongecorrigeerde resultaten te gebruiken. Bij de conclusies moet dan wel gekeken worden of er zich grote wijzigingen voordoen in de achtergrondvariabelen, die wellicht een nieuwe kijk geven op de gevonden resultaten.

De berekeningen worden uitgevoerd op de latente trek met behulp een speciaal programma voor regressieanalyse van vaardigheidsscores geschat met behulp van OPLM, te weten SAUL: Structural Analysis of Univariate Latent variabels (Verhelst & Verstralen, 2002).

2.3.2 Analyses voor de leerlingen in 2014

Voor het vergelijken van de resultaten over de jaren heen wordt uiteraard gebruikgemaakt van data waarin alle jaren vertegenwoordigd zijn. Voor de analyses van de effecten van de achtergrondgegevens in 2014 wordt alleen gebruikgemaakt van gegevens uit 2014.

Voor de variabelen die in 2014 zijn geanalyseerd, zijn de gegevens van de gemiddelden van de te onderscheiden groepen vergeleken op de vaardigheidsschalen. Deze zijn direct vergelijkbaar met de ongecorrigeerde effecten. Doordat de vaardigheden in 2014 op dezelfde schaal liggen als in de voorgaande jaren zijn de resultaten hiermee direct vergelijkbaar over de jaren heen. De gepresenteerde effectschattingen betreffen voor de variabelen echter die van het gecorrigeerde model, ook omwille van de vergelijkbaarheid met de vorige jaren. Het gaat dan om de effecten binnen het model waarin de variabelen 'stratum', 'geslacht', 'leertijd' en 'formatiegewicht' altijd zijn opgenomen (het basismodel). In aanvullende analyses zijn de overige variabelen, zoals regio en thuistaal, telkens als additionele variabele opgenomen, wat verschillende modellen met vijf variabelen oplevert.

Door gebruik te maken van deze gecorrigeerde effecten is het mogelijk een inschatting te maken van het additieve effect van een variabele, zoals de gesproken thuistaal, naast het effect dat al gevonden wordt op basis van de vooropleiding van de ouders, zoals gerepresenteerd in het formatiegewicht.

In de eerdere verslagen wordt dieper ingegaan op verschillen tussen de gecorrigeerde modellen en de niet gecorrigeerde modellen en de verschillen in interpretatie. Ook de gevolgen voor de vergelijking over de jaren heen en binnen het analysejaar zijn daar uiteengezet.

2.3.3 Significantie en effectgrootte

Of een gemiddeld vaardigheidsverschil tussen twee jaren statistische significantie oplevert, hangt in belangrijke mate af van de steekproefgrootte. Hoe groter de steekproef, hoe eerder een verschil statistisch significant is. De steekproef voor jaargroep 8 is dusdanig groot dat verschillen tussen groepen al heel snel significant zijn. Om de resultaten toch zinvol met elkaar te kunnen vergelijken rapporteren wij behalve de statistische significantie ook de zogeheten effectgrootte. De effectgrootte wordt in ons geval berekend als het verschil tussen de gemiddelden van de twee jaren (of twee subgroepen) gedeeld door de (gepoolde)

standaardafwijking van de twee groepen die onderling worden vergeleken. Bij de interpretatie van de effectgrootte hanteren we de vuistregel van Cohen (1988) die is afgebeeld in tabel 2.1. Alles met een effectgrootte boven de 0,20 noemen we hier betekenisvol.

Tabel 2.1 *Kwalificatie van effectgrootten*

Effectgrootte (zowel plus als min) Kwalificatie

0,0 tot 0,2	geen effect
0,2 tot 0,5	klein effect
0,5 tot 0,8	matig effect
0,8 of groter	groot effect

2.3.4 De gekozen rapportageschaal

Elke vaardigheid in dit onderzoek is getransformeerd naar een schaal met een gemiddelde van 250 en een standaarddeviatie van 50. Dit is conform de werkwijze bij PPO (zie bijv. Janssen, Van der Schoot & Hemker, 2005). De startwaarde is voor elke schaal het gemiddelde dat we in 2008 hebben aangetroffen. Dat gemiddelde is op 250 gesteld (zie verder Hemker & Van Weerden, 2009). De transformatie heeft als voordeel dat we de prestaties voor verschillende vaardigheden en voor verschillende jaren naast elkaar kunnen zetten op een en dezelfde schaal. Om de resultaten goed te kunnen beschrijven richten we ons in de rapportage niet alleen op de gemiddelde leerling, maar ook op andere groepen leerlingen in de vaardigheidsverdeling. Voor dit doel zijn vijf typische leerlingen gedefinieerd, gekoppeld aan kenmerkende percentielpunten (zie tabel 2.2). In de rapportage gaan we uit van het basisjaar en rapporteren in de jaren daarop de eventuele wijziging in percentielwaarde. Zo zal bij een eventuele vooruitgang in vaardigheid blijken dat meer leerlingen de schaalwaarde 250 bereiken, maar ook kunnen meer leerlingen de schaalwaarde 186 bereiken. Teruggerekend betekent dat een wijziging in het percentage leerlingen dat tot een groep typische leerlingen behoort ten opzichte van 2008.

Toen in 2008 de Jaarlijkse Meting Taal & Rekenen (voorheen JPON) van start ging, werd al gedacht aan het meenemen van de referentieniveaus. Deze waren echter nog niet gedefinieerd. De Eindtoetsen van 2008 konden dus inhoudelijk nog niet afgestemd worden op de referentieniveaus. Ook waren in 2008 de officiële grenswaarden voor 1F of 1S nog niet vastgesteld, waardoor ook nog geen standaard voor niveau 1F en 1S was vastgelegd.

In 2008 is er daarom voor gekozen verschillende punten op de vaardigheidsschaal als ijkpunten te nemen, te weten de percentielpunten P10, P25, P50, P75 en P90. Zie voor de aanduiding en de schaalwaarden van de deze ijkpunten tabel 2.2.

Tabel 2.2 *Definiëring typische leerlingen in de vaardigheidsverdeling*

Aanduiding	Afkorting	Percentiel 2008	% kandidaten 2008 van minstens dit niveau	Schaalwaarde
Zeër laag vaardig	ZLV	P10	90%	186
Laag vaardig	LV	P25	75%	216
Gemiddelde/standaard	G/S	P50	50%	250
Hoog vaardig	HV	P75	25%	284
Zeër hoog vaardig	ZHV	P90	10%	314

Aan de schaalwaarde horende bij de laag vaardige leerling kan een speciale betekenis kan worden toegekend. Dit is het punt op de vaardigheidsschaal dat door minstens 75% van de leerlingen wordt bereikt. In het advies van de Expertgroep Doorlopende Leerlijnen Rekenen en Taal (SLO, 2008) wordt dit punt regelmatig genoemd als referentieniveau voor 1F, zowel bij taal als bij rekenen. Het referentieniveau 1S zou dan voor rekenen overeenkomen met P50. Bij taal is P75 het uitgangspunt. Gekoppeld aan de specifieke inhoud van de schalen die we hier rapporteren zijn er echter nog geen uitspraken gedaan over het gewenste niveau. Opgemerkt moet worden dat deze grenspunten P50 en P75 niet de definitieve niveaus 1F en 1S zijn, maar (grove) voorlopige inschattingen. Deze punten op de schaal zijn door middel van een populatiereferentie (2008) gemaakt, terwijl de echte niveaus 1F en 1S op basis van inhoudelijke kenmerken worden vastgesteld.

In 2014 was er nog geen inhoudelijke koppeling tussen de Eindtoets en de inhoudelijke kenmerken van de referentieniveaus. Dat betekent dat ook in dit verslag over 2014 er ook nog geen exacte koppeling gemaakt kan worden tussen prestaties op de Eindtoets van 2014 enerzijds en de referentieniveaus anderzijds. Er is weliswaar meer bekend over de definitieve referentieniveaus maar deze zijn nog niet inhoudelijk gekoppeld aan de opgaven en gegevens van de Eindtoets. Dat zal met ingang van 2015 wel mogelijk moeten worden, als voor de eerste keer de Centrale Eindtoets wordt afgenomen.

3 De resultaten voor jaargroep 8

3.1 De vergelijking over de jaren

Ervan uitgaande dat steekproeffluctuaties geen rol spelen (zie paragraaf 2.2), vergelijken we in deze paragraaf de resultaten van jaargroep 8 in 2014 met voorgaande jaren vanaf 2008. We presenteren hier dus de ongecorrigeerde verschillen.

In tabel 3.1 zijn de uitkomsten voor de zeven opeenvolgende jaren weergegeven als gemiddelde op de vaste schaal. Het gemiddelde voor 2008 is per definitie 250 en de standaarddeviatie is 50. De schaalwaarde van 250 is vastgelegd in 2008 en vormt het startpunt waarmee we de uitkomsten van 2009 en later vergelijken (zie verder Hemker & Van Weerden, 2009).

In tabel 3.2 geven we weer hoe groot de verschillen van jaar tot jaar zijn in termen van effectgrootte. Bovendien geven we de effectgrootte voor het contrast 2014-2008, dat wil zeggen het verschil in uitkomst van de meting in 2014 met die van de eerste meting in 2008, dus over de afgelopen 6 jaar. Als het verschil tussen 2014 en 2013 statistisch significant is, is het effect vetgedrukt. Evenzo geldt dat voor het verschil tussen 2014 en 2008.

Vergelijken we de uitkomsten voor 2014 met 2013 dan zien we dat de schaalwaarden over de hele linie hetzelfde zijn of een punt lager. Alleen voor Woordenschat is het verschil significant ($\alpha = 0,01$). Deze uitkomst betekent dat alleen bij Woordenschat sprake is van een significante teruggang in resultaat vergeleken met 2013, al valt daarbij de kanttekening te maken dat dit resultaat niet significant is bij een α van 0,001. Bij de andere vaardigheden is geen wijziging te constateren, noch in positieve, noch in negatieve zin.

Vergelijken we 2014 met 2008 dan zien we dat voor alle vaardigheden nog wel steeds hetzelfde geldt als wat we in 2013 al constateerden, namelijk dat het in het laatste jaar bereikte gemiddelde hoger is dan in 2008, het uitgangsjaar. In tabel 3.2 zien we tevens dat die positieve verschillen allemaal significant zijn. De effectgrootte varieert van 0,04 tot 0,08 en is het grootst bij twee rekenonderwerpen, respectievelijk Breuken, procenten en verhoudingen en Meten, meetkunde, tijd en geld.

Tabel 3.1 Gemiddelde schaalwaarden jaargroep 8*

Vaardigheden	2008	2009	2010	2011	2012	2013	2014
T: Woordenschat	250	249	257	251	252	254	253
T: Spelling	250	250	252	252	256	254	254
T: Begrijpend lezen (anker)	250	252	254	257	253	253	252
RW: Rekenen: totaalschaal	250	250	253	253	256	253	253
RW: Getallen en bewerkingen	250	249	252	252	255	252	252
RW: Breuken, procenten en verhoudingen	250	250	254	254	257	254	254
RW: Meten, meetkunde, tijd en geld	250	250	254	254	256	255	254

* Standaarddeviaties variëren van 49 tot 53; in 2008 per definitie 50.

* Merk op dat er kleine verschillen met eerder gerapporteerde tabellen kunnen ontstaan. Dit is te wijten aan de herberekeningen die effect kunnen hebben op afrondingen. Deze hebben geen effect op de eerder getrokken conclusies. De verschillen in de controle data waarbij meer (mogelijk speciaal onderwijs) scholen zaten, zijn o klein dat deze niet tot een andere duiding van de Tabel 3.1 zouden leiden.

Er is gecontroleerd of de selectie van leerlingen de waarvan niet geheel zeker was of deze tot de onderzoekspopulatie behoorden (zie paragraaf 2.2.2) enig effect op de resultaten had. Daarbij is de passing van het model op de kleinere dataset beter. De resultaten op de grotere dataset, waarbij mogelijk scholen uit het speciaal onderwijs opgenomen, zijn iets slechter (niet weergegeven in een tabel), maar de verschillen zijn zeer klein.

Tabel 3.2 Effectgroottes van de jaarverschillen jaargroep 8*

Vaardigheden	2009-2008	2010-2009	2011-2010	2012-2011	2013-2012	2014-2013	2014-2008
T: Woordenschat	-0,02	0,15	-0,11	0,02	0,04	-0,03	0,06
T: Spelling	0,00	0,03	0,00	0,07	-0,03	0,00	0,06
T: Begrijpend lezen (anker)	0,04	0,04	0,07	-0,09	0,00	-0,01	0,05
RW: Rekenen: totaalschaal	-0,01	0,07	0,00	0,06	-0,06	0,00	0,06
RW: Getallen en bewerkingen	-0,01	0,05	0,00	0,07	-0,06	-0,01	0,04
RW: Breuken, procenten en verhoudingen	-0,01	0,08	0,01	0,06	-0,07	0,01	0,08
RW: Meten, meetkunde, tijd en geld	0,01	0,07	0,00	0,05	-0,04	-0,01	0,08

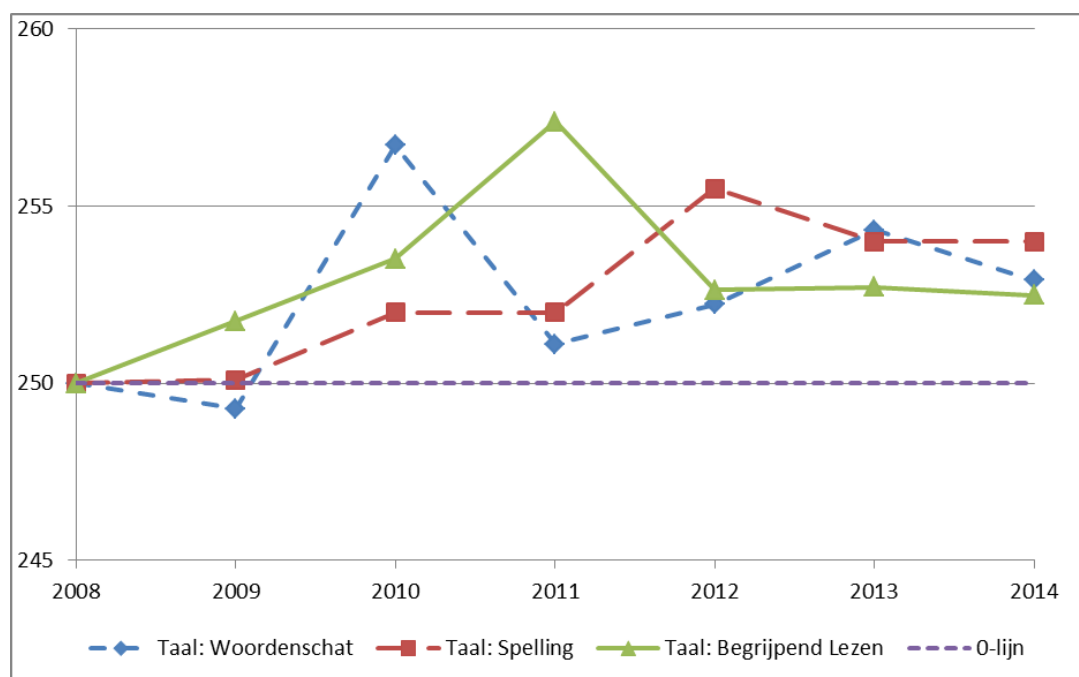
* Vet = significant verschil met voorgaand jaar of, in laatste kolom, met 2008 ($\alpha < 0,01$).

3.1.1 Taalvaardigheden

De ontwikkeling in schaalscores is afzonderlijk in beeld gebracht voor de drie verschillende taalvaardigheden (figuur 3.1). We zien dat, hoewel de trendlijnen van elkaar verschillen, de uiteindelijke groei van 2008 naar 2014 vergelijkbaar is. Bij de vergelijking van 2014 met 2013 zien we weinig verschuivingen, zeker in vergelijking met eerdere jaren. Alleen bij Woordenschat is er sprake van een licht dalende trend sinds 2013.

Het niet-lineaire verloop van de trend in de resultaten bij Woordenschat is al nader besproken in de rapportage van 2011, waarbij gewezen is op het verschil tussen jongens en meisjes bij de uitkomsten. De trendlijn bij Begrijpend lezen is gebaseerd op de resultaten zoals gevonden met de ankertoets. Dat is ook de wijze waarop de eerdere jaarvergelijkingen gedaan zijn (zie Van Weerden e.a., 2013). De vergelijking met de toets Basisvaardigheden, zoals in de rapportage van 2012 verkend is, was ook mogelijk geweest, maar bleek lastiger te interpreteren. In 2013 is deze daarom niet meer gerapporteerd. Wat vooral opvalt in de resultaten van 2014 ten opzichte van de vorige jaren is een zekere convergentie van de drie onderwerpen en een recente stabilisatie in resultaat. Het niveau ligt daarbij wel een aantal punten hoger dan in 2008.

Figuur 3.1 Trends over de jaren voor de taalonderwerpen in jaargroep 8



Tabel 3.3 Vergelijking over de jaren heen in percentage per leerlingengroep voor de taalonderwerpen in jaargroep 8

Leerlingengroep*	Percentiel-grens	Woordenschat: % leerlingen						
		2008	2009	2010	2011	2012	2013	2014
> dan ZLV	P10	90	90	92	90	91	91	91
> LV	P25	75	74	79	76	76	78	77
> Standaard 2008	P50	50	49	55	51	52	53	52
> HV	P75	25	24	29	26	26	28	27
> ZHV	P90	10	10	13	10	11	12	11

Leerlingengroep*	Percentiel-grens	Spelling: % leerlingen						
		2008	2009	2010	2011	2012	2013	2014
> dan ZLV	P10	90	90	91	91	92	91	91
> LV	P25	75	75	76	76	78	77	77
> Standaard 2008	P50	50	50	52	52	54	53	53
> HV	P75	25	25	26	26	29	28	28
> ZHV	P90	10	10	11	11	12	11	11

Leerlingengroep*	Percentiel-grens	Begrijpend lezen: % leerlingen						
		2008	2009	2010	2011	2012	2013	2014
> dan ZLV	P10	90	91	91	92	91	91	91
> LV	P25	75	76	77	79	77	77	77
> Standaard 2008	P50	50	52	53	56	52	52	52
> HV	P75	25	26	27	30	27	27	27
> ZHV	P90	10	11	11	13	11	11	11

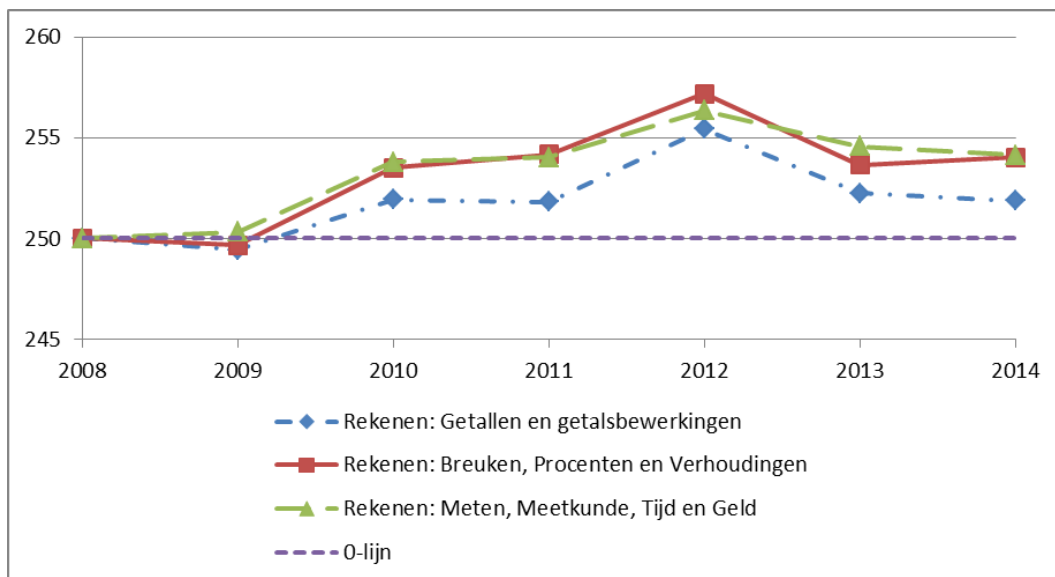
* Zie tabel 2.2; % leerlingen: percentage leerlingen minstens op niveau leerlinggroep

In tabel 3.3 is weergegeven welk percentage van de leerlingen een bepaalde percentielscore behaalt, waarbij de startwaarde gebaseerd is op de situatie van 2008. Zo was het percentage leerlingen dat percentielwaarde 10 behaald, de zeer laag vaardige leerling, in 2008 per definitie 90%. Het gaat dan concreet om het percentage leerlingen dat een schaalwaarde van 186 of hoger heeft op de schaal van 250 (sd=50), zie ook tabel 2.2. In 2014 blijkt dat percentage nu bij alle vaardigheden op 91% uit te komen, hetgeen betekent dat iets meer leerlingen dat niveau in 2014 bereiken dan in 2008. Het percentage leerlingen dat in 2014 in percentiel 25 (van 2008) valt, is bij alle vaardigheden 77%, terwijl dat in 2008 (per definitie) 75% was. Bij de gemiddelde leerling, de p50-leerling, varieert dat in 2014 van 52% tot 53%.

3.1.2 Rekenen

Voor Rekenen is het mogelijk om zowel een totaalschaal te presenteren als drie afzonderlijke schalen. De onderlinge samenhang is dermate groot dat hier statistisch geen enkel bezwaar tegen is. Voor de volledigheid presenteren we hier de afzonderlijke schalen. De sterke onderlinge samenhang van de drie onderwerpen zien we in de grafische weergave terug (figuur 3.2). Over de hele linie zien we ongeveer hetzelfde beeld voor alle drie de schalen: in 2010 gaan alle schalen enkele punten omhoog, in 2011 is er nauwelijks verschil, maar in 2012 is er weer een duidelijke vooruitgang. In 2013 gaan alle schalen weer naar beneden en in 2014 blijven ze op ongeveer hetzelfde niveau. Over de hele linie bekeken is de uitkomst in 2014 evident lager dan in 2012. Maar ten opzichte van 2008 blijft er nog steeds sprake van een positief verschil, waarbij Getallen en bewerkingen relatief het laagst scoort van de drie onderwerpen. Alleen bij Meten, meetkunde, tijd en geld zien we een klein verschil en wel in negatieve zin. Net als bij de taalonderwerpen lijkt er sprake van een zekere stabilisering op een wat hoger niveau dan in 2008, maar lager dan in tussenliggende jaren.

Figuur 3.2 Trends over de jaren voor de rekenonderwerpen in jaargroep 8



Tabel 3.4 Vergelijking over de jaren heen in percentage per leerlingengroep voor de rekenonderwerpen in jaargroep 8

Leerlingengroep*	Percentielgrens	Getallen en bewerkingen: % leerlingen						
		2008	2009	2010	2011	2012	2013	2014
> dan ZLV	P10	90	90	91	91	92	91	91
> LV	P25	75	74	76	76	78	76	76
> Standaard 2008	P50	50	49	52	51	54	52	52
> HV	P75	25	24	26	26	29	26	26
> ZHV	P90	10	10	11	11	12	11	11

Leerlingengroep*	Percentielgrens	Breuken, procenten en verhoudingen: % leerlingen						
		2008	2009	2010	2011	2012	2013	2014
> dan ZLV	P10	90	90	91	91	92	91	91
> LV	P25	75	75	77	78	79	77	77
> Standaard 2008	P50	50	50	53	53	56	53	53
> HV	P75	25	25	27	28	30	27	28
> ZHV	P90	10	10	11	12	13	11	11

Leerlingengroep*	Percentielgrens	Meten, meetkunde, tijd en geld: % leerlingen						
		2008	2009	2010	2011	2012	2013	2014
> dan ZLV	P10	90	90	91	91	92	92	91
> LV	P25	75	75	77	77	79	78	78
> Standaard 2008	P50	50	50	53	53	55	54	53
> HV	P75	25	25	27	28	29	28	28
> ZHV	P90	10	10	11	11	12	12	12

* Zie tabel 2.2; % leerlingen: percentage leerlingen minstens op niveau leerlingengroep

Ook voor Rekenen zijn de percentages leerlingen in kaart gebracht voor de verschillende percentielwaarden: p10, p25, p50, p75 en p90. Dit is weergegeven in tabel 3.4. Ten opzichte van 2013 is er echter weinig veranderd. De toename in 2014 ten opzichte van 2008 varieert van 1 tot 3 procent. Het percentage zeer laag vaardige leerlingen (ZLV) is met 1 procent afgenomen: waar eerst 90% beter presteerde dan de leerlingen op het laagste niveau, is dat nu 91%. Eveneens zien we dat de best presterende groep iets groter is geworden. Ook zien we dat in 2014 een hoger percentage leerlingen boven het niveau van de typische standaardleerling van 2008 zit.

3.2 Verschillen voor achtergrondvariabelen

In deze paragraaf rapporteren we de verschillen tussen groepen leerlingen. De verschillende groepen zijn gebaseerd op de categorisering naar achtergrondvariabele. We presenteren hier in de eerste plaats de gevonden effecten in 2014 en gaan bij enkele variabelen in op de trend over zes jaar.

Anders dan in 2013 en 2012 zijn er in 2014 weer gegevens beschikbaar met betrekking tot de variabele thuistaal. De variabelen "code I" en "code IJK" zijn daardoor minder belangrijk geworden.

Dit jaar gaat er speciale aandacht uit naar de resultaten bij de variabele leertijd. Binnen deze variabele zijn twee groepen onderscheiden: vertraagde leerlingen en niet vertraagde (reguliere) leerlingen. Vertraagde leerlingen zijn leerlingen die gezien hun leeftijd al van de basisschool af hadden kunnen zijn.

Deze leerlingen hebben dus een verlengde leertijd gehad in het primair onderwijs. De groep niet vertraagde leerlingen betreft leerlingen waarvan op basis van hun geboortedatum verwacht wordt dat ze in groep 8 zitten en leerlingen die een jaar (of meer) hebben overgeslagen. De extra aandacht die we aan deze variabele besteden wordt niet zozeer ingegeven door de onverwachte verschillen in resultaat bij deze groepen, maar komt voort uit de waarneming dat het aantal vertraagde leerlingen in de loop der tijd gestaag blijkt af te nemen. Welke effect dat kan hebben op interpretatie van de resultaten van de jaarvergelijking bespreken we onder het kopje leertijd.

We onderscheiden in de rapportage variabelen op leerlingniveau, die leerlingen categoriseren los van de school waar ze op zitten en variabelen op schoolniveau, waarvoor geldt dat alle leerlingen van dezelfde school dezelfde categorisering hebben.

Het gebruikte basismodel voor de analyses van 2014 bevatte de verklarende variabelen geslacht, leertijd, formatiegewicht en stratum. Deze laatste variabele definiëren we als het percentage gewichtenleerlingen op een school, onderscheiden naar drie categorieën. Ieder van de geschatte effecten per variabele wordt gecorrigeerd voor effecten van de overige variabelen. Hiermee kan het unieke effect van de variabele bepaald worden. Aan dit basismodel zijn in de verdere analyses variabelen toegevoegd. Dit betreft de variabelen toetsvariant (Eindtoets of Niveautoets), advies VO, schoolgrootte, regio en verstedelijking. Dit levert een vijftal extra modellen op, waarbij de basisvariabelen telkens aangevuld worden met één extra variabele. De gerapporteerde effecten zijn telkens gecorrigeerde effecten. Voor een weergave van de gemiddelden, standaarddeviaties, overschrijdingskansen en effectgroottes per achtergrondvariabele verwijzen we naar de tabellen in de bijlagen 1 tot en met 4.

In tabel 3.5 zijn de effectgroottes voor de variabelen op leerlingniveau weergegeven voor geslacht, leertijd formatiegewicht, advies VO en type toets dat de leerling gemaakt heeft. In deze tabel worden drie kleuren groen gebruikt: de meest lichte kleur groen geeft de effecten weer die als klein worden geïnterpreteerd (effectgrootte in absolute waarde vanaf 0,20 tot 0,50); de iets donkerder kleur groen geeft de effecten weer die als matig worden geïnterpreteerd (effectgrootte in absolute waarde vanaf 0,50 tot 0,80) en de donkerste kleur groen geeft de effecten weer die als groot worden geïnterpreteerd (effect in absolute waarde van 0,80 en groter). Niet gearceerde effecten zijn in absolute waarde kleiner dan 0,20 en worden als niet relevant geïnterpreteerd. De resultaten voor thuistaal staan in tabel 3.12. Merk op dat voor de variabele advies VO alleen de resultaten van de hoofdniveaus zijn gepresenteerd. De resultaten van de tussenniveaus liggen daar tussenin.

Tabel 3.5 Gecorrigeerde effecten op leerlingniveau in jaargroep 8 in 2014*

Variabele	Contrast	Taal			Rekenen		
		WS	Sp	BL	G/B	BPV	MMTG
Geslacht	Meisjes - Jongens	-0,04	0,28	0,10	-0,37	-0,40	-0,30
Leertijd	Regulier - Vertraagd	0,52	0,70	0,62	0,70	0,71	0,70
Formatiegewicht	0.0 - 0.3	0,48	0,36	0,60	0,43	0,47	0,51
	0.3 - 1.2	0,26	-0,22	0,01	-0,20	-0,15	-0,12
	0.0 - 1.2	0,75	0,14	0,61	0,23	0,32	0,38
Advies VO	vmbo-KB - vmbo-BB	0,77	0,68	1,23	1,14	1,14	1,20
	vmbo-GT - vmbo-KB	0,43	0,39	0,67	0,59	0,61	0,57
	havo - vmbo-GT	0,94	0,86	1,16	1,13	1,19	1,20
	vwo - havo	1,45	1,41	1,64	1,53	1,59	1,58
	vwo - vmbo-BB	3,59	3,34	4,70	4,39	4,54	4,54
Toets	Eindtoets-Niveautoets	1,00	0,99	1,37	1,35	1,33	1,32

* vet is significant ($p=0,01$);

3.2.1 Effecten op leerlingniveau

Geslacht

Voor het verschil in resultaten tussen jongens en meisjes zien we in alle gevallen significante en betekenisvolle effecten. Deze zijn bij de rekenonderwerpen het grootst en in het voordeel van de jongens. Bij de taalonderwerpen zien we meer variatie in richting. Jongens hebben een hogere vaardigheid bij Woordenschat, maar een lagere vaardigheid bij Spelling en Begrijpend lezen.

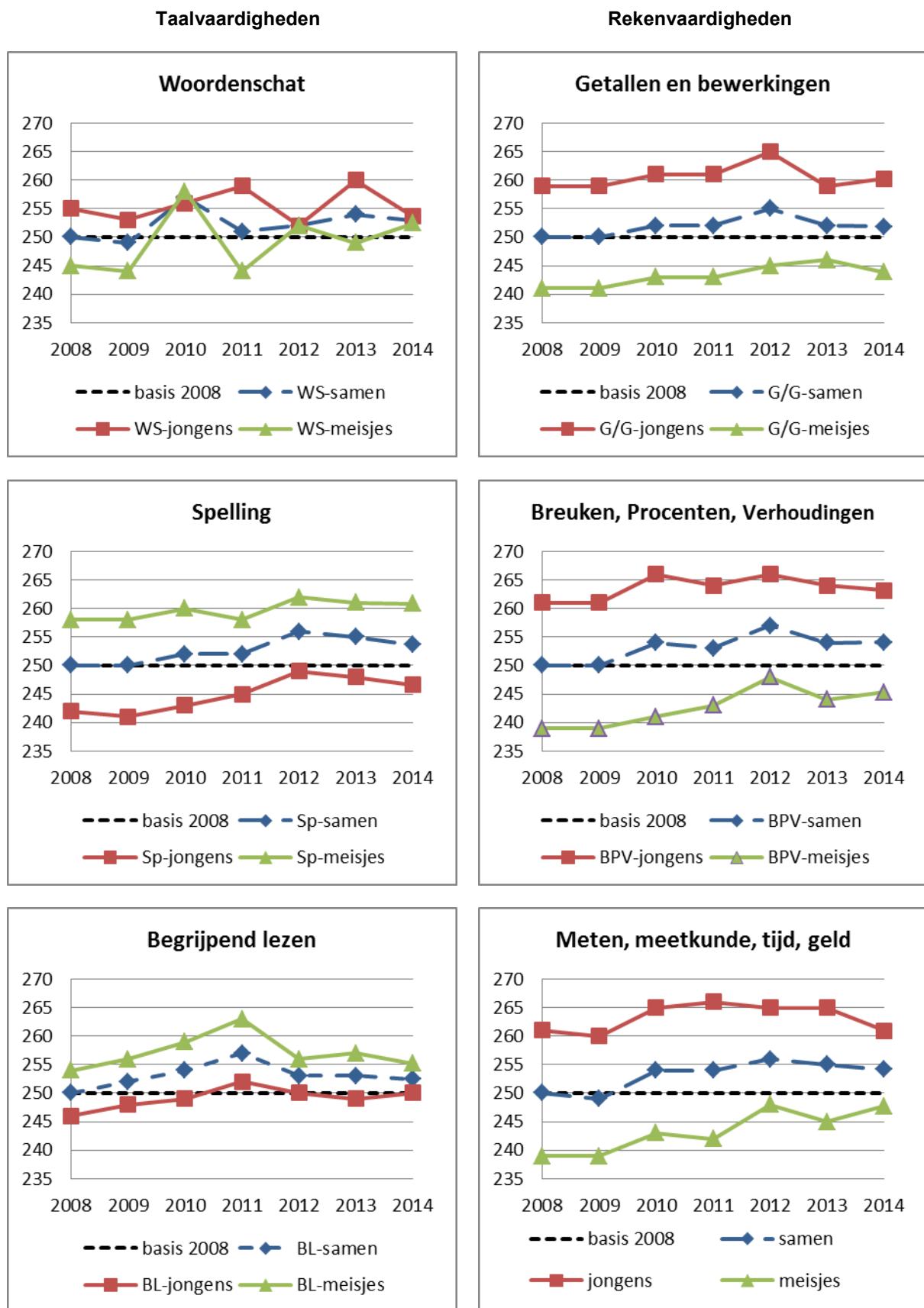
In tabel 3.6 hebben we deze resultaten onderscheiden naar geslacht bij elkaar gezet. De resultaten zijn in beeld gebracht met een grafische weergave per vaardigheid in figuur 3.3. In tabel 3.7 is het verschil in vaardigheid tussen jongens en meisjes weergegeven (per afnamejaar: gemiddelde vaardigheid jongens min gemiddelde vaardigheid meisjes). Die verschillen zijn ook afgebeeld in figuur 3.4.

Bij de meeste vaardigheden zien we dat dan wel jongens dan wel meisjes een hogere vaardigheid hebben. Bij de rekenvaardigheden zijn dat de jongens, en bij Begrijpend lezen en Spelling zijn dat de meisjes. Een uitzondering is de vaardigheid Woordenschat waarbij jongens in veel van de jaren een hogere vaardigheid laten zien, maar niet in alle jaren. We zien daar ook een grillig beeld over de jaren heen, waarbij soms de vaardigheid van jongens daalt en die van meisjes stijgt of omgekeerd. Daar waar bij de andere vaardigheden gevonden wordt dat een stijging of daling in vaardigheid over de jaren heen bij de ene groep ook vaak gepaard gaat met eenzelfde type verandering bij de andere groep (de vaardigheden van jongens en meisjes zijn positief gecorreleerd), is dat niet het geval bij Woordenschat. De verschillen tussen de twee groepen worden wel kleiner over de jaren heen. Dat wil zeggen: bij alle vaardigheden zien we een trend over de jaren heen dat het (absolute) verschil tussen de vaardigheid van de jongens en meisjes kleiner wordt. Dat houdt niet in dat dit bij elke vaardigheid in alle jaren gevonden wordt, maar als de trendlijn getrokken wordt, is te zien dat deze voor alle vaardigheden richting de nullijn gaat.

Tabel 3.6 Vaardigheidsscores per onderwerp naar geslacht van 2008 tot en met 2014

Taal	Geslacht	Gemiddelde vaardigheidsscore							Verskil 2014-2008
		2008	2009	2010	2011	2012	2013	2014	
Woordenschat	samen	250	249	257	251	252	254	253	3
	jongens	255	253	256	259	252	260	254	-1
	meisjes	245	244	258	244	252	249	252	7
Spelling	samen	250	250	252	252	256	255	254	4
	jongens	242	241	243	245	249	248	247	5
	meisjes	258	258	260	258	262	261	261	3
Begrijpend lezen	samen	250	252	254	257	253	253	252	2
	jongens	246	248	249	252	250	249	250	4
	meisjes	254	256	259	263	256	257	255	1
Rekenen-Wiskunde									
Getallen en bewerkingen	samen	250	250	252	252	255	252	252	2
	jongens	259	259	261	261	265	259	260	1
	meisjes	241	241	243	243	245	246	244	3
Breuken, procenten en verhoudingen	samen	250	250	254	253	257	254	254	4
	jongens	261	261	266	264	266	264	263	2
	meisjes	239	239	241	243	248	244	245	6
Meetkunde, meten, tijd en geld	samen	250	249	254	254	256	255	254	4
	jongens	261	260	265	266	265	265	261	0
	meisjes	239	239	243	242	248	245	248	9

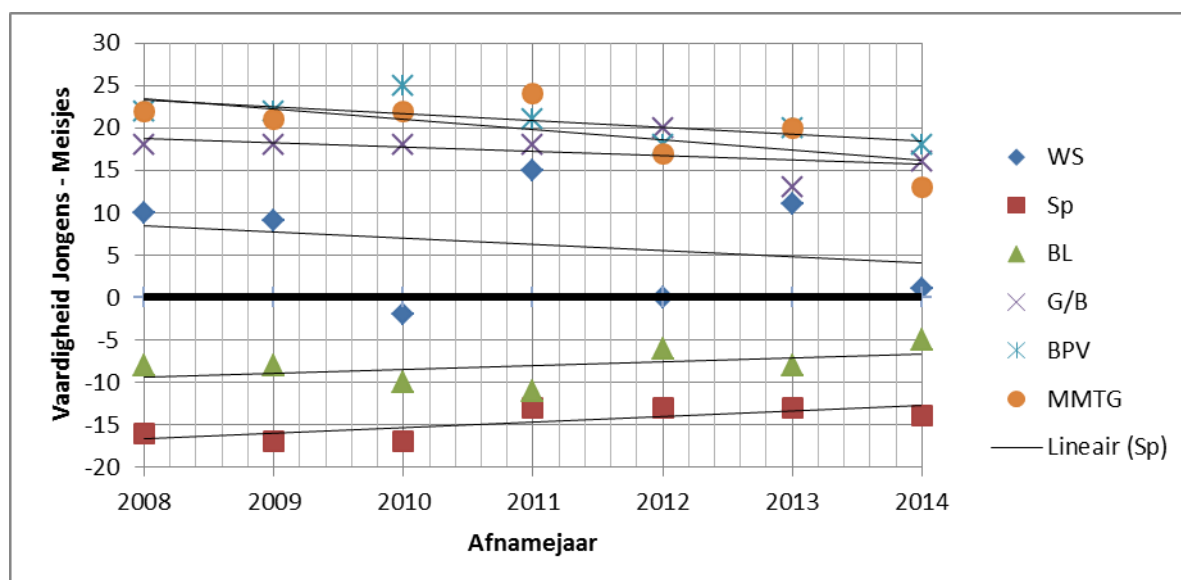
Figuur 3.3 Trends in vaardigheden naar geslacht



Tabel 3.7 Wijziging in het verschil tussen jongens en meisjes van 2008 tot en met 2014 weergegeven in punten op de schaal.

Verschil J - M	Taal			Rekenen		
	Jaar	WS	Sp	BL	G/B	BPV
2008	10	-16	-8	18	22	22
2009	9	-17	-8	18	22	21
2010	-2	-17	-10	18	25	22
2011	15	-13	-11	18	21	24
2012	0	-13	-6	20	18	17
2013	11	-13	-8	13	20	20
2014	1	-14	-5	16	18	13
Gemiddeld	6	-15	-8	17	21	20

Figuur 3.4 Verschil tussen jongens en meisjes van 2008 tot en met 2014, en de bijbehorende lineaire trendlijnen



Leertijd

Zoals gesteld wordt de variabele leertijd hier, en in eerdere rapporten, geanalyseerd voor twee groepen: de groep vertraagde leerlingen die gezien hun leeftijd (teldatum 1 oktober) het basisonderwijs al verlaten hadden kunnen hebben, en de groep die gezien hun leeftijd nog thuis hoort in het basisonderwijs. De vertraagde leerlingen hebben een verlengde leertijd in het basisonderwijs gehad ten opzichte van de reguliere leerlingen. In welk leerjaar ze de verlengde leertijd hebben gekregen is daarbij niet bekend. Een relatief groot deel zal dit mogelijk al in de kleuterperiode (groep 1 of 2) hebben opgelopen. Ook is er een relatief grote groep leerlingen geboren in de maanden augustus en september die dit overkomt. Dit zouden vergeleken met hun jaargroepgenoten jonge leerlingen zijn als ze het reguliere traject zouden volgen. Een overzicht van de gevonden effecten onder het gecorrigeerde model is gegeven in tabel 3.8.

Over de jaren heen is gebleken dat vertraagde leerlingen op in dit onderzoek gemeten vaardigheden gemiddeld minder goed presteren dan niet vertraagde leerlingen. Bij Woordenschat is het effect gemiddeld 0,50, bij Spelling 0,60 en bij Begrijpend lezen en de rekenvaardigheden rond de 0,70. Gemiddeld genomen is dat te kwalificeren als een matig effect.

Over de jaren heen is er relatief weinig variatie binnen een vaardigheid voor wat betreft het effect van de leertijd. Bij Woordenschat is die nog het grootst (een verschil van 0,10 bij effecten variërend van 0,44 tot en met 0,54, met een gemiddelde van 0,50), maar bij de overige vaardigheden varieert dat gemiddeld niet meer dan 0,05 effectpunten.

Tabel 3.8 *Het gevonden effect per jaar voor de variabele leertijd berekend met het model waarbij binnen een jaar gecorrigeerd is voor geslacht, formatiegewicht en stratum (licht oranje gemarkeerde cellen: eerste en meest recente peiljaar)*

	2008	2009	2010	2011	2012	2013	2014
Woordenschat	0,54	0,54	0,44	0,47	0,50	0,49	0,52
Spelling	0,69	0,69	0,71	0,68	0,69	0,72	0,70
Begrijpend lezen	0,60	0,60	0,60	0,58	0,60	0,58	0,62
Rekenen: G/B	0,67	0,67	0,70	0,67	0,67	0,67	0,70
Rekenen: BPV	0,71	0,71	0,75	0,68	0,71	0,69	0,71
Rekenen: MMTG	0,69	0,69	0,71	0,67	0,71	0,65	0,70

Tabel 3.9 *Gemiddelde vaardigheid voor de reguliere en de vertraagde leerlingen van 2008 tot en met 2014 (licht oranje gemarkeerde cellen: eerste en meest recente peiljaar).*

Vaardigheid	Groep	Gemiddelde vaardigheidsscore							Vershil
		2008	2009	2010	2011	2012	2013	2014	2014-2008
Taal									
Woordenschat	samen (2014)	250	249	257	251	252	254	253	3
	regulier	256	254	262	256	257	259	257	1
	vertraagd	225	225	234	229	228	231	227	2
Spelling	samen (2014)	250	250	252	252	256	254	254	4
	regulier	257	256	258	258	262	261	259	2
	vertraagd	221	221	222	223	227	224	224	3
Begrijpend Lezen	samen (2014)	250	252	254	257	253	253	252	2
	regulier	257	258	260	263	259	258	258	1
	vertraagd	223	224	226	230	225	225	223	0
Rekenen-Wiskunde									
Getallen en bewerkingen	samen (2014)	250	249	252	252	255	252	252	2
	regulier	257	256	259	258	261	258	257	0
	vertraagd	224	224	224	225	228	225	223	-1
Breuken, procenten en verhoudingen	samen (2014)	250	250	254	254	257	254	254	4
	regulier	257	256	260	259	263	260	260	3
	vertraagd	222	222	225	226	228	225	224	2
Meetkunde, meten, tijd en geld	samen (2014)	250	250	254	254	256	255	254	4
	regulier	257	256	260	259	262	260	260	3
	vertraagd	223	222	226	227	226	228	224	1

Ook als we kijken naar de gevonden gemiddelde waarden over de jaren heen zien we dat er weinig variatie is. In tabel 3.9 worden de gemiddelde vaardigheden gegeven voor iedere leertijdgroep. Over het algemeen volgen de subgroepen Regulier en Vertraagd dezelfde trend (gemiddelde correlatie is 0,88, en de laagste 0,70 bij rekenen: Meten, meetkunde, tijd en geld). Als we nu kijken naar de het verschil tussen 2014 en

2008 voor de gehele groep en de voor de subgroepen valt op dat de groei in vaardigheid voor de subgroepen kleiner is dan voor de gehele groep. Bij rekenen – Getallen en bewerkingen – is er zelfs sprake van een daling in vaardigheid bij de vertraagde leerlingen. Dit komt vreemd over, maar is te verklaren vanuit de veranderende verdeling van de reguliere en de vertraagde leerlingen in de onderzoekspopulatie over de jaren heen

Er is in de periode 2008-2014 een duidelijk dalende trend in het percentage vertraagde leerlingen binnen de onderzoekspopulatie van leerlingen in jaargroep 8 van het reguliere basisonderwijs. Deze gegevens zijn gebaseerd op de leerlingen uit de onderzoekspopulatie die een eindtoets gemaakt hebben, maar ze komen overeen met gegevens uit andere bronnen. In een publicatie van de inspectie (Paper_Zittenblijven_ORD_2013_vs3.pdf) wordt informatie gegeven over het percentage zittenblijvers in 2011. Daarin wordt gesproken over 18,0% leerlingen dat minimaal 1 jaar vertraagd was. Dat komt exact overeen met het percentage dat in 2011 gevonden is de JPON-populatie (zie tabel 3.10).

Tabel 3.10 Percentage reguliere en vertraagde leerlingen van 2008 tot en met 2014 in de JMTR-populatie

Leertijdcategorie \ Jaar	2008	2009	2010	2011	2012	2013	2014
Regulier	79,8	80,9	81,6	82,0	82,9	83,1	84,3
Vertraagd	20,2	19,1	18,4	18,0	17,1	16,9	15,7

Een andere bron waarmee de gevonden trend van het dalend percentage vertraagde en dus oudere leerlingen te vergelijken is, zijn de bestanden die DUO beschikbaar stelt. Hoewel in de gegevens niet opgenomen is in welk leerjaar de leerlingen zitten (en daarmee niet bekend is of een leerling vertraagd is), is wel de verdeling over de verschillende leeftijdsgroepen bekend. Als binnen dat bestand alleen naar de verhouding leerlingen van 11 jaar in vergelijking met die van 12 jaar en ouder gekeken wordt, dan wordt eenzelfde trend gevonden.

Tabel 3.11 Percentage 11-jarigen ten opzichte van de 12 jaar en oudere leerlingen in de DUO-bestanden

Leeftijd \ Jaar	2008	2009	2010	2011	2012	2013	2014
11 jaar	81,3	82,2	82,6	82,7	83,5	83,6	84,4
12 jaar of ouder	18,7	17,8	17,4	17,3	16,5	16,4	15,6

De percentages zijn hier uiteraard niet geheel vergelijkbaar met die in tabel 3.10, aangezien niet alle 11-jarigen in jaargroep 8 zitten: een deel van hen is gedoubleerd en zit in jaargroep 7 (of zelfs in jaargroep 6). Daarnaast zitten er ook voorlijke leerlingen van 10 jaar al in jaargroep 8.

De resultaten in tabel 3.9 laten duidelijk zien dat de prestaties van de niet vertraagde (reguliere) leerlingen beter zijn dan die van de vertraagde leerlingen. Ook is te zien dat de stijgende jaartrend minder groot, of zelfs afwezig, is binnen de groep vertraagde en niet vertraagde leerlingen dan in de gehele populatie. Die sterker stijgende trend in de gehele populatie is dan ook voor een heel groot deel te danken aan de afname van het percentage vertraagde leerlingen.

Als we het percentage vertraagde leerlingen in 2008 vergelijken met dat in 2014 zien we dat er een afname is van 4,5% op de 20,2% van 2008. Het is een interessante vraag hoe dit te interpreteren. Is het zo dat de leerlingen in de jaren ervoor beter opgeleid worden zodat er minder noodzaak is voor leerlingen om te

blijven zitten? Of is het beleid geworden om leerlingen minder vaak te laten doubleren of minder vaak langer te laten 'kleuteren'? Wat zijn de eigenschappen van de 4,5% leerlingen die in 2008 wel doubleren maar in 2014 blijkbaar niet? Hoe verhoudt de vaardigheid van die groep zich tot die van de vertraagde leerlingen die in 2014 gedoubleerd zouden zijn? En wat zou in vaardigheid zijn in groep 8 als die 4,5% leerlingen in 2014 wel waren blijven zitten en pas het jaar erop gemeten zouden zijn? Zou dat extra jaar voor die leerlingen een hogere vaardigheid opleveren en zou dat betekenen dat dan de gevonden stijging in vaardigheid in jaren eigenlijk nog hoger is? De meeste van deze vragen zijn op basis van de hier nu beschikbare gegevens of aannames helaas niet te beantwoorden. Dit fenomeen zou verder onderzocht moeten worden door aanvullende dataverzameling, of wellicht dat andere rapporten waarin onderzoek naar doubleren is gedaan hier meer licht op kunnen schijnen.

Er kan wel gesteld worden dat er op grond van de huidige gegevens gesproken kan worden over een groei in vaardigheid over de jaren heen. Een extra resultaat is dat deze leerwinst gemeten wordt met een duidelijke afname aan leertijd. Als we uitgaan van een leertijd van 8 jaar voor de reguliere leerlingen en 9 jaar voor de vertraagde, betekent dat een verkorting van de gemiddelde leertijd van 0,045 jaar. Met 150.000 leerlingen is dat in totaal van 6750 leerjaar die minder besteed is aan onderwijs, terwijl de prestatie beter wordt.

De groep reguliere leerlingen kan ook nog uitgesplitst worden in twee relevante groepen: de vervroegde leerlingen die op 1 oktober nog geen 11 zijn, en de leerlingen die dan wel 11 zijn³. Bij die tweedeling was in 2008 van de 79,9% 'reguliere leerlingen' het percentage vervroegde leerlingen 4,2%, en het percentage 11-jarigen 75,7%. In 2014 was het percentage reguliere leerlingen 84,4%, met 8,0% vervroegde leerlingen, tegenover 76,4% leerlingen van 11 jaar. De grootste stijging in het aantal reguliere leerlingen komt dus door leerlingen uit de vervroegde groep. Er is dus sprake van een stijging van 3,8% van leerlingen die een jaar korter onderwijs heeft gehad dan in 2008 door de toename van vervroegde leerlingen, en 4,5% daling in de vertraagde leerlingen (12 jaar en ouder op 1 oktober) die een jaar langer onderwijs heeft gehad. Dat betekent een stijging van 8,3% leerlingen die een jaar korter onderwijs heeft gehad dan in 2008 gebruikelijk was. Met 150.000 leerlingen is dat in totaal bijna 12500 leerjaren minder besteed aan onderwijs, terwijl de prestaties beter worden.

Formatiegewicht

Voor de variabele formatiegewicht verdelen we de leerlingen in drie categorieën: leerlingen zonder gewicht (0.00), met een laag gewicht (0.30) en met een hoog gewicht (1.20). Het contrast tussen leerlingen zonder gewicht en met een laag gewicht levert overal een significant en betekenisvol effect op. Leerlingen zonder gewicht hebben steeds een hogere vaardigheid. De effectgrootte is klein en in twee gevallen matig te noemen, namelijk bij de onderwerpen Begrijpend lezen en Meten, meetkunde, tijd en geld).

De contrasten zijn gegeven in tabel 3.5 – dit zijn gecorrigeerde effecten – waarbij gecorrigeerd wordt voor stratum, leertijd en geslacht. Daarin is te zien dat het contrast 'geen gewicht' versus de 'gewichtsl leerlingen' (0.30 dan wel 1.20) bijna overal positieve effecten laat zien met een gemiddelde tussen de 0,4 en de 0,5. Dat staat voor een klein, maar bijna matig effect. Opvallend is dat het verschil tussen de **geen** gewicht en **laag**gewicht leerlingen minder variatie vertoont dan het verschil tussen de **geen** gewicht en **laag** gewicht leerlingen. Het verschil tussen de leerlingen met gewicht 0.00 en 0.30 varieert van 0,36 (bij Spelling; klein effect) tot 0,60 (bij Begrijpend lezen; matig effect), terwijl dat bij de leerlingen met gewicht 0.00 en 1.20 varieert van 0,14 (bij Spelling; niet significant, geen effect) tot 0,75 (bij Woordenschat; bijna een groot effect). Kijken we naar de verschillen tussen de leerlingen met laag en hoog gewicht dan presteren de leerlingen met gewicht 0.30 alleen bij Woordenschat significant beter (klein effect), terwijl bij Spelling en Getallen en bewerkingen de leerlingen met gewicht 1.20 significant beter presteren (klein effect). Bij de overige rekenvaardigheden en Begrijpend lezen was er sprake van een effect kleiner dan 0,20, dat in twee van de drie gevallen ook significant was.

³ De leerlingen die ouder dan 11 zijn op 1 oktober zijn de vertraagde leerlingen.

Net als voorgaande jaren vinden we het grootste verschil bij het contrast geen gewicht versus hoog gewicht voor het onderwerp Woordenschat. De berekende effectgrootte is met 0,75 als matig te classificeren, maar zit wel net als in eerdere jaren dicht bij 0,80 en is dus bijna groot te noemen. Voor Begrijpend lezen is het effect ook matig en voor Spelling zelfs betekenisloos. Bij Rekenen leveren al deze drie contrasten kleine effecten op.

Het contrast tussen leerlingen met een laag versus een hoog gewicht levert een wisselend beeld op. Als we corrigeren voor geslacht, leertijd en stratum zien we dat de leerlingen met het laagste gewicht alleen slechter presteren bij Woordenschat, maar zelfs beter presteren bij Spelling. Ook als we naar de ongecorrigeerde effecten kijken door de gemiddelde geobserveerde vaardigheid te vergelijken zien we dat de leerlingen met gewicht 1.20 beter presteren dan leerlingen met gewicht 0.30, zij het dat het effect dan wel -0,16 wordt, en daarmee niet betekenisvol. Bij de overige vaardigheden is het verschil (zowel gecorrigeerd als niet gecorrigeerd) niet betekenisvol. Merk op dat het gevonden effect bij Getallen en bewerkingen afgerond wel -0,20 haalt, maar niet afgerond de grens van 0,20 niet haalt.

Alle gevonden betekenisvolle effecten zijn significant. De gevonden effecten lijken sterk op de effecten die in eerdere jaren gevonden zijn.

VO-advies

Voor deze variabele onderscheiden we een vijftal vo-adviezen: vmbo-BB, vmbo-KB, vmbo-GT, havo en vwo. Gegevens over de tussenliggende adviezen zijn wel beschikbaar, maar de verschillen liggen doorgaans halverwege de naastliggende adviezen. Het advies havo/vwo bijvoorbeeld ligt halverwege havo- en vwo-advies wat betreft de behaalde prestaties en dat geldt ook voor de grootte van de effecten. Ter bevordering van de eenvoud van de weergave hebben we de tussenliggende adviezen niet in de tabel gezet (zie tabel 3.5). Bij deze variabele komen we zeer grote verschillen tegen. We zien dat er sprake is van contrasten die overall significant zijn en overall betekenisvolle effectgroottes opleveren. Die effectgroottes variëren van klein tot (zeer) groot. Bij het grootste contrast, dat tussen vwo en vmbo-BB, zien we effectgroottes tot 4,70, hetgeen betekent dat het verschil tussen deze groepen van leerlingen, gecorrigeerd voor geslacht, leertijd, en formatiegewicht, bijna vijf standaarddeviaties bedraagt. Het illustreert de grote spreiding in vaardigheid die we in jaargroep 8 al aantreffen. Opvallend in dit verband is ook de aanzienlijke effectgrootte bij het contrast vwo-havo. Deze leerlingen zitten in het VO soms nog bij elkaar in een klas, maar ook hier is bij alle onderdelen het verschil aanzienlijk. De kleinste verschillen vinden we bij het contrast vmbo-GT en vmbo-KB. Daar variëren de effectgroottes van 0,39 bij Spelling tot 0,67 bij Begrijpend lezen.

Sinds 2012 zijn de effecten op twee verschillende manieren uitgerekend, waarbij op een verschillende wijze met de ontbrekende waarden is omgegaan. De waarden van de alternatieve rekenwijze zijn groter en nauwkeuriger omdat er beter rekening gehouden wordt met de kleinere standaarddeviatie in de verschillende subgroepen. Deze waarden zijn dan ook sinds 2012 in de tabellen gegeven. Voor de vergelijkbaarheid over de jaren heen zijn in bijlage 3 beide berekeningswijzen gerapporteerd. In de vergelijking met de voorgaande jaren lijken in 2014 vooral bij de rekenvaardigheden de afstanden tussen de groepen met een verschillend VO-advies iets kleiner geworden, maar de effecten blijven nog steeds matig tot (meestal) groot.

Toets

Ook de toetsvorm die is gemaakt kan als variabele worden gezien. Naast de reguliere Eindtoets is er een alternatieve, gemakkelijker eindtoets beschikbaar. Tot 2013 was deze toets, aangeduid als Niveautoets (NT), alleen digitaal beschikbaar en bedoeld voor leerlingen met een leerachterstand van anderhalf jaar of meer. Vanaf 2013 is er een papieren alternatief ingevoerd, zodat de leerlingen gewoon in de klas met de andere leerlingen de toets kunnen maken. Deze versie, de Eindtoets Niveau (EN), is bestemd voor leerlingen waarvan de leerkracht verwacht dat ze doorstromen naar vmbo-BB of vmbo-KB. Deze groep is dus groter dan de groep die voor 2013 de Niveautoets maakte en bij deze groep leerlingen is er per definitie

een duidelijk verband met het VO-advies. Voor wat betreft de gevonden effecten is het dan ook weinig verrassend dat hier bij alle vaardigheden een groot effect gevonden wordt.

In de vergelijking over de jaren heen zijn voor deze variabele eigenlijk alleen 2013 en 2014 goed met elkaar te vergelijken (zie tabel 3.5). De definitie van groep die de EN maakt, is duidelijk anders dan die van de groep die vroeger de NT maakte. Er is ook te zien dat de er veel meer leerlingen de EN in 2013 en 2014 maken dan er in het verleden de NT maakten. Het effect tussen de NT en de reguliere eindtoets was in het verleden – behalve bij Spelling – ook groter dan dat nu het geval is tussen de EN en de EB. In vergelijking met 2013 is de afstand tussen de leerlingen die de EN maken en de leerlingen die de EB maken iets kleiner, maar het effect is met rond de 1,0 voor Spelling en Woordenschat en meer dan 1,3 voor Begrijpend lezen en de drie rekenonderwerpen nog steeds groot.

Thuis taal

Nadat deze achtergrondvariabele in 2012 en 2013 niet was opgevraagd, is dat in 2014 wel weer gebeurd. Bij deze variabele is naar twee verschillende opdelingen gekeken. De eerste opdeling is een driedeling met de volgende categorieën: (1) de leerling spreekt thuis alleen Nederlands, (2) de leerling spreekt naast Nederlands ook een andere taal, of (3) de leerling spreekt thuis geen Nederlands. De tweede opdeling maakt ook onderscheid naar de verschillende thuistalen en kent vijf categorieën: Nederlands, Arabisch, Turks, Surinaams en Antilliaans en overige talen. Om de groepen voldoende groot te houden wordt hierbij geen onderscheid meer gemaakt tussen leerlingen die thuis alleen die andere taal spreken of naast die taal thuis ook Nederlands spreken. Een verschil met de jaren 2008-2011 is dat de categorie ‘Westerse taal’ niet meer is opgevraagd en nu in de categorie ‘Overige talen’ is opgenomen. Deze categorie is daardoor niet met eerdere jaren te vergelijken.

In tabel 3.12 worden de gecorrigeerde effecten gegeven, wat betekent dat hier de aanvullende effecten onderzocht worden gecorrigeerd voor formatiegewicht, stratum, geslacht en leertijd. Hierbij valt op dat de taal die thuis gesproken wordt bij Spelling en de drie rekenvaardigheden geen effect van betekenis heeft. Wel zien we kleine effecten bij Woordenschat en Begrijpend lezen, en een matig effect bij Woordenschat waar het gaat om het contrast ‘alleen Nederlands’ versus ‘alleen een andere taal’.

Tabel 3.12 Gecorrigeerde effecten voor de variabele thuis taal in jaargroep 8 in 2014*

Variabele	Contrast	Taal			Rekenen		
		WS	Sp	BL	G/B	BPV	MMTG
Taal							
	Alleen NLS - NLS+Ander	0,39	-0,13	0,32	-0,02	0,05	0,03
	NLS+Ander - Alleen Ander	0,18	0,00	-0,02	-0,03	-0,05	0,04
	Alleen NLS - Alleen Ander	0,56	-0,13	0,31	-0,05	-0,01	0,07
	Alleen NLS - Turks(+NLS)**	0,88	-0,03	0,52	-0,08	0,02	0,06
	Alleen NLS - Arabisch(+NLS)	0,50	-0,10	0,28	0,00	0,00	0,08
	Alleen NLS - Sur./Ant.(+NLS)	0,63	0,21	0,72	0,55	0,60	0,60
	Alleen NLS - Ov. ander taal(+NLS)	0,34	-0,26	0,14	-0,16	-0,11	-0,02

* vet is significant (p = 0,01);

** (+NLS): eventueel wordt naast de genoemde taal thuis ook Nederlands gesproken

Meest lichte kleur groen geeft de effecten weer die als klein worden geïnterpreteerd; de iets donkerder kleur groen geeft de effecten weer die als matig worden geïnterpreteerd en de donkerste kleur groen geeft de effecten weer die als groot worden geïnterpreteerd (zie tabel 2.1). Niet gearceerde effecten zijn in absolute waarde kleiner dan 0,20 en worden als niet relevant geïnterpreteerd.

De enige uitzondering is de groep die (al dan niet naast Nederlands) een taal uit Suriname of de Antillen spreekt: die groep presteert – ook na de genoemde correcties – lager op alle genoemde vaardigheden, met een klein effect bij Spelling en een matig effect bij de andere vaardigheden. Het grootste effect zien we overigens bij het contrast ‘alleen Nederlands’ versus ‘Turks’⁴ voor Woordenschat. Daar is met 0,88 sprake van een effectgrootte die als groot gekwalificeerd kan worden. Voor alle onderscheiden contrasten worden van de zes vaardigheden bij Woordenschat de grootste effecten gevonden. Alleen bij het contrast ‘alleen Nederlands’ versus ‘Surinaams/Antilliaans’ wordt een groter effect gevonden bij Begrijpend lezen dan bij Woordenschat. Tot slot valt op dat bij Spelling de categorie ‘alleen Nederlands’ vaak een niet positief effect oplevert bij de gecorrigeerde effecten. Bij het contrast ‘alleen Nederlands’ versus ‘overige andere talen’ levert dat zelfs een significant verschil met een klein effect op (-0,26).

Bij Woordenschat is wel te zien dat de thuistaal invloed heeft op de vaardigheid: hoe meer Nederlands er gesproken wordt, hoe beter de woordenschat, al is het verschil tussen ‘alleen een andere taal’ en ‘Nederlands en een andere taal’ niet betekenisvol. Dat is ook het geval bij Begrijpend lezen, waarbij een zelfde beeld te zien is als bij Woordenschat, zij het dat de effecten in de meeste gevallen iets kleiner zijn.

In de tabellen 3.13 en 3.14 zijn de ongecorrigeerde gemiddelde vaardigheden gegeven voor de verschillende groepen over de jaren heen. Hierbij ontbreken dus de gegevens van 2012 en 2013 omdat in die jaren geen gegevens over thuistaal zijn verzameld. Om een idee te krijgen van hoe de resultaten van 2014 zich verhouden tot eerdere peilingen is ook het verschil gegeven van 2014 ten opzichte van 2008. Zo is de trend vanaf het eerste meetpunt te zien. Een negatieve waarde is indicatief voor een achteruitgang binnen een groep. Merk op dat dalingen binnen een groep gepaard kunnen gaan met een stijging over de gehele groep, als de samenstelling van de gehele groep over de jaren verschilt.

In de tabellen 3.13 en 3.14 zien we wel dat er bij de rekenvaardigheden en spelling in de meeste gevallen wel verschillen gevonden worden, maar gezien het ontbreken van effecten in het gecorrigeerde model lijken deze verschillen vooral verklaard te kunnen worden door een oververtegenwoordiging van leerlingen met een gewicht 0.30 of 1.20 bij de groepen die niet alleen Nederlands spreken. Bij de groep die alleen Nederlands spreekt, zien we dan ook een percentage van rond de 90% leerlingen zonder gewicht, terwijl dat bij overige groepen aanzienlijk lager ligt met overal percentages van onder de 70%.

⁴ Deze leerlingen spreken thuis Turks, maar kunnen thuis naast Turks eventueel ook Nederlands spreken.

Tabel 3.13 Gemiddelde vaardigheidsscores uitgesplitst naar thuistaal (driedeling) per jaar

Thuistaal Algemeen		Gemiddelde vaardigheid per jaar					Vershil
Taalvaardigheden	Thuistaal	2008	2009	2010	2011	2014	2014 - 2008
Woordenschat	Alleen Nederlands	255	253	262	256	257	2
	NLs + andere taal	220	223	231	223	223	3
	Alleen een andere taal	201	208	209	207	210	9
Spelling	Alleen Nederlands	251	251	252	252	254	3
	NLs + andere taal	247	249	251	255	252	5
	Alleen een andere taal	238	243	245	244	249	11
Begrijpend lezen	Alleen NLs	254	256	257	261	255	1
	NLs + andere taal	230	228	234	237	224	-6
	Alleen een andere taal	219	217	220	225	220	1
Rekenvaardigheden	Thuistaal	2008	2009	2010	2011	2014	Vershil
Getallen & bewerkingen	Alleen Nederlands	252	251	254	253	253	0
	NLs + andere taal	244	244	243	247	244	-1
	Alleen een andere taal	239	239	238	239	243	4
Breuken, procenten & verhoudingen	Alleen Nederlands	252	252	256	256	255	1
	NLs + andere taal	239	239	243	244	241	0
	Alleen een andere taal	233	233	235	237	240	6
Meten, meetkunde, tijd & geld	Alleen Nederlands	252	252	256	256	256	2
	NLs + andere taal	237	236	242	242	241	2
	Alleen een andere taal	232	232	232	234	236	4

In de tabellen 3.13 en 3.14 is lichtblauw gemarkeerd wanneer in 2014 de hoogste vaardigheid voor een groep gevonden is, in vergelijking met de jaren 2008 tot en met 2011, en lichtpaars wanneer in 2014 de laagste vaardigheid is gevonden. Op die manier is te zien dat de groep die alleen een andere taal spreekt in 2014 bij de meeste vaardigheden nog wel steeds het minst presteert, maar minder laag dan in de eerdere jaren. Vergeleken met 2008 zijn de waarden vaak gestegen, behalve dan bij Begrijpend lezen.

Opvallend lager dan in voorgaande jaren zijn de uitkomsten voor Begrijpend lezen bij leerlingen die naast Nederlands thuis nog een andere taal spreken. Uitgesplitst naar thuistaal valt op dat de leerlingen die thuis Arabisch spreken beter zijn gaan presteren in Spelling en de drie rekenschalen, maar nauwelijks in Begrijpend lezen. De leerlingen die thuis Surinaams/Antilliaans spreken presteren nu minder dan in 2011, vooral bij Woordenschat en Begrijpend lezen.

Tabel 3.14 Gemiddelde vaardigheden voor de groep die thuis alleen Nederlands spreekt en verschillende groepen die thuis (ook) een andere taal dan Nederlands spreken

Thuis taal gedetailleerd		Gemiddelde vaardigheid per jaar					Vershil 2014 -
Taalvaardigheden	Thuis taal*	2008	2009	2010	2011	2014	gem('08-'11)
Woordenschat	Alleen Nederlands	255	253	262	256	257	0
	Turks	184	190	190	189	192	4
	Arabisch	201	208	213	208	211	4
	Surinaams/Antilliaans	212	216	225	220	212	-7
	<i>Overige talen</i>					227	30
Spelling	Alleen Nederlands	251	251	252	252	254	3
	Turks	238	244	242	243	243	1
	Arabisch	235	241	244	243	248	7
	Surinaams/Antilliaans	230	231	236	240	232	-2
	<i>Overige talen</i>					258	-4
Begrijpend lezen	Alleen Nederlands	254	256	257	261	255	-2
	Turks	209	205	211	215	206	-4
	Arabisch	215	212	219	225	219	1
	Surinaams/Antilliaans	212	210	213	222	203	-11
	<i>Overige talen</i>					235	21
Rekenvaardigheden	Thuis taal*	2008	2009	2010	2011	2014	Vershil
Getallen & bewerkingen	Alleen Nederlands	252	251	254	253	253	0
	Turks	240	240	237	240	242	2
	Arabisch	235	236	238	241	240	3
	Surinaams/Antilliaans	219	218	214	215	216	0
	<i>Overige talen</i>					251	2
Breuken, procenten & verhoudingen	Alleen Nederlands	252	252	256	256	255	1
	Turks	230	230	234	234	236	4
	Arabisch	231	231	235	236	240	6
	Surinaams/Antilliaans	213	213	213	218	215	0
	<i>Overige talen</i>					249	6
Meten, meetkunde, tijd & geld	Alleen Nederlands	252	252	256	256	256	2
	Turks	229	228	230	231	234	4
	Arabisch	227	227	231	234	235	5
	Surinaams/Antilliaans	211	211	211	215	214	2
	<i>Overige talen</i>					245	11

* Thuis taal: bij de andere talen dan Nederlands kan naast die andere taal thuis ook Nederlands gesproken worden

De variabele IJK

De variabele die wordt aangeduid met de letters IJK, komt voort uit de mogelijkheid voor de leerkracht om de leerling een indicatie mee te geven, waaruit blijkt dat de leerling moeite kan hebben met de reguliere Eindtoets. Omdat in 2012 en 2013 de code voor de thuistaal ontbrak op het antwoordblad is deze variabele toen als alternatief gebruikt. De categorisering voor de IJK-code komt voort uit de volgende aanduiding⁵:

- Code I: (Allochtone) leerlingen die aan het begin van jaargroep 8 vier jaar of korter in Nederland zijn en die het Nederlands onvoldoende beheersen om de opgaven in de Eindtoets goed te kunnen lezen.
- Code J: Leerlingen die naar verwachting naar het (voortgezet) speciaal onderwijs of naar het praktijkonderwijs (pro) gaan.
- Code K: Leerlingen die naar verwachting in aanmerking komen voor het leerwegondersteunend onderwijs (lwoo).

Tabel 3.15 Effecten van de variabelen IJK, gemiddelden en standaarddeviaties in 2011-2014*

Jaar		2014			2013			2012			2011		
		Eff.	gem	sd	Eff.	gem	sd	Eff.	gem	sd	Eff.	gem	sd
Taalvaardigheden													
Woordenschat	Geen IJK		257	48		258	50		256	48		255	48
	Wel IJK	-0,92	202	50	-0,91	204	50	-0,92	201	50	-1,10	202	48
Spelling	Geen IJK		257	48		258	50		260	48		255	48
	Wel IJK	-0,87	207	49	-0,93	205	49	-0,81	213	49	-0,98	208	48
Begrijpend lezen	Geen IJK		257	47		258	50		258	47		262	47
	Wel IJK	-1,19	190	49	-1,14	192	50	-1,15	193	49	-1,34	198	47
Rekenvaardigheden													
Getallen en bewerkingen	Geen IJK		256	47		257	47		260	47		257	47
	Wel IJK	-1,16	194	49	-1,14	195	48	-1,25	194	48	-1,37	192	47
Breuken, procenten en verhoudingen	Geen IJK		258	47		258	47		262	46		257	47
	Wel IJK	-1,13	196	49	-1,16	195	48	-1,22	197	47	-1,27	197	47
Meten, meetkunde, tijd en geld	Geen IJK		258	47		259	47		261	47		258	48
	Wel IJK	-1,12	196	49	-1,12	197	49	-1,16	197	48	-1,26	199	48

* Alle vetgedrukte verschillen zijn significant ($p = 0,01$)

Leerlingen in de categorieën I en J doen doorgaans niet mee aan de Eindtoets Basis, maar kunnen de Niveautoets voorgelegd krijgen. Voor leerlingen in de categorie K is dat zeker het geval. Daarvan vinden we er ook het meest terug in de data. Er is een grote samenhang tussen formatiegewicht en deze speciale codes. Leerlingen met de code K komen procentueel het meest voor bij successievelijk 1.20- en 0.30-leerlingen. Dat geldt ook voor de codes I en J, maar dan met veel lagere percentages.

In tabel 3.15 zien we dat bij alle onderwerpen het contrast tussen leerlingen die wel een IJK-code hebben en leerlingen die dat niet hebben, significant en betekenisvol is. De effectgroottes zijn zonder uitzondering groot en hebben in de meeste gevallen de omvang van meer dan een hele standaarddeviatie. Alleen bij Spelling en Woordenschat is het effect iets kleiner, maar nog steeds groot te noemen ($>0,80$).

⁵ Deze codes zijn door de leerkracht ingevuld.

In tabel 3.16 zijn de resultaten voor alleen 2014 weergegeven. De gegevens met de contrasten van de IJK-codes zijn gelijk aan die in tabel 3.15, maar voor de volledigheid zijn ook de effecten van alleen de variabele I in kaart gebracht. Deze variabele focust op de leerlingen die te kort in Nederland zijn om de taal goed te beheersen. Deze groep is heel klein. Bij de meeste vaardigheden zijn de verschillen niet significant en zijn de effectgroottes niet betekenisvol. Dat is alleen het geval bij Woordenschat. Deze geringe effecten kunnen deels ook verklaard worden uit een over-representatie van gewichtsleerlingen in deze groep. De ongecorrigeerde effecten zijn ruim 0,20 lager. Dat betekent dat als niet gecorrigeerd voor formatiegewicht, stratum, leertijd en geslacht het verschil bij Woordenschat en Begrijpend lezen groter wordt. Bij de overige vaardigheden is te zien dat de I-groep minder goed presteert, maar nergens is het effect groot genoeg om een betekenis te hebben. Daar waar het in de lijn der verwachtingen zou liggen dat de groep met code I zich zou gedragen als een achterstandsgroep, vanwege het korte verblijf in Nederland, zien we dat niet echt terug in de resultaten – althans niet meer dan al door het formatiegewicht verklaard wordt. Alleen bij Woordenschat is er een echt additioneel effect. Deze kleine groep (ongeveer een half procent van de leerlingen) lijkt daarmee geen echte achterstandsgroep.

Tabel 3.16 Effecten van de variabelen IJK en de variabele I, gemiddelden en standaarddeviaties in 2014*

2014	IJK			Alleen I				
	wel /geen	Eff.	Gem.	sd.	wel /geen	Eff.	Gem.	sd.
Taalvaardigheden	Geen IJK		257	48	Geen I		253	50
	Wel IJK	-0,92	202	50	Wel I	-0,47	218	55
Spelling	Geen IJK		257	48	Geen I		254	50
	Wel IJK	-0,87	207	49	Wel I	0,02	247	51
Begrijpend lezen	Geen IJK		257	47	Geen I		253	50
	Wel IJK	-1,19	190	49	Wel I	-0,11	235	54
Rekvaardigheden	wel /geen	Eff.	Gem.	Sd.	wel /geen	Eff.	Gem.	Sd.
Getallen en bewerkingen	Geen IJK		256	47	Geen I		252	50
	Wel IJK	-1,16	194	49	Wel I	0,07	244	52
Breuken, procenten en verhoudingen	Geen IJK		258	47	Geen I		254	50
	Wel IJK	-1,13	196	49	Wel I	0,11	247	53
Meten, meetkunde, tijd en geld	Geen IJK		258	47	Geen I		254	50
	Wel IJK	-1,12	196	49	Wel I	0,06	245	53

* Alle vetgedrukte verschillen zijn significant ($p=0,01$)

Meest lichte kleur groen geeft de effecten weer die als klein worden geïnterpreteerd; de iets donkerder kleur groen geeft de effecten weer die als matig worden geïnterpreteerd en de donkerste kleur groen geeft de effecten weer die als groot worden geïnterpreteerd (zie tabel 2.1). Niet gearceerde effecten zijn in absolute waarde kleiner dan 0,20 en worden als niet relevant geïnterpreteerd.

3.2.2 Effecten op schoolniveau

Bij de analyse van de invloed van de achtergrondvariabelen stratum, schoolgrootte, regio en verstedelijking is een correctie toegepast. Bij die variabelen is er duidelijk een relatie met andere variabelen, die maakt dat voor een goede vergelijking een gezuiverd effect moet worden berekend. Daarin zijn de variabelen geslacht, leertijd en formatiegewicht als correctiefactoren meegenomen. De gecorrigeerde effecten zijn te vinden in tabel 3.17.

Stratum

De indeling van scholen naar stratum is gebaseerd op de indeling naar formatiegewicht en we zien dat in de effecten bij de contrasten terug. Het beeld voor stratum is vergelijkbaar met dat van formatiegewicht, maar de verschillen zijn kleiner, niet altijd significant en het aantal betekenisvolle effecten is minder. Zo zien we dat bij twee van de drie rekenvaardigheden er alleen een effect is tussen stratum 1 (minder dan 10% van de gehele schoolpopulatie heeft een formatiegewicht) en stratum 3 (25% of meer van de leerlingen op school heeft een formatiegewicht 0.30 of 1.20). De leerlingen op stratum-1-scholen presteren beter, maar het gecorrigeerde effect is klein te noemen. Geen van de andere effecten bij Rekenen is betekenisvol.

Ook bij Spelling is geen van de effecten betekenisvol, maar voor Begrijpend lezen zijn alle effecten dat wel. Het effect van het contrast Stratum 1 - Stratum 2 (10 tot 25% gewichtsl leerlingen op school) is klein, evenals dat van Stratum 2 - Stratum 3. Het effect van Stratum 1 - Stratum 3 is de som van de eerste twee effecten. Met 0,45 is dat nog steeds geclassificeerd als een klein effect, maar het is bijna als een matig effect te kwalificeren. Bij Woordenschat is een vergelijkbaar beeld te zien, alleen het effect van het contrast Stratum 1 - Stratum 2 is niet betekenisvol.

Tabel 3.17 Effecten van de variabelen stratum, schoolgrootte, regio en verstedelijking (na correctie)*

Variabele	Contrast	Taal			Rekenen		
		WS	Sp	BL	G/B	BPV	MMTG
Stratum							
	Stratum 1 - Stratum 2	0,15	0,12	0,21	0,12	0,14	0,15
	Stratum 2 - Stratum 3	0,33	-0,08	0,24	0,01	0,08	0,10
	Stratum 1 - Stratum 3	0,48	0,03	0,45	0,12	0,22	0,24
Regio							
	Oost - Noord	0,00	0,01	0,03	0,05	0,05	0,03
	West - Noord	0,08	0,09	0,12	0,10	0,12	0,10
	Zuid - Noord	0,06	0,14	0,15	0,18	0,19	0,16
	West - Oost	0,07	0,08	0,09	0,06	0,07	0,07
	Zuid - Oost	0,06	0,13	0,11	0,13	0,14	0,13
	Zuid - West	-0,02	0,05	0,03	0,08	0,06	0,06
Verstedelijking							
	Sterk - Zeer sterk	-0,04	-0,10	-0,05	-0,07	-0,07	-0,06
	Matig - Sterk	0,03	-0,02	0,02	0,04	0,04	0,05
	Weinig - Matig	-0,07	-0,05	-0,05	-0,07	-0,04	-0,05
	Niet - Weinig	0,03	-0,02	-0,04	-0,01	-0,02	0,01
Schoolgrootte							
	Groot (>200 ln) - Klein (1-200 ln)	0,03	0,13	0,09	0,08	0,06	0,06

* Alle vetgedrukte verschillen zijn significant (p=0,01)

Meest lichte kleur groen geeft de effecten weer die als klein worden geïnterpreteerd; de iets donkerder kleur groen geeft de effecten weer die als matig worden geïnterpreteerd en de donkerste kleur groen geeft de effecten weer die als groot worden geïnterpreteerd (zie tabel 2.1). Niet gearceerde effecten zijn in absolute waarde kleiner dan 0,20 en worden als niet relevant geïnterpreteerd.

De ongecorrigeerde effecten zijn uiteraard groter, omdat bij de hier gepresenteerde resultaten gecorrigeerd is voor de verdeling van formatiegewicht van de individuele leerling, wat per definitie samenhangt met de variabele stratum. De effecten bij stratum kunnen zodoende gezien worden als het additionele effect als er relatief veel schoolgenoten een gewicht hebben, bovenop het effect wanneer de leerling zelf al dan niet een formatiegewicht heeft. Zonder correcties zijn de effecten tussen Stratum 1 en Stratum 2 ongeveer 0,06 effectpunten groter. Bij het contrast tussen Stratum 2 en Stratum 3 is te zien dat het effect ongecorrigeerd in het voordeel is van Stratum 2 met een verschil van 0,10 bij Spelling (-,08 wordt dan 0,02) tot 0,22 bij Woordenschat (0,33 wordt een effect van 0,55).

De ongecorrigeerde vaardigheidsschattingen laten eenzelfde beeld zien: de vaardigheden zijn stabiel over de jaren heen (zie bijlage 3). Daarbij valt op te merken dat de jaren 2008 en 2009 lastig te vergelijken zijn met de jaren erna door de verandering van de definitie van formatiegewicht, wat effectief gezien uiteraard ook tot een veranderde definitie van stratum heeft geleid.

Schoollocatie: Regio en Verstedelijking

De locatie van de school en de mate van verstedelijking leveren geen toegevoegd effect bij de variabele formatiegewicht, stratum, leertijd en geslacht: geen van de gecorrigeerde effecten is betekenisvol. Als we naar de ongecorrigeerde effecten kijken dan zien we dat leerlingen op scholen in zeer sterk stedelijke gebieden bij Woordenschat iets minder goed presteren, maar daar zijn ook meer leerlingen die thuis een andere taal spreken en een hoger formatiegewicht hebben. Over het algemeen geldt: hoe minder stedelijk het gebied, hoe groter het percentage leerlingen dat thuis alleen Nederlands spreekt en hoe groter het percentage leerlingen zonder formatiegewicht.

Schoolgrootte en denominatie

De variabele schoolgrootte is sinds 2011 in de analyses meegenomen, waarbij een school tot 200 leerlingen gedefinieerd is als een kleine school en een school met 200 of meer leerlingen een grote school is. De richting van het effect was in bijna alle gevallen wel in het voordeel van de grotere scholen waarbij het gecorrigeerde effect positief was, maar in geen van de jaren is er voor een van de vaardigheden een betekenisvol effect uitgekomen. Door de jaren heen werd het kleinste effect altijd gevonden bij Woordenschat (gemiddeld $< ,01$), en het grootste bij Spelling (gemiddeld ongeveer 0,12). Alle andere effecten lagen hier tussenin.

De ongecorrigeerde effecten zijn iets groter (gemiddeld 0,04 effectpunt), maar nog steeds is geen van de effecten betekenisvol. De gevonden effecten voor de verschillende vaardigheden verschillen ook weinig met die van vorig jaar.

Er is in 2014 een klein onderzoek uitgevoerd naar het effect van de denominatie van de school. Daarbij is de opdeling in drie categorieën gebruikt: Rooms-Katholiek (36,1% van de leerlingen), Openbaar (30,6%), en Protestants-Christelijk (23,4%). Bijna 90% van de leerlingen zit op een van deze scholen. In dit onderzoek is geen onderscheid gemaakt tussen de verschillende rekenschalen. De gecorrigeerde effecten laten zien dat geen van de effecten betekenisvol is (zie tabel 3.18). Als we kijken naar de marginalen of de ongecorrigeerde effecten dan levert dat weinig verschil op wat betreft de leerlingen op de rooms-katholieke scholen versus die op de protestant-christelijke scholen. Bij de openbare scholen pakken de ongecorrigeerde effecten ten opzichte van de andere twee soorten negatief uit. Dit is o.a. te verklaren vanuit het feit dat openbare scholen relatief veel leerlingen met een formatiegewicht hebben. Als daar niet voor wordt gecorrigeerd is er dus wel degelijk een verschil.

Tabel 3.18 Effecten voor de variabele denominatie van de school

Contrast	Taal			Rekenen*
	WS	Sp	BL	
Openbaar – Rooms K.	0,01	-0,02	-0,02	-0,07
Protestants Chr. – Rooms K.	-0,01	-0,03	-0,05	-0,09
Protestants Chr. – Openbaar	-0,02	-0,02	-0,03	-0,02

* Ongedeelde schaal voor rekenen

** Alle vetgedrukte verschillen zijn significant ($p=0,01$)

4 Conclusies

4.1 Algemeen beeld

Bij een vergelijking van de resultaten van 2014 met die van 2013 blijken er nauwelijks verschillen te zijn. De uitkomsten voor de verschillende taalonderwerpen zijn sinds 2011 sterk naar elkaar toegetrokken. Terwijl er in de periode 2009-2012 sterke wisselingen optraden bij de drie taalvaardigheden, met ups en downs, kunnen we nu constateren dat er in de laatste drie jaar een stabiel beeld is ontstaan. Daarbij liggen de waarden een paar punten hoger dan bij het begin van de metingen in 2008.

Voor de rekenonderwerpen zien we ook een stabilisatie, op een iets hoger niveau dan in 2008. Voor de drie onderwerpen Getallen en bewerkingen, Breuken, procenten en verhoudingen en Meten, meetkunde, tijd en geld geldt echter dat ze sinds het begin van de meting weinig onderlinge fluctuatie hebben vertoond. Wel zien we dat na een stijging in niveau in respectievelijk 2010 en 2012 er nu een stabilisatie heeft plaatsgevonden op een lager niveau, vergelijkbaar met dat van 2010.

Bij elkaar genomen blijkt dat in 2014 zowel taal als rekenen een stabiel en consistent beeld geven. De gemiddelde groei in vaardigheid van 2008 naar 2014 is ook sterk vergelijkbaar over de taal- en rekenschalen. Deze gemiddelde groei is per jaar gemiddeld iets minder dan 1% van de standaarddeviatie (0,01 effectgroottepunt). Vergeleken met 2008 ligt het vaardigheidsniveau bij de verschillende schalen een paar punten hoger, tussen 252 en 254, maar er zijn bij iedere vaardigheid eerder in de verslagperiode betere resultaten behaald.

4.2 Invloed van achtergrondvariabelen

Geslacht

Jongens in jaargroep 8 behalen hogere scores in Rekenen-Wiskunde en Woordenschat dan meisjes. Meisjes presteren beter in Spelling en Begrijpend lezen. In de loop van de tijd zien we de prestaties van jongens en meisjes dichterbij elkaar komen. Het verschil tussen jongens en meisjes worden dus kleiner op de verschillende vakken.

Leertijd

Vertraagde leerlingen blijken in alle gevallen een grote achterstand te hebben. De belangrijkste verschuiving die we hier waarnemen is echter niet een verandering in vaardigheid, maar een geleidelijke verlaging van het percentage vertraagde leerlingen op school. We zien dat in vergelijking met 2008 in plaats van iets meer dan 20% nu bijna 16% van de leerlingen vertraagd is. Opvallend is dat de afname van het aantal vertraagde leerlingen in de populatie niet heeft geleid tot een daling in de vaardigheid.

Formatiegewicht

In jaargroep 8 zien we dat leerlingen zonder formatiegewicht (0.00) het op alle vaardigheden beter doen dan leerlingen mét een formatiegewicht, laag (0.30) of hoog (1.20). Het gaat dan om kleine en matige effectgroottes. Woordenschat is de enige vaardigheid waarbij leerlingen met een laag formatiegewicht significant hoger scoren dan leerlingen met een hoog formatiegewicht. Er is dan sprake van een klein effect. Bij Spelling en Getallen en bewerkingen is het effect andersom: leerlingen met een hoog gewicht presteren beter dan leerlingen met een laag formatiegewicht.

Advies voortgezet onderwijs

Door de jaren heen is het niveau van de leerlingen die geadviseerd worden op specifieke niveaus van het voortgezet in te stromen niet fundamenteel gewijzigd. Het verschil in vaardigheid verschilt enorm tussen

leerlingen met de verschillende doorstroomadviezen. Gemiddeld is het verschil tussen de vaardigheid van een vmbo-BB leerling en een vwo-leerling meer dan vier standaarddeviaties. Het geeft aan hoe groot de vaardigheidsverschillen aan het eind van de basisschool al zijn. Voor Woordenschat en Spelling zijn de verschillen nog het kleinst, nl. circa 3,5 standaarddeviaties. Vooral het verschil tussen de leerlingen met adviezen vwo en havo levert een grote bijdrage aan de grootte van het totaaleffect.

Thuis taal

We zien dat thuis taal bij Spelling en de rekenvaardigheden geen toegevoegd effect heeft naast formatiegewicht, stratum, leertijd en geslacht. Bij Begrijpend lezen is dat effect er wel en nog meer bij Woordenschat. Over het algemeen geldt voor deze vaardigheden: hoe meer Nederlands de leerling thuis spreekt, hoe hoger de vaardigheid.

Stratum

Deze variabele is geconstrueerd op basis de verdeling in formatiegewichten van de leerlingen op school. De verschillen en effecten zijn dan ook grotendeels vergelijkbaar met die bij de variabele formatiegewicht. In jaargroep 8 zijn de verschillen nagenoeg overal significant, maar niet in alle gevallen met betekenisvolle effectgroottes. Die zien we vooral bij taalonderwerpen, waar leerlingen in stratum 1 hogere resultaten behalen dan leerlingen in stratum 3.

Regio en verstedelijking

De variabelen regio en verstedelijking hebben geen toegevoegd effect, naast de variabelen formatiegewicht, leertijd, geslacht en stratum. Betekenisvolle effecten zijn er wel voor de ongecorrigeerde vergelijking. In zeer stedelijke gebieden zijn de prestaties voor Woordenschat bijvoorbeeld minder hoog, maar daar zijn in verhouding ook meer leerlingen die thuis geen Nederlands spreken en leerlingen met een hoog formatiegewicht.

Schoolgrootte en denominatie

Deze twee variabelen leveren geen betekenisvolle effecten op. Grotere scholen lijken het beter te doen, maar het effect is niet van betekenis. Voor de denominatie van de school geldt hetzelfde, maar daar zien we bij de ongecorrigeerde vergelijking wel een negatief effect voor openbare scholen. Deze scholen hebben relatief meer leerlingen met een formatiegewicht.

4.3 Discussie

De resultaten voor de gehele populatie van 2014 waren sterk vergelijkbaar met die van 2013. Alleen bij Woordenschat waren de prestaties iets minder goed. Waar in het verleden de verschillende vaardigheden nog wel eens grotere onderlinge verschillen vertoonden, is dat de laatste jaren ook minder het geval. Dit heeft bij alle vaardigheden jaren opgeleverd met piekprestaties, maar ook met mindere prestaties. Bij elkaar genomen blijkt dat in 2014 zowel taal als rekenen een stabiel en consistent beeld geven. De trendlijn door de jaren heen is in de afgelopen jaren in totaal positief te noemen, waarbij de groei gemiddeld over vaardigheden en jaren iets minder dan 1% van de standaarddeviatie per jaar is. Op zich lijkt een dergelijke groei niet sensationeel, maar over de jaren en vaardigheden heen is dit toch een acceptabel resultaat.

Dat de laatste jaren de resultaten wat stabiel zijn, kan de vraag oproepen of hier sprake is van stabilisatie of van stagnatie. De initiatieven om het reken- en taalonderwijs te stimuleren zijn niet gestopt. Het is dus de vraag of er een plafond is bereikt. De resultaten die in het verleden werden behaald suggereren dat er meer mogelijk is. Positief is echter wel dat het erop lijkt dat het onderwijs efficiënter is geworden. Het percentages dublures daalt, terwijl het percentage voorlijke leerlingen juist stijgt zonder dat dit ten koste lijkt te gaan van de vaardigheid van de leerlingen – sterker nog, deze is vanaf 2008 zelfs licht gestegen.

Hoewel deze evaluatie gestart is in het kader van de Kwaliteitsagenda van het Ministerie van OC&W (2008), waarin o.a. het opbrengstgericht werken wordt gestimuleerd, vooral bij taal en rekenen, kan op grond van de uitkomsten van dit onderzoek slechts ten dele een adequate beleidsevaluatie worden gedaan. Voor het vinden van verklaringen zijn gegevens nodig over wat scholen precies in hun onderwijs hebben gedaan op het gebied van de gemeten vaardigheden. Omdat deze gegevens ontbreken kunnen we dus niet vaststellen in hoeverre hierin wijzigingen zijn opgetreden in de loop der jaren.

Het verdient daarom aanbeveling om, net als bij het regulier peilingsonderzoek (PPON) gebruikelijk is, het onderwijsaanbod of veranderingen daarin te inventariseren zodanig dat gegevens over de leerlingprestaties daaraan gekoppeld kunnen worden. Met behulp van informatie over bijvoorbeeld het gebruik van andere methoden of het besteden van meer tijd aan bepaalde onderwerpen kan wellicht een plausibele interpretatie worden gegeven voor de verandering in resultaten bij leerlingen.

Literatuur

Berkel, S. van, M. Hilde, R. Engelen, F. Kamphuis, F. Kleintjes, R. Krom (2010). *Woordenschat Groep 3 t/m 5. Wetenschappelijke verantwoording*. Arnhem: Cito.

Boxtel, H. van, R. Engelen, A. de Wijs (2012). *Wetenschappelijke verantwoording van de Eindtoets Basisonderwijs 2010*. Arnhem: Cito.

Cito (2012). *Terugblik en resultaten 2012. Eindtoets basisonderwijs jaargroep 8*. Arnhem: Cito.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (second ed.)*. Lawrence Erlbaum Associates.

Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008). *Over de drempels met taal en rekenen*. Enschede: SLO.

Feenstra, H., F. Kamphuis, F. Kleintjes, R. Krom (2010). *Wetenschappelijke verantwoording Begrijpend lezen voor groep 3 tot en met 6*. Arnhem: Cito.

Hemker, B.T. (2012). *The impact of motivation: modeling motivation in educational measurement. Presentation presented July 4, 2012 at the ITC conference, Amsterdam*.

Hemker, B.T., J. Kordes & J.J. van Weerden (2011). *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2010 - Jaarlijks Peilingsonderzoek van het Onderwijsniveau*. Arnhem: Cito Arnhem. (zie www.cito.nl, Onderzoek en wetenschap, PPON)

Hemker, B.T., J.B. Kuhlemeier & J.J. van Weerden (2010). *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2009 - Jaarlijks Peilingsonderzoek van het Onderwijsniveau*. Arnhem: Cito. (zie www.cito.nl, Onderzoek en wetenschap, PPON)

Hemker, B.T. & J.J. van Weerden (2009). *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2008 - Jaarlijks Peilingsonderzoek van het Onderwijsniveau - Technische rapportage*. Arnhem: Cito. (<http://www.minocw.nl/documenten/133682d.pdf>)

Hemker, B.T. & J.J. van Weerden (2012). *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2011 - Jaarlijks Peilingsonderzoek van het Onderwijsniveau - Technische rapportage*. Arnhem: Cito. (zie www.cito.nl, Onderzoek en wetenschap, PPON)

Inspectie van het onderwijs (2012). *Monitor verbetertrajecten taal en rekenen 2008/2009, 2009/2010 en 2010/2011*. Utrecht: Inspectie van het Onderwijs/OCW.

Janssen, Jan, Frank van der Schoot, Bas Hemker (2005). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 4. Uitkomst van de vierde peiling in 2004*. PPON-reeks nummer 32. Arnhem: Cito. (http://www.cito.nl/po/ppon/rekwisk/eind_fr.htm)

Janssen, J., N. Verhelst, R. Engelen en F. Scheltens (2010). *Wetenschappelijke verantwoording van de toetsen LOVS Rekenen-Wiskunde voor groep 3 tot en met 8*. Arnhem: Cito.

Scheltens, F., Hemker, B.T., & Vermeulen, J. (2013). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 5. Uitkomst van de vijfde peiling in 2011*. Arnhem: Cito.

Van Weerden, J.J., B.T. Hemker, H. Straat & K. Mulder (2013). *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2012 - Jaarlijks Peilingsonderzoek van het Onderwijsniveau*. Arnhem: Cito (zie www.cito.nl, Onderzoek en wetenschap, PPON).

Verhelst, N.D. (1993). Itemresponstheorie. In: T.J.H.M. Eggen & P.F. Sanders (red.). *Psychometrie in de praktijk*. Arnhem: Cito (p. 83-178).

Verhelst, N.D., C.A.W. Glas & H.H.F.M. Verstralen (1995). *OPLM: One Parameter Logistic Model. Computer program and manual*. Arnhem: CITO.

Verhelst, N.D. & H.H.F.M. Verstralen (2002). *Structural Analysis of a Univariate Latent Variable (SAUL); Theory and a Computer Program*. Arnhem: Cito.

Wijs, A. de, F. Kamphuis, F. Kleintjes, M. Tomesen (2010). *Wetenschappelijke verantwoording Spelling voor groep 3 tot en met 6*. Arnhem: Cito.

Relevante websites:

www.cito.nl
www.toetswijzer.nl
<http://ppon.cito.nl>
www.minocw.nl
www.owinsp.nl

Bijlagen

Bijlage 1 Gemiddelden en standaarddeviaties per vaardigheid gecategoriseerd naar achtergrondvariabele in jaargroep 8 in 2014

Variabele	Vaardigheid Categorie	Taal		Rekenen-Wiskunde*									
		Woorden- schat	Spelling	Begrijpend lezen	G/B	BPV	MMTG						
		Gem.	SD	Gem.	SD	Gem.	SD	Gem.	SD	Gem.	SD	Gem.	SD
Jaar	2014	253	50	254	50	252	50	252	50	254	50	254	50
Formatiegewicht	F 0.00	257	48	256	50	257	48	254	49	257	49	257	49
	F 0.30	225	49	233	50	219	49	228	50	227	50	226	50
	F 1.20	204	49	241	51	210	49	233	50	229	50	226	50
Geslacht	Jongens	254	50	247	50	250	50	260	49	263	49	261	50
	Meisjes	252	50	261	49	255	50	244	49	245	49	248	49
Leertijd	Regulier	257	49	259	48	258	48	257	48	260	48	260	48
	Vertraagd	227	50	224	48	223	49	223	49	224	49	224	49
Stratum	Stratum (lovs*) 1	260	48	257	50	260	48	256	49	259	49	259	49
	Stratum (lovs*) 2	249	49	248	50	246	49	247	50	249	50	249	50
	Stratum (lovs*) 3	222	51	247	51	224	51	240	51	238	51	236	51
Soort toets	Eindtoets Basis	261	46	261	46	263	44	262	44	264	44	264	44
	Eindtoets Niveau	207	47	210	47	196	45	197	45	199	45	199	45
Advies	vmbo-BB	188	35	195	37	174	30	179	32	180	31	178	31
voortgezet	vmbo-KB	217	35	219	36	212	30	213	32	214	31	215	31
onderwijs	vmbo-GT	232	34	233	36	232	30	231	32	233	31	232	31
	havo	265	34	264	36	267	30	266	32	269	31	269	31
	vwo	314	34	313	36	316	30	313	32	317	31	317	31
	Regio	Noord	253	49	251	50	250	49	248	50	250	50	251
	Oost	252	49	251	50	251	49	250	50	252	50	252	50
	West	253	51	254	50	252	50	252	50	254	50	254	50
	Zuid	254	50	257	50	256	50	256	50	258	50	258	50
Urbanisatiegraad	Zeer sterk	245	53	257	50	246	52	251	51	251	51	250	51
	Sterk	252	50	254	50	252	50	251	50	253	50	253	50
	Matig	256	49	254	50	255	50	254	50	256	50	256	50
	Weinig	254	49	252	50	254	49	252	50	255	50	255	50
	Niet	256	49	251	50	253	49	251	50	254	49	256	49
Thuis taal - algemeen	Alleen Nederlands	257	48	254	50	255	49	253	50	255	50	256	50
	Nederlands+Anders	223	51	252	51	224	51	244	51	241	51	241	51
	Alleen een andere taal	210	51	249	51	220	51	243	51	240	51	236	51
Thuis taal - detail	Alleen Nederlands	257	48	254	50	255	49	253	50	255	50	256	50
	Turks (+)	192	50	243	51	206	51	242	51	236	51	234	51
	Arabisch (+)	211	50	248	51	219	51	240	51	240	51	235	51
	Surinaams/Antilliaans (+)	212	50	232	51	203	50	216	51	215	52	214	51
	Overige andere taal (+)	227	50	258	51	235	51	251	51	249	52	245	51
I*	Langer dan 4 jr in NL	253	50	254	50	253	50	252	50	254	50	254	50
	Korter dan 4 jr in NL	218	55	247	51	235	54	244	52	247	53	245	53
IJK*	Geen opmerking	257	48	257	48	257	47	256	47	258	47	258	47
	Opmerking bij leerling*	202	50	207	49	190	49	194	49	196	49	196	49
Schoolgrootte	Klein (1-200 ln)	250	50	248	50	247	50	248	50	251	50	251	50
	Groot (>200 ln)	254	50	256	50	255	50	253	50	256	50	256	50

* G/B = Getallen en bewerkingen; BPV = Breuken, procenten, verhoudingen; MMTG = Meten, meetkunde, tijd en geld.

** I of IJK = Opmerking bij een leerling: I=korter dan 4 jaar in NL, J=gaat naar VSO of PrO; K=krijgt LWOO

Bijlage 2 Effectschattingen voor alle vaardigheden gecategoriseerd naar achtergrondvariabelen in jaargroep 8 in 2014

Effecten 2014		Taalvaardigheden						Rekenvaardigheden					
		WS		Sp		BL		G/B		BPV		MMTG	
Variabele	Contrast	Eff.gr	z	Eff.gr	z	Eff.gr	z	Eff.gr	z	Eff.gr	z	Eff.gr	z
Formatiegewicht	F 0.00 – F 0.30	0,48	***	0,36	***	0,6	***	0,43	***	0,47	***	0,51	**
	F 0.30 – F 1.20	0,26	***	-0,22	***	0,01		-0,2	***	-0,15	**	-0,12	*
	F 0.00 – F 1.20	0,75	***	0,14	***	0,61	***	0,23	***	0,32	***	0,38	**
Geslacht	Meisjes - Jongens	-0,04	*	0,28	***	0,1	***	-0,37	***	-0,4	***	-0,3	**
Leertijd	Regulier - Vertraagd	0,52	***	0,7	***	0,62	***	0,7	***	0,71	***	0,7	**
Stratum	Stratum: Nieuw 1 – Nieuw 2	0,15	***	0,12	***	0,21	***	0,12	***	0,14	***	0,15	**
	Stratum: Nieuw 2 – Nieuw 3	0,33	***	-0,08	*	0,24	***	0,01		0,08	*	0,1	**
	Stratum: Nieuw 1 – Nieuw 3	0,48	***	0,03		0,45	***	0,12	***	0,22	***	0,24	**
Soort toets	Eindtoets Basis - Niveau	1	***	0,99	***	1,37	***	1,35	***	1,33	***	1,32	**
Advies VO	vmbo-KB - vmbo-BB	0,77	***	0,68	***	1,23	***	1,14	***	1,14	***	1,2	**
	vmbo-GT - vmbo-KB	0,43	***	0,39	***	0,67	***	0,59	***	0,61	***	0,57	**
	havo - vmbo-GT	0,94	***	0,86	***	1,16	***	1,13	***	1,19	***	1,2	**
	vwo - havo	1,45	***	1,41	***	1,64	***	1,53	***	1,59	***	1,58	**
	vwo - vmbo BB	3,59	***	3,34	***	4,7	***	4,39	***	4,54	***	4,54	**
Regio	Oost - Noord	0		0,01		0,03		0,05		0,05		0,03	
	West - Noord	0,08	*	0,09	**	0,12	***	0,1	***	0,12	***	0,1	**
	Zuid - Noord	0,06		0,14	***	0,15	***	0,18	***	0,19	***	0,16	**
	West - Oost	0,07	***	0,08	***	0,09	***	0,06	**	0,07	***	0,07	**
	Zuid - Oost	0,06	*	0,13	***	0,11	***	0,13	***	0,14	***	0,13	**
	Zuid - West	-0,02		0,05	*	0,03		0,08	***	0,06	**	0,06	*
Mate van Verstedelijking	Sterk - Zeer sterk	-0,04	***	-0,1		-0,05	**	-0,07	*	-0,07	*	-0,06	
	Matig - Sterk	0,03		-0,02		0,02		0,04		0,04		0,05	
	Weinig - Matig	-0,07	*	-0,05	*	-0,05		-0,07	**	-0,04		-0,05	
	Niet - Weinig	0,03		-0,02		-0,04		-0,01		-0,02		0,01	
Thuis taal - algemeen	Alleen NLS - NLS+Ander	0,39	***	-0,13		0,32	***	-0,02		0,05		0,03	
	NLS+Ander - Alleen Ander	0,18	*	0		-0,02		-0,03		-0,05		0,04	
	Alleen NLS - Alleen Ander	0,56	***	-0,13	***	0,31	***	-0,05		-0,01		0,07	
Thuis taal - detail: Alleen Nederlands tegenover –	Turks (+)	0,88	***	-0,03		0,52	***	-0,08		0,02		0,06	
	Arabisch (+)	0,5	***	-0,1		0,28	***	0		0		0,08	
	Surinaams/Antilliaans (+)	0,63	***	0,21		0,72	***	0,55	***	0,6	***	0,6	**
	Overige andere taal (+)	0,34	***	-0,26	***	0,14	**	-0,16	***	-0,11	*	-0,02	
I	Geen - Wel I code	0,47	***	-0,02		0,11		-0,07		-0,11		-0,06	
IJK	Geen - Wel IJK code	0,92	***	0,87	***	1,19	***	1,16	***	1,13	***	1,12	**
Schoolgrootte	Groot - Klein	0,03		0,13	***	0,09	***	0,08	***	0,06	***	0,06	**
Denominatie	Openbaar - RK	0,01		-0,02		-0,02		-0,07			***		
	PC - RK	-0,01		-0,03		-0,05	*	-0,09			***		
	PC -Openbaar	-0,02		-0,02		-0,03		-0,02					
	ABZ - RK	0,1	*	0,04		0,17	***	0,02					
	ABZ -Openbaar	0,09		0,05		0,19	***	0,09			*		
	ABZ - PC	0,1	*	0,07		0,22	***	0,11				**	

- geclassificeerd als "geen effect", maar wel significant: * : $0,01 < \alpha < 0,001$; ** : $0,001 < \alpha < 0,0001$; *** : $\alpha < 0,0001$
- Sp = Spelling; BL-A = Begrijpend lezen (anker); WS = Woordenschat; G/B = Getallen en bewerkingen; BPV = Breuken, procenten, verhoudingen; MMTG = Meten, meetkunde, tijd en geld.
- (+) bij Thuis taal - details betekent: de genoemde andere taal, al dan niet in combinatie met Nederlands
- Meest lichte kleur groen geeft de effecten weer die als klein worden geïnterpreteerd; de iets donkerder kleur groen geeft de effecten weer die als matig worden geïnterpreteerd en de donkerste kleur groen geeft de effecten weer die als groot worden geïnterpreteerd (zie tabel 2.1). Niet gearceerde effecten zijn in absolute waarde kleiner dan 0,20 en worden als niet relevant geïnterpreteerd.

Bijlage 3 De effecten per vaardigheid per jaar over de jaren heen

Effecten bij Woordenschat

Variabele	Contrast	2008/2009	2010	2011	2012	2013	2014
Geslacht	Meisjes - Jongens	-0,23	0,01	-0,35	0,00	-0,24	-0,04
Leertijd	Regulier - Vertraagd	0,54	0,44	0,47	0,50	0,49	0,52
Stratum	Stratum Oud: S1 - S2	0,08	0,07
	Stratum Nieuw: S1 - S2	.	0,15	0,15	0,12	0,11	0,15
	Stratum Oud: S2 - S3	0,15	0,42
	Stratum Nieuw: S2 - S3	.	0,34	0,30	0,28	0,37	0,33
	Stratum Oud: S1 - S3	0,23	0,48
	Stratum Nieuw: S1 - S3	.	0,48	0,45	0,40	0,47	0,48
Formatiegewicht	F 1.00 – F1.25 (08/09)	0,54/0,56
	F 0.00 – F 0.30	.	0,48	0,50	0,48	0,39	0,48
	F 1.25 – F1.90 (08/09)	0,51/0,20
	F 0.30 – F1.20	.	0,27	0,32	0,28	0,42	0,26
	F 1.00 – F1.90 (08/09)	1,05/0,76
	F 0.00 – F1.20	.	0,76	0,82	0,76	0,81	0,75
Soort toets	EB - NT	1,42	1,21	1,19	1,10	1,13	1,00
Advies VO	vmbo-KB - vmbo-BB	0,64	0,51	0,52	0,65	0,51	0,62
	vmbo-GT - vmbo-KB	0,46	0,41	0,44	0,37	0,44	0,34
	havo - vmbo-GT	0,77	0,71	0,69	0,79	0,74	0,80
	vwo - havo	1,24	1,15	1,27	1,17	1,18	1,27
	vwo - vmbo-BB	3,11	2,78	2,92	2,98	2,87	3,03
Advies VO (alternatief)	vmbo-KB - vmbo-BB	.	.	.	0,80	0,63	0,77
	vmbo-GT - vmbo-KB	.	.	.	0,47	0,53	0,43
	havo - vmbo-GT	.	.	.	0,93	0,85	0,94
	vwo - havo	.	.	.	1,34	1,32	1,45
	vwo- vmbo BB	.	.	.	3,53	3,31	3,59
Thuisstaal	Alleen NLs - NLs+Ander	0,28	0,33	0,42	.	.	0,39
	Alleen Ander - NLs+Ander	0,19	0,33	0,18	.	.	0,18
	Alleen NLs - Alleen Ander	0,47	0,66	0,60	.	.	0,56
Regio	Oost - Noord	-0,02	-0,07	-0,11	-0,05	-0,10	0,00
	West - Noord	0,07	0,00	0,00	0,00	-0,06	0,08
	Zuid - Noord	0,03	-0,03	-0,12	-0,07	-0,13	0,06
	West - Oost	0,09	0,07	0,10	0,05	0,05	0,07
	Zuid - Oost	0,05	0,04	-0,01	-0,02	-0,02	0,06
	Zuid - West	-0,04	-0,03	-0,11	-0,07	-0,07	-0,02
Urbanisatiegraad	Sterk - Zeer sterk	-0,04	0,03	0,11	-0,01	-0,04	-0,04
	Matig - Sterk	-0,01	-0,05	-0,01	0,00	0,02	0,03
	Weinig - Matig	-0,06	-0,04	-0,03	-0,01	-0,04	-0,07
	Niet - Weinig	-0,04	-0,10	-0,06	0,00	0,02	0,03
	Niet - Zeer Sterk	-0,15	-0,16	0,01	-0,02	-0,04	-0,05
Schoolgrootte	Groot - Klein	.	.	0,00	-0,03	0,01	0,03

Effecten bij Spelling

Variabele	Contrast	2008/2009	2010	2011	2012	2013	2014
Geslacht	Meisjes - Jongens	0,33	0,33	0,23	0,25	0,27	0,28
Leertijd	Regulier - Vertraagd	0,69	0,71	0,68	0,69	0,72	0,70
Stratum	Stratum Oud: S1 - S2	0,06	0,04
	Stratum Nieuw: S1 - S2	.	0,11	0,08	0,07	0,10	0,12
	Stratum Oud: S2 - S3	0,02	0,01
	Stratum Nieuw: S2 - S3	.	-0,04	-0,08	0,05	0,02	-0,08
	Stratum Oud: S1 - S3	0,07	0,04
	Stratum Nieuw: S1 - S3	.	0,07	0,00	0,12	0,12	0,03
Formatiegewicht	F 1.00 – F1.25 (08/09)	0,41/0,38
	F 0.00 – F 0.30	.	0,31	0,36	0,33	0,36	0,36
	F 1.25 – F1.90 (08/09)	-0,22/-0,36
	F 0.30 – F1.20	.	-0,14	-0,21	-0,27	-0,27	-0,22
	F 1.00 – F1.90 (08/09)	0,19/ 0,02	.	.	0,00	.	.
	F 0.00 – F1.20	.	0,17	0,15	0,06	0,10	0,14
Soort toets	EB - NT	1,04	1,04	0,91	0,79	1,14	0,99
Advies VO	vmbo-KB - vmbo-BB	0,47	0,43	0,37	0,44	0,53	0,55
	vmbo-GT - vmbo-KB	0,39	0,28	0,39	0,34	0,39	0,30
	havo - vmbo-GT	0,77	0,72	0,80	0,78	0,81	0,74
	vwo - havo	1,45	1,36	1,41	1,36	1,31	1,25
	vwo - vmbo-BB	3,08	2,79	2,97	2,92	3,04	2,84
Advies VO (alternatief)	vmbo-KB - vmbo-BB	.	.	.	0,54	0,67	0,68
	vmbo-GT - vmbo-KB	.	.	.	0,42	0,49	0,39
	havo - vmbo-GT	.	.	.	0,90	0,94	0,86
	vwo - havo	.	.	.	1,51	1,49	1,41
	vwo - vmbo-BB	.	.	.	3,38	3,58	3,34
Thuistaal	Alleen NLs - NLs+Ander	-0,10	-0,18	-0,18	.	.	-0,13
	Alleen Ander - NLs+Ander	0,05	0,03	0,11	.	.	0,00
	Alleen NLs - Alleen Ander	-0,05	-0,15	-0,08	.	.	-0,13
Regio	Oost - Noord	0,05	-0,07	-0,01	-0,01	-0,01	0,01
	West - Noord	0,06	-0,10	0,07	0,06	0,12	0,09
	Zuid - Noord	0,15	-0,08	0,05	0,11	0,18	0,14
	West - Oost	0,01	-0,03	0,09	0,07	0,13	0,08
	Zuid - Oost	0,10	0,00	0,07	0,12	0,19	0,13
	Zuid - West	0,09	0,03	-0,02	0,05	0,06	0,05
Urbanisatiegraad	Sterk - Zeer sterk	-0,09	-0,02	-0,07	-0,11	-0,06	-0,10
	Matig - Sterk	-0,01	-0,06	-0,01	0,01	-0,05	-0,02
	Weinig - Matig	-0,07	-0,01	-0,04	-0,05	-0,04	-0,05
	Niet - Weinig	-0,04	-0,04	-0,04	-0,06	-0,05	-0,02
	Niet - Zeer Sterk	-0,21	-0,13	-0,16	-0,21	-0,20	-0,19
Schoolgrootte	Groot - Klein	.	.	0,12	0,13	0,11	0,13

Effecten bij Begrijpend lezen

Variabele	Contrast	2008/2009	2010	2011	2012	2013	2014
Geslacht	Meisjes - Jongens	0,16	0,18	0,20	0,12	0,15	0,10
Leertijd	Regulier - Vertraagd	0,60	0,60	0,58	0,60	0,58	0,62
Stratum	Stratum Oud: S1 - S2	0,11	0,10
	Stratum Nieuw: S1 - S2	.	0,15	0,16	0,17	0,19	0,21
	Stratum Oud: S2 - S3	0,15	0,28
	Stratum Nieuw: S2 - S3	.	0,24	0,22	0,29	0,37	0,24
	Stratum Oud: S1 - S3	0,27	0,38
	Stratum Nieuw: S1 - S3	.	0,39	0,38	0,45	0,56	0,45
Formatiegewicht	F 1.00 – F1.25 (08/09)	0,55/ 0,61
	F 0.00 – F 0.30	.	0,57	0,56	0,59	0,52	0,60
	F 1.25 – F1.90 (08/09)	-0,02/ 0,08
	F 0.30 – F1.20	.	0,07	0,02	-0,03	0,17	0,01
	F 1.00 – F1.90 (08/09)	0,54/ 0,69
	F 0.00 – F1.20	.	0,64	0,59	0,57	0,69	0,61
Soort toets	EB - NT	1,60	1,42	1,55	1,45	1,41	1,37
Advies VO	vmbo-KB - vmbo-BB	0,77	0,72	0,87	0,85	0,84	0,93
	vmbo-GT - vmbo-KB	0,60	0,50	0,50	0,54	0,56	0,50
	havo - vmbo-GT	0,95	0,96	0,94	1,07	0,93	0,91
	vwo - havo	1,33	1,36	1,33	1,23	1,45	1,32
	vwo - vmbo-BB	3,65	3,54	3,64	3,69	3,77	3,66
Advies VO (alternatief)	vmbo-KB - vmbo-BB	.	.	.	1,09	1,09	1,23
	vmbo-GT - vmbo-KB	.	.	.	0,71	0,74	0,67
	havo - vmbo-GT	.	.	.	1,34	1,15	1,16
	vwo - havo	.	.	.	1,52	1,75	1,64
	vwo - vmbo-BB	.	.	.	4,66	4,73	4,70
Thuistaal	Alleen NLs - NLs+Ander	0,17	0,14	0,24	.	.	0,32
	Alleen Ander - NLs+Ander	0,06	0,16	0,07	.	.	-0,02
	Alleen NLs - Alleen Ander	0,23	0,30	0,31	.	.	0,31
Regio	Oost - Noord	0,06	0,06	0,04	0,10	0,04	0,03
	West - Noord	0,11	0,13	0,08	0,13	0,09	0,12
	Zuid - Noord	0,14	0,11	0,08	0,14	0,09	0,15
	West - Oost	0,05	0,06	0,04	0,03	0,05	0,09
	Zuid - Oost	0,07	0,05	0,04	0,04	0,05	0,11
	Zuid - West	0,03	-0,02	0,00	0,01	0,00	0,03
Urbanisatiegraad	Sterk - Zeer sterk	-0,06	-0,01	-0,01	-0,01	-0,02	-0,05
	Matig - Sterk	-0,01	-0,02	-0,01	0,02	-0,02	0,02
	Weinig - Matig	-0,07	-0,05	-0,03	-0,02	-0,06	-0,05
	Niet - Weinig	-0,06	-0,07	-0,04	-0,04	-0,06	-0,04
	Niet - Zeer Sterk	-0,20	-0,15	-0,09	-0,05	-0,16	-0,12
Schoolgrootte	Groot - Klein	.	.	0,04	0,07	0,07	0,09

Effecten bij Rekenen met getallen en bewerkingen

Variabele	Contrast	2008/2009	2010	2011	2012	2013	2014
Geslacht	Meisjes - Jongens	-0,43	-0,40	-0,41	-0,43	-0,30	-0,37
Leertijd	Regulier - Vertraagd	0,67	0,70	0,67	0,67	0,67	0,70
Stratum	Stratum Oud: S1 - S2	0,06	0,05
	Stratum Nieuw: S1 - S2	.	0,12	0,08	0,09	0,10	0,12
	Stratum Oud: S2 - S3	0,02	0,06
	Stratum Nieuw: S2 - S3	.	0,03	0,04	0,04	0,05	0,01
	Stratum Oud: S1 - S3	0,08	0,12
	Stratum Nieuw: S1 - S3	.	0,15	0,12	0,14	0,14	0,12
Formatiegewicht	F 1.00 – F1.25 (08/09)	0,48
	F 0.00 – F 0.30	.	0,47	0,46	0,49	0,44	0,43
	F 1.25 – F1.90 (08/09)	-0,24
	F 0.30 – F1.20	.	-0,14	-0,17	-0,24	-0,24	-0,20
	F 1.00 – F1.90 (08/09)	0,24
	F 0.00 – F1.20	.	0,34	0,28	0,27	0,20	0,23
Soort toets	EB - NT	1,75	1,77	1,82	1,86	1,39	1,35
Advies VO	vmbo-KB - vmbo-BB	0,86	0,80	0,94	0,93	0,78	0,89
	vmbo-GT - vmbo-KB	0,61	0,54	0,55	0,86	0,53	0,44
	havo - vmbo-GT	0,93	1,03	0,97	0,98	0,97	0,91
	vwo - havo	1,27	1,26	1,20	1,20	1,28	1,27
	vwo - vmbo-BB	3,67	3,63	3,66	3,97	3,57	3,51
Advies VO (alternatief)	vmbo-KB - vmbo-BB	.	.	.	1,17	1,02	1,14
	vmbo-GT - vmbo-KB	.	.	.	0,65	0,70	0,59
	havo - vmbo-GT	.	.	.	1,20	1,20	1,13
	vwo - havo	.	.	.	1,44	1,56	1,53
	vwo - vmbo-BB	.	.	.	4,46	4,48	4,39
Thuistaal	Alleen NLs - NLs+Ander	-0,06	-0,02	-0,05	.	.	-0,02
	Alleen Ander - NLs+Ander	-0,03	0,00	0,04	.	.	-0,03
	Alleen NLs - Alleen Ander	-0,10	-0,02	-0,01	.	.	-0,05
Regio	Oost - Noord	0,11	0,04	0,03	0,04	0,00	0,05
	West - Noord	0,14	0,09	0,07	0,09	0,08	0,10
	Zuid - Noord	0,24	0,14	0,11	0,11	0,14	0,18
	West - Oost	0,03	0,05	0,04	0,05	0,08	0,06
	Zuid - Oost	0,12	0,10	0,08	0,06	0,14	0,13
	Zuid - West	0,09	0,05	0,04	0,01	0,06	0,08
Urbanisatiegraad	Sterk - Zeer sterk	-0,07	0,00	0,00	-0,03	-0,05	-0,07
	Matig - Sterk	0,01	0,01	-0,01	0,01	-0,02	0,04
	Weinig - Matig	-0,04	-0,04	0,01	-0,03	-0,04	-0,07
	Niet - Weinig	-0,01	-0,01	-0,04	-0,05	0,00	-0,01
	Niet - Zeer Sterk	-0,11	-0,04	-0,04	-0,10	-0,11	-0,11
Schoolgrootte	Groot - Klein	.	.	0,06	0,08	0,08	0,08

Effecten bij Rekenen met breuken, procenten en verhoudingen

Variabele	Contrast	2008/2009	2010	2011	2012	2013	2014
Geslacht	Meisjes - Jongens	-0,51	-0,58	-0,48	-0,43	-0,46	-0,40
Leertijd	Regulier - Vertraagd	0,71	0,75	0,68	0,71	0,69	0,71
Stratum	Stratum Oud: S1 - S2	0,09	0,08
	Stratum Nieuw: S1 - S2	.	0,13	0,12	0,12	0,13	0,14
	Stratum Oud: S2 - S3	0,07	0,10
	Stratum Nieuw: S2 - S3	.	0,08	0,10	0,11	0,14	0,08
	Stratum Oud: S1 - S3	0,16	0,18
	Stratum Nieuw: S1 - S3	.	0,21	0,21	0,24	0,27	0,22
Formatiegewicht	F 1.00 – F1.25 (08/09)	0,50
	F 0.00 – F 0.30	.	0,48	0,48	0,56	0,47	0,47
	F 1.25 – F1.90 (08/09)	-0,15	.	0,00	.	.	.
	F 0.30 – F1.20	.	-0,10	-0,13	-0,20	-0,15	-0,15
	F 1.00 – F1.90 (08/09)	0,35
	F 0.00 – F1.20	.	0,38	0,36	0,36	0,33	0,32
Soort toets	EB - NT	1,76	1,64	1,56	1,62	1,45	1,33
Advies VO	vmbo-KB - vmbo-BB	0,83	0,78	0,82	0,89	0,87	0,88
	vmbo-GT - vmbo-KB	0,63	0,60	0,58	0,55	0,56	0,45
	havo - vmbo-GT	1,02	1,01	1,06	1,07	0,99	0,94
	vwo - havo	1,37	1,30	1,35	1,33	1,36	1,30
	vwo - vmbo-BB	3,85	3,69	3,81	3,84	3,78	3,57
Advies VO (alternatief)	vmbo-KB - vmbo-BB	.	.	.	1,14	1,13	1,14
	vmbo-GT - vmbo-KB	.	.	.	0,72	0,75	0,61
	havo - vmbo-GT	.	.	.	1,32	1,24	1,19
	vwo - havo	.	.	.	1,61	1,66	1,59
	vwo - vmbo-BB	.	.	.	4,79	4,77	4,54
Thuistaal	Alleen NLs - NLs+Ander	0,00	0,00	0,04	.	.	0,05
	Alleen Ander - NLs+Ander	-0,02	0,06	0,01	.	.	-0,05
	Alleen NLs - Alleen Ander	-0,03	0,06	0,04	.	.	-0,01
Regio	Oost - Noord	0,10	0,03	0,05	0,06	0,03	0,05
	West - Noord	0,13	0,08	0,07	0,09	0,11	0,12
	Zuid - Noord	0,19	0,11	0,11	0,09	0,13	0,19
	West - Oost	0,02	0,05	0,03	0,02	0,08	0,07
	Zuid - Oost	0,08	0,08	0,06	0,03	0,10	0,14
	Zuid - West	0,06	0,03	0,03	0,01	0,02	0,06
Urbanisatiegraad	Sterk - Zeer sterk	-0,07	-0,01	-0,01	-0,04	-0,06	-0,07
	Matig - Sterk	0,02	0,01	-0,02	0,01	-0,03	0,04
	Weinig - Matig	-0,04	-0,03	0,01	-0,02	-0,03	-0,04
	Niet - Weinig	-0,03	-0,05	-0,05	-0,05	0,00	-0,02
	Niet - Zeer Sterk	-0,12	-0,08	-0,07	-0,10	-0,12	-0,09
Schoolgrootte	Groot - Klein	.	.	0,06	0,07	0,09	0,06

Effecten bij Rekenen met Meten, meetkunde, tijd en geld

Variabele	Contrast	2008/2009	2010	2011	2012	2013	2014
Geslacht	Meisjes - Jongens	-0,51	-0,52	-0,55	-0,39	-0,44	-0,30
Leertijd	Regulier - Vertraagd	0,69	0,71	0,67	0,71	0,65	0,70
Stratum	Stratum Oud: S1 - S2	0,10	0,08
	Stratum Nieuw: S1 - S2	.	.	0,13	0,12	0,12	0,15
	Stratum Oud: S2 - S3	0,10	0,17
	Stratum Nieuw: S2 - S3	.	.	0,12	0,15	0,16	0,10
	Stratum Oud: S1 - S3	0,20	0,26
	Stratum Nieuw: S1 - S3	.	.	0,25	0,27	0,28	0,24
Formatiegewicht	F 1.00 – F1.25 (08/09)	0,52
	F 0.00 – F 0.30	.	0,48	0,46	0,56	0,46	0,51
	F 1.25 – F1.90 (08/09)	-0,16
	F 0.30 – F1.20	0,00	-0,03	-0,05	-0,12	-0,12	-0,12
	F 1.00 – F1.90 (08/09)	0,36
	F 0.00 – F1.20	0,00	0,45	0,41	0,43	0,34	0,38
Soort Toets	EB - NT	1,86	1,69	1,67	1,71	1,39	1,32
Advies VO	vmbo-KB - vmbo-BB	0,83	0,72	0,76	1,13	0,75	0,93
	vmbo-GT - vmbo-KB	0,58	0,59	0,59	0,67	0,53	0,41
	havo - vmbo-GT	1,03	1,06	1,03	1,36	1,00	0,95
	vwo - havo	1,51	1,41	1,45	1,83	1,46	1,29
	vwo - vmbo-BB	3,95	3,78	3,83	4,99	3,73	3,59
Advies VO (alternatief)	vmbo-KB - vmbo-BB	.	.	.	1,13	1,00	1,20
	vmbo-GT - vmbo-KB	.	.	.	0,67	0,71	0,57
	havo - vmbo-GT	.	.	.	1,36	1,25	1,20
	vwo - havo	.	.	.	1,83	1,78	1,58
	vwo - vmbo-BB	.	.	.	4,99	4,73	4,54
Thuistaal	Alleen NLs - NLs+Ander	0,02	0,01	0,06	.	.	0,03
	Alleen Ander - NLs+Ander	-0,06	0,08	0,03	.	.	0,04
	Alleen NLs - Alleen Ander	-0,03	0,09	0,09	.	.	0,07
Regio	Oost - Noord	0,09	-0,03	0,02	0,06	0,00	0,03
	West - Noord	0,14	0,04	0,06	0,09	0,06	0,10
	Zuid - Noord	0,22	0,07	0,08	0,10	0,08	0,16
	West - Oost	0,05	0,07	0,04	0,03	0,06	0,07
	Zuid - Oost	0,13	0,10	0,07	0,04	0,08	0,13
	Zuid - West	0,08	0,03	0,03	0,01	0,02	0,06
Urbanisatiegraad	Sterk - Zeer sterk	-0,02	0,00	0,02	-0,04	-0,02	-0,06
	Matig - Sterk	0,00	0,02	0,00	0,03	-0,02	0,05
	Weinig - Matig	-0,04	-0,03	0,03	-0,05	-0,02	-0,05
	Niet - Weinig	-0,01	-0,03	-0,05	0,01	0,02	0,01
	Niet - Zeer Sterk	-0,07	-0,04	0,00	-0,05	-0,04	-0,05
Schoolgrootte	Groot - Klein	.	.	0,03	0,05	0,04	0,06

Bijlage 4 De gemiddelde vaardigheid per vaardigheid per jaar over de jaren heen

Gemiddelde vaardigheid per jaar en per subgroep bij Woordenschat

Variabele	Niveau	2008	2009	2010	2011	2012	2013	2014
Jaar	.	250	249	257	251	252	254	253
Geslacht	Jongens	255	253	256	259	252	260	254
	Meisjes	245	244	258	244	252	249	252
Leertijd	Regulier	256	254	262	256	257	259	257
	Vertraagd	225	225	234	229	228	231	227
Stratum	Stratum (oud) 1	257	255	263
	Stratum (oud) 2	247	245	257
	Stratum (oud) 3	218	222	225
	Stratum (lovs*) 1	.	.	265	259	259	261	260
	Stratum (lovs*) 2	.	.	255	249	250	253	249
	Stratum (lovs*) 3	.	.	228	223	225	224	222
Formatiegewicht	F 1.00	259	256
	F 1.25	229	225
	F 1.90	199	210
	F 0.00	.	.	262	256	257	260	257
	F 0.30	.	.	232	225	227	233	225
	F 1.20	.	.	208	202	206	204	204
Soort toets	EB	251	250	258	253	254	263	261
	NT	172	174	187	182	187	204	207
Advies VO	vmbo-BB	182	183	193	187	185	192	188
	vmbo-KB	211	210	217	211	214	217	217
	vmbo-GT	233	230	237	232	231	236	232
	havo	265	262	268	261	264	267	265
	vwo	313	310	315	312	312	314	314
Thuis taal	Alleen NLS	255	253	262	256	.	.	257
	NLS + bu	220	223	231	223	.	.	223
	Alleen bu	201	208	209	207	.	.	210
Regio	Noord	250	248	259	255	256	261	253
	Oost	249	248	256	251	253	255	252
	West	250	249	256	251	251	253	253
	Zuid	251	249	258	250	251	253	254
Urbanisatiegraad	Zeer sterk	238	240	245	239	243	245	245
	Sterk	250	249	258	252	251	253	252
	Matig	254	252	260	255	254	257	256
	Weinig	252	250	259	254	254	257	254
	Niet	252	249	257	253	256	259	256
Schoolgrootte	Klein (<200 ln)	.	.	.	249	251	252	250
	Groot (200+ ln)	.	.	.	252	252	255	254

Gemiddelde vaardigheid per jaar en per subgroep bij Spelling

Variabele	Niveau	2008	2009	2010	2011	2012	2013	2014
Jaar	.	250	250	252	252	256	255	254
Geslacht	Jongens	242	241	243	245	249	248	247
	Meisjes	258	258	260	258	262	261	261
Leertijd	Regulier	257	256	258	258	262	261	259
	Vertraagd	221	221	222	223	227	224	224
Stratum	Stratum (oud) 1	253	252	254
	Stratum (oud) 2	247	247	250
	Stratum (oud) 3	239	243	243
	Stratum (lovs*) 1	.	.	262	254	259	258	257
	Stratum (lovs*) 2	.	.	252	248	253	251	248
	Stratum (lovs*) 3	.	.	229	246	247	245	247
Formatiegewicht	F 1.00	254	253
	F 1.25	230	231
	F 1.90	236	243
	F 0.00	.	.	254	254	258	257	256
	F 0.30	.	.	235	232	237	234	233
	F 1.20	.	.	237	240	246	242	241
Soort toets	EB	251	251	252	253	257	263	261
	NT	187	190	192	199	208	205	210
Advies VO	vmbo-BB	195	195	198	198	199	195	195
	vmbo-KB	215	215	217	214	218	217	219
	vmbo-GT	231	230	230	230	232	234	233
	havo	262	261	260	263	265	266	264
	vwo	318	316	316	318	318	317	313
Thuis taal	Alleen NLS	251	251	252	252	.	.	254
	NLS + bu	247	249	251	255	.	.	252
	Alleen bu	238	243	245	244	.	.	249
Regio	Noord	246	245	255	249	253	251	251
	Oost	249	249	252	249	253	250	251
	West	250	250	251	254	257	256	254
	Zuid	253	253	252	252	258	259	257
Urbanisatiegraad	Zeer sterk	248	250	249	254	257	255	257
	Sterk	251	251	253	252	256	256	254
	Matig	252	252	252	252	258	255	254
	Weinig	250	249	252	250	255	254	252
	Niet	247	246	251	249	253	252	251
Schoolgrootte	Klein (<200 ln)	.	.	.	247	251	250	248
	Groot (200+ ln)	.	.	.	253	258	256	256

Gemiddelde vaardigheid per jaar en per subgroep bij Begrijpend lezen

Variabele	Niveau	2008	2009	2010	2011	2012	2013	2014
Jaar	.	250	252	254	257	253	253	252
Geslacht	Jongens	246	248	249	252	250	249	250
	Meisjes	254	256	259	263	256	257	255
Leertijd	Regulier	257	258	260	263	259	258	258
	Vertraagd	223	224	226	230	225	225	223
Stratum	Stratum (oud) 1	256	259	260
	Stratum (oud) 2	246	248	252
	Stratum (oud) 3	225	224	227
	Stratum (lovs*) 1	.	.	255	265	261	261	260
	Stratum (lovs*) 2	.	.	248	254	249	248	246
	Stratum (lovs*) 3	.	.	244	233	226	221	224
Formatiegewicht	F 1.00	256	260
	F 1.25	224	225
	F 1.90	218	214
	F 0.00	.	.	259	262	259	258	257
	F 0.30	.	.	225	227	222	225	219
	F 1.20	.	.	213	219	215	207	210
Soort toets	EB	251	253	255	260	255	263	263
	NT	163	167	174	174	175	194	196
Advies VO	vmbo-BB	176	177	180	181	175	176	174
	vmbo-KB	207	208	209	215	209	210	212
	vmbo-GT	231	232	231	236	231	233	232
	havo	267	269	268	273	272	267	267
	vwo	316	318	320	322	317	319	316
Thuis taal	Alleen NLs	254	256	257	261	.	.	255
	NLs + bu	230	228	234	237	.	.	224
	Alleen bu	219	217	220	225	.	.	220
Regio	Noord	246	248	250	256	250	252	250
	Oost	249	252	254	258	254	253	251
	West	250	251	254	257	252	252	252
	Zuid	253	254	256	259	255	255	256
Urbanisatiegraad	Zeer sterk	242	243	244	249	244	244	246
	Sterk	251	253	255	258	253	253	252
	Matig	253	256	257	260	256	256	255
	Weinig	251	254	256	259	255	254	254
	Niet	248	251	253	258	254	253	253
Schoolgrootte	Klein (<200 ln)	.	.	.	254	249	248	247
	Groot (200+ ln)	.	.	.	259	255	255	255

Gemiddelde vaardigheid per jaar en per subgroep bij Rekenen met Getallen en bewerkingen

Variabele	Niveau	2008	2009	2010	2011	2012	2013	2014
Jaar	.	250	250	252	252	255	252	252
Geslacht	Jongens	259	259	261	261	265	259	260
	Meisjes	241	241	243	243	245	246	244
Leertijd	Regulier	257	256	259	258	261	258	257
	Vertraagd	224	224	224	225	228	225	223
Stratum	Stratum (oud) 1	254	253	255
	Stratum (oud) 2	246	246	250
	Stratum (oud) 3	238	238	238
	Stratum (lovs*) 1	.	.	257	256	259	256	256
	Stratum (lovs*) 2	.	.	249	249	251	249	247
	Stratum (lovs*) 3	.	.	239	240	243	240	240
Formatiegewicht	F 1.00	255	254
	F 1.25	227	226
	F 1.90	234	234
	F 0.00	.	.	256	255	259	256	254
	F 0.30	.	.	228	227	229	228	228
	F 1.20	.	.	229	232	236	235	233
Soort toets	EB	252	251	253	254	258	262	262
	NT	157	159	158	158	161	195	197
Advies VO	vmbo-BB	177	176	177	175	179	180	179
	vmbo-KB	209	208	209	210	214	211	213
	vmbo-GT	232	231	229	232	234	232	231
	havo	268	266	269	269	271	268	266
	vwo	314	313	315	314	316	315	313
Thuis taal	Alleen NLS	252	251	254	253	.	.	253
	NLS + bu	244	244	243	247	.	.	244
	Alleen bu	239	239	238	239	.	.	243
Regio	Noord	244	243	248	250	252	250	248
	Oost	249	249	251	251	254	249	250
	West	250	250	252	251	256	253	252
	Zuid	254	254	255	255	257	255	256
Urbanisatiegraad	Zeer sterk	247	247	246	247	252	251	251
	Sterk	250	250	252	252	256	252	251
	Matig	252	252	254	253	257	253	254
	Weinig	251	251	253	253	255	252	252
	Niet	249	249	252	252	253	252	251
Schoolgrootte	Klein (<200 ln)	.	.	.	249	251	248	248
	Groot (200+ ln)	.	.	.	253	257	254	253

Gemiddelde vaardigheid per jaar en per subgroep bij Rekenen met Breuken, procenten en verhoudingen

Variabele	Niveau	2008	2009	2010	2011	2012	2013	2014
Jaar	.	250	250	254	253	257	254	254
Geslacht	Jongens	261	261	266	264	266	264	263
	Meisjes	239	239	241	243	248	244	245
Leertijd	Regulier	257	256	260	259	263	260	260
	Vertraagd	222	222	225	226	228	225	224
Stratum	Stratum (oud) 1	255	255	258
	Stratum (oud) 2	246	246	251
	Stratum (oud) 3	233	233	237
	Stratum (lovs*) 1	.	.	259	259	263	259	259
	Stratum (lovs*) 2	.	.	251	250	253	250	249
	Stratum (lovs*) 3	.	.	238	237	240	235	238
Formatiegewicht	F 1.00	256	255
	F 1.25	226	226
	F 1.90	228	228
	F 0.00	.	.	257	257	261	258	257
	F 0.30	.	.	229	227	228	227	227
	F 1.20	.	.	227	228	232	229	229
Soort toets	EB	252	251	255	255	259	264	264
	NT	157	159	165	170	173	195	199
Advies VO	vmbo-BB	177	176	178	179	180	179	180
	vmbo-KB	207	207	209	210	213	212	214
	vmbo-GT	231	230	231	232	235	233	233
	havo	269	267	269	271	274	269	269
	vwo	317	316	316	320	322	318	317
Thuis taal	Alleen NLs	252	252	256	256	.	.	255
	NLs + bu	239	239	243	244	.	.	241
	Alleen bu	233	233	235	237	.	.	240
Regio	Noord	245	245	250	251	254	252	250
	Oost	249	250	253	253	257	252	252
	West	250	250	253	253	257	254	254
	Zuid	253	253	256	256	258	256	258
Urbanisatiegraad	Zeer sterk	244	245	247	248	252	251	251
	Sterk	249	249	253	254	257	254	253
	Matig	252	252	256	255	259	255	256
	Weinig	252	251	255	255	258	254	255
	Niet	250	249	253	254	256	255	254
Schoolgrootte	Klein (<200 ln)	.	.	.	250	253	249	251
	Groot (200+ ln)	.	.	.	255	259	256	256

Gemiddelde vaardigheid per jaar en per subgroep bij Rekenen met Meten, meetkunde, tijd en geld

Variabele	Niveau	2008	2009	2010	2011	2012	2013	2014
Jaar	.	250	249	254	254	256	255	254
Geslacht	Jongens	261	260	265	266	265	265	261
	Meisjes	239	239	243	242	248	245	248
Leertijd	Regulier	257	256	260	259	262	260	260
	Vertraagd	223	222	226	227	226	228	224
Stratum	Stratum (oud) 1	255	254	259
	Stratum (oud) 2	246	245	251
	Stratum (oud) 3	231	231	234
	Stratum (lovs*) 1	.	.	260	259	262	260	259
	Stratum (lovs*) 2	.	.	251	250	252	251	249
	Stratum (lovs*) 3	.	.	235	236	237	236	236
Formatiegewicht	F 1.00	256	255
	F 1.25	225	224
	F 1.90	226	226
	F 0.00	.	.	258	257	261	259	257
	F 0.30	.	.	229	228	227	229	226
	F 1.20	.	.	222	225	227	229	226
Soort toets	EB	252	251	255	256	259	264	264
	NT	153	155	162	166	167	197	199
Advies VO	vmbo-BB	178	177	179	181	179	184	178
	vmbo-KB	208	207	208	209	212	213	215
	vmbo-GT	230	228	230	231	232	233	232
	havo	268	266	270	269	272	269	269
	vwo	320	319	320	321	325	321	317
Thuis taal	Alleen NLs	252	252	256	256	.	.	256
	NLs + bu	237	236	242	242	.	.	241
	Alleen bu	232	232	232	234	.	.	236
Regio	Noord	245	244	254	253	254	255	251
	Oost	249	248	252	253	256	253	252
	West	249	249	253	252	256	255	254
	Zuid	254	253	257	256	257	256	258
Urbanisatiegraad	Zeer sterk	242	242	245	246	250	250	250
	Sterk	250	249	253	253	256	255	253
	Matig	252	252	257	255	259	256	256
	Weinig	252	251	256	257	257	255	255
	Niet	251	250	255	255	257	257	256
Schoolgrootte	Klein (<200 ln)	.	.	.	251	253	252	251
	Groot (200+ ln)	.	.	.	254	257	256	256

Bijlage 5 Verdeling van de achtergrondvariabelen in de onderzoekspopulatie in 2014 die een eindtoets gemaakt heeft

N = aantal leerlingen in de onderzoekspopulatie in 2014

%A = percentage leerlingen per niveau van de achtergrondvariabele

%B = als %A, maar dan als de ontbrekende waarden niet meetellen in de percentages

NB: als er geen %B gegeven is, dan is %B gelijk aan %A (geen/nauwelijks ontbrekende waarden)

Peilingskenmerk

Jaar	N	%A
Totaal in 2014	148314	100

Leerlingkenmerken

Formatiegewicht (F4)	N	%A	%B
0.00	109546	73,9	87,4
0.30	8822	5,9	7,0
1.20	7012	4,7	5,6
<i>niet ingevuld</i>	22934	15,5	

Formatiegewicht (F3)*	N	%A
0.00/ontbrekend	132480	89,3
0.30	8822	5,9
1.20	7012	4,7

* Verdeling als ontbrekende informatie over fgw gelijk gezien wordt als een gewicht 0.00

Geslacht	N	%A	%B
Jongen	72945	49,2	49,7
Meisje	73763	49,7	50,3
<i>Onbekend</i>	1606	1,1	

Leertijd	N	%A	%B
Regulier (10-11jr)	123983	83,6	84,3
Vertraagd (12-13jr)	23027	15,5	15,7
<i>Onbekend</i> **	1304	0,9	

** of te jong (jonger dan 10) of te oud (ouder dan 13) op 1 oktober 2013

Leeftijd	N	%A	%B
11 jaar of jonger	124122	83,7	84,3
12 jaar of ouder	23032	15,5	15,7
<i>Onbekend</i>	1160	0,8	

Eindtoets	N	%A
Basis	121846	82,2
Niveau	22715	15,3
Anker	3753	2,5

Code I	N	%A	
Langer dan 4 jaar in NL	147577	99,5	
Korter dan 4 jaar in NL	737	0,5	
Code IJK	N	%A	
Geen I, J of K	138310	93,3	
IJK-code	10004	6,7	
Advies VO	N	%A	%B
vmbo-BB	6681	4,5	6,2
vmbo-BB/KB	4870	3,3	4,5
vmbo-KB	11725	7,9	10,9
vmbo-KB/GT	875	0,6	0,8
vmbo-GT	21799	14,7	20,3
vmbo-GT/Havo	13482	9,1	12,6
Havo	17796	12,0	16,6
Havo/vwo	12573	8,5	11,7
vwo	17322	11,7	16,2
Onbekend	41191	27,8	
Thuis taal (algemeen)	N	%A	%B
Alleen Nederlands	129521	87,3	91,1
Nederlands + Andere	1953	1,3	1,4
Alleen Andere taal	10661	7,2	7,5
Onbekend	6179	4,2	
Thuis taal (detail)	N	%A	%B
Alleen Nederlands	129521	87,3	91,1
Turks (+)	3444	2,3	2,4
Arabisch (+)	3442	2,3	2,4
Surinaams/Antilliaans (+)	768	0,5	0,5
Anders (+)	4960	3,3	3,5
Onbekend	6179	4,2	

(+): alleen die taal of eventueel in combinatie met Nederlands

Schoolkenmerken

Stratum***	N	%A
S1: 0% tot 10% fgw lln	97158	65,5
S2: 10% tot 25% fgw lln	34461	23,2
S3: 25% of meer fgw lln	16695	11,3

* stratum is gedefinieerd door percentage formatiegewicht leerlingen (fgw lln) op school

Regio****	N	%A
Noord	14893	10,0
Oost	36339	24,5
West	62692	42,3
Zuid	34388	23,2

Verstedelijking****	N	%A
Zeer sterk stedelijk	22668	15,3
Sterk stedelijk	36157	24,4
Matig stedelijk	31582	21,3
Weinig stedelijk	37850	25,5
Niet stedelijk	20055	13,5
**** 2 leerlingen onbekend (geen effect op %)		

Schoolgrootte	N	%A
Klein (1-200)	46326	31,2
Groot (>200)	101988	68,8

Cito helpt je inzicht te krijgen in je ontwikkeling en mogelijkheden. Door kennis, vaardigheden en competenties objectief meetbaar te maken en de ontwikkeling er van te volgen, kun je het beste uit jezelf halen, verantwoorde keuzes maken en beter richting geven aan je toekomst. Cito draagt daaraan bij door wereldwijd werk te maken van goed en eerlijk toetsen, vanuit de kernwaarden kundig, toonaangevend, integer, innovatief en betrokken.

Cito

Amsterdamseweg 13
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11
www.cito.nl

Fotografie: Ron Steemers