



Meting taal en rekenen 2011

**Derde meting: een indicatie van leerprestaties in het basis- en voortgezet onderwijs en het
middelbaar beroepsonderwijs in termen van het referentiekader**

Wetenschappelijke verantwoording

A. van Til, B.T. Hemker en M. van Onna
© Stichting Cito Instituut voor Toetsontwikkeling Arnhem, 2012

Inhoud

Samenvatting	3
1 Inleiding	6
2 Methode van onderzoek	8
2.1 Eindtoets Basisonderwijs	8
2.2 Diagnostische toetsen	9
2.3 Pilot Rekentoets VO	9
3 Populatiebeschrijving	11
3.1 Populatiebeschrijving Eindtoets Basisonderwijs	11
3.2 Populatiebeschrijving diagnostische toetsen in VO en mbo	12
3.3 Populatiebeschrijving Rekentoets VO	13
4 Resultaten	15
4.1 IJking	15
4.2 Verdeling van de referentieniveaus in de populaties	16
4.3 Vergelijking resultaten over jaren heen	19
5 Verschillen tussen leerlingen	25
6 Voorbehoud	30
Literatuur	31
Bijlage 1: Aantallen leerlingen diagnostische toetsen	32
Bijlage 2: Eigenschappen van de gebruikte toetsen	40
Bijlage 3: Vergelijkingen van prestaties over de jaren heen	45
Bijlage 4: Achtergrondvariabelen en deelpopulaties	47

Samenvatting

In dit verslag rapporteren we prestaties van leerlingen in 2011 uit het basis- en voortgezet onderwijs en het mbo ten aanzien van het Referentiekader taal en rekenen. Daarnaast geven we weer hoe de prestaties van leerlingen uit 2011 zich verhouden tot voorgaande jaren. Ten slotte laten we zien hoe leerlingen uit verschillende deelpopulaties scoren ten opzichte van elkaar. Daaruit is bijvoorbeeld af te leiden of jongens het beter doen dan meisjes of omgekeerd.

Toetsen

De resultaten in dit verslag zijn uitsluitend gebaseerd op toetsen die door Cito zijn afgenomen in 2011. Het gaat om de volgende toetsen:

- Eindtoets Basisonderwijs
- Diagnostische toetsen taal en rekenen (afgenomen in VO en mbo)
- Pilot Rekentoets VO

In deze publicatie wordt niet gerapporteerd over toetsen die in 2011 onder verantwoordelijkheid van CVE zijn afgenomen zoals de centraal ontwikkelde examens in het mbo. Dit behoort tot het taakgebied van CVE.

Taakverdeling CVE en Cito

Met ingang van het schooljaar 2011-2012 bepaalt CVE de samenstelling en normering van de volgende toetsen taal en rekenen:

- Eindexamens Nederlands in het voortgezet onderwijs (vanaf het schooljaar 2013-2014 gekoppeld aan het Referentiekader taal en rekenen)
- Rekentoets VO (maakt vanaf schooljaar 2013-2014 deel uit van het eindexamen)
- Centraal ontwikkelde examens in het mbo (worden ingevoerd vanaf schooljaar 2013-2014)

Omdat niet alle toetsen vanaf het schooljaar 2011-2012 al wettelijk zijn voorgeschreven, gaat het in sommige gevallen om een pilot. Cito zal de constructie van de toetsen voor zijn rekening nemen.

Koppeling toetsen aan referentieniveaus

De Eindtoets Basisonderwijs 2011 was inhoudelijk niet afgestemd op de referentieniveaus. Daarnaast zijn de officiële grenswaarden voor 1F nog niet bepaald. Er is zodoende nog geen standaard voor niveau 1F vastgelegd, wat concreet betekent dat we niet weten hoeveel opgaven een leerling goed moet hebben om niveau 1F te halen. Aangezien er nog geen harde koppeling gemaakt kan worden tussen prestaties op de Eindtoets enerzijds en de referentieniveaus anderzijds, gebruiken we de verdeling van de resultaten in 2008 als referentie. De vaardigheid die door 75 procent van de leerlingen behaald werd in 2008 was het referentiepunt. Het is de verwachting dat dit punt niet ver af ligt van niveau 1F.

Voor de overige toetsen geldt dat ze wel gebaseerd zijn op de referentieniveaus en dat we met behulp van de prestaties op de toetsen de beheersingsniveaus voor 2F of 3F in beeld kunnen brengen. Om dat te kunnen doen moesten de toetsen eerst zijn genormeerd. De normering van de diagnostische toetsen 2011 konden we realiseren door de ijkpunten van de eerste indicatieve ijking van 2009 over te brengen op de toetsen van 2011.

Voorbehoud

Hoewel het onderzoek dat aan dit verslag ten grondslag ligt zorgvuldig is opgezet, moeten de resultaten met voorzichtigheid worden geïnterpreteerd. Zoals hiervoor vermeld, is in de eerste plaats de Eindtoets Basisonderwijs inhoudelijk nog niet gebaseerd op de referentieniveaus en zijn de standaarden voor niveau 1F nog niet vastgelegd. In de tweede plaats is de normering die we hebben toegepast op de diagnostische toetsen en de pilot rekentoets VO 2011 nog niet definitief. Pas als een toets een aantal jaren operationeel is, kan een definitieve normering worden vastgesteld. Als deze definitieve normering strenger of soepeler is dan de voorlopige normering, zal dit van invloed zijn op de hier gerapporteerde prestaties. In de derde plaats zijn er, met uitzondering van de Eindtoets Basisonderwijs, nog geen belangrijke beslissingen verbonden aan de prestaties van leerlingen op de in het onderzoek gebruikte toetsen. Dit kan een negatief effect hebben op de motivatie en de prestaties van leerlingen, ook omdat er dikwijls nog geen gericht onderwijsaanbod heeft plaatsgevonden. Ten vierde beschrijven de referentieniveaus de gewenste kennis- en vaardigheidsniveaus aan het einde van een onderwijsperiode. Alle toetsresultaten zijn echter al in het voorjaar verzameld. Dit maakt ook dat de gerapporteerde prestaties mogelijk zijn onderschat. Tenslotte geldt voor de steekproeven in het VO dat ze over het algemeen wel redelijk, maar niet helemaal perfect overeenkomen met landelijke verdelingen. Voor het mbo geldt dat de representativiteit niet kan worden vastgesteld. Waarom het punt van de representativiteit voor dit onderzoek van belang is en in welke mate de steekproeven afwijken, zetten we hieronder uiteen. In ieder geval kunnen we stellen dat er een aantal redenen is om een slag om de arm te houden bij het duiden van de gepresenteerde resultaten.

Representativiteit

Het onderzoek waarover we in dit verslag rapporteren, is een vorm van peilingsonderzoek. Bij dit soort onderzoek is het gebruikelijk om een zorgvuldige steekproef te trekken, zodat men ervan uit kan gaan dat de prestaties van de deelnemende leerlingen gegeneraliseerd kunnen worden naar de hele populatie. Alle bovengenoemde toetsen zijn echter afgenomen met een ander doel dan het in kaart brengen van de landelijke prestaties ten aanzien van taal en rekenen. Scholen konden zich vrijwillig inschrijven voor de toetsen en daarom kunnen we er niet bij voorbaat van uitgaan dat onze gegevensverzamelingen representatief zijn voor landelijke populaties. Daarom hebben we eerst in kaart gebracht in hoeverre de verdeling van de deelnemende leerlingen over achtergrondkenmerken als bijvoorbeeld regio en verstedelijking overeenkomt met landelijke verdelingen. De resultaten zetten we hieronder op een rijtje.

Eindtoets basisonderwijs

De leerlingen die deelnamen aan de Eindtoets basisonderwijs zijn representatief voor geslacht en formatiegewicht. Naar regio worden kleine verschillen gevonden. Desondanks verwachten we dat de resultaten gegeneraliseerd kunnen worden. Nadere analyses wijzen namelijk uit dat regio en verstedelijking niet direct samenhangen met de prestaties van leerlingen ten aanzien van taal en rekenen.

Diagnostische toetsen

De spreiding van de deelnemende leerlingen over regio en verstedelijking kwam redelijk overeen met landelijke verdelingen. Er was alleen sprake van een lichte ondervertegenwoordiging van leerlingen uit de zeer verstedelijkte gebieden en van leerlingen uit het westen van het land. Voor mbo kon de representativiteit niet worden vastgesteld wegens een gebrek aan landelijke gegevens. Wel kunnen we melden dat voor mbo alle regio's hebben meegedaan aan de toetsen.

Rekentoets VO

Ook bij de Rekentoets VO komt de groep deelnemende scholen qua spreiding over regio en verstedelijking redelijk overeen met de landelijke populatie. Er is sprake van een lichte oververtegenwoordiging van de regio's noord en west. Daarnaast komen de deelnemers aan de Rekentoets VO relatief iets minder vaak uit sterk verstedelijkte gebieden en juist iets vaker uit weinig of niet verstedelijkte gebieden.

Prestaties van leerlingen ten aanzien van de referentieniveaus in 2011

Hieronder brengen we per sector in beeld hoe leerlingen hebben gepresteerd ten aanzien van de referentieniveaus.

Basisonderwijs

In de vergelijking van de resultaten over de jaren heen is gebruik gemaakt van de resultaten van het project JPON (Jaarlijks Peilingsonderzoek van het Onderwijsniveau). Dit onderzoek is gericht op het monitoren van het onderwijsniveau op het gebied van taal- en rekenvaardigheid in het basisonderwijs. Bij de start van het project, in 2008, hebben we voor verschillende opgavencategorieën prestatieniveaus gedefinieerd, waaronder een niveau dat door 75 procent van de leerlingen werd beheerst. Dit 75-procentniveau is dus niet bepaald op inhoudelijke gronden, maar puur op basis van hoe leerlingen in 2008 hebben gescoord. Het functioneert in het project JPON als een soort ijkpunt, zodat we na kunnen gaan hoe leerlingen zich in de loop der jaren ontwikkelen. Als we nagaan hoe de Eindtoetsleerlingen presteren ten aanzien van dit punt, dan zien we dat er het best is gescoord op de opgavencategorie begrijpend lezen. In deze categorie beheerst 79 procent van de leerlingen het gedefinieerde prestatieniveau. Voor de overige opgavencategorieën ligt het beheersingspercentage op 76 of 77 procent.

VO

Onze meting laat een grote variatie zien tussen de taal- en rekenprestaties bij de verschillende vmbo-leerwegen. Voor vmbo-BB en vmbo-BB+ (dit omvat het leerwegondersteunend en Praktijkonderwijs) lijkt het percentage leerlingen dat niveau 2F beheerst voor geen van de vaardigheden boven de 50 procent te liggen. Voor rekenen en taalverzorging wordt in ons onderzoek het minst goed gepresteerd. Het percentage onderzochte BB-leerlingen dat voor deze vaardigheid niveau 2F beheerst, is respectievelijk 18 en 21 procent. Daarnaast lopen de prestaties van de deelnemende leerlingen uit leerjaar 4 KB erg uiteen. Voor lees- en luistervaardigheid beheerst ongeveer 75 procent van deze leerlingen niveau 2F, voor taalverzorging en rekenen ligt het beheersingspercentage rond de 50 procent. Voor de onderzochte leerlingen uit klas 4 vmbo GT geldt dat het percentage leerlingen dat 2F beheerst voor alle vaardigheden boven de 70 procent ligt. Van de havoleerlingen uit het onderzoek bereikt minstens driekwart van de leerlingen niveau 3F voor leesvaardigheid, luistervaardigheid en taalverzorging. In vergelijking hiermee blijven de prestaties voor rekenen iets achter, want voor deze vaardigheid beheerst slechts 68 procent van

de deelnemende havoleerlingen niveau 3F. Bovenstaande resultaten zijn gebaseerd op afnames van de diagnostische toetsen. Voor zover de afnames van de Rekentoets VO vergelijkbaar zijn met de diagnostische toetsen, komen de resultaten overeen, met uitzondering van het havo. Voor dit schooltype vallen de resultaten van de leerlingen op de Rekentoets VO 10 procent lager uit, mogelijk als gevolg van het feit dat het om een digitale afname gaat, wat leerlingen over het algemeen als iets moeilijker ervaren.

mbo

De afnameresultaten van de diagnostische toetsen laten zien dat de mbo-leerlingen het best lijken te presteren op leesvaardigheid. Voor dit onderdeel haalt 60 tot 94 procent van de deelnemende leerlingen de betreffende referentieniveaus. De meeste moeite lijken mbo-leerlingen te hebben met rekenen en taalverzorging. Het percentage leerlingen dat voor deze leergebieden het betreffende referentieniveau haalt, varieert van 32 tot 65 procent.

Jaarvergelijkingen

Om te bepalen hoe leerlingen uit verschillende leerjaren ten opzichte van elkaar presteren, hebben we de afnamegegevens met elkaar vergeleken. Voor het basisonderwijs blijkt dat leerlingen in 2011 zowel voor taal als voor rekenen gemiddeld beter presteren dan in 2008. Het gaat maar om een klein verschil in prestatie, maar dit verschil is voor alle opgavencategorieën (m.u.v. woordenschat) wel significant, dus kan niet aan toeval worden toegeschreven. Bovendien kunnen we vanwege het grote aantal leerlingen dat meedoet aan de Eindtoets ook een relatief klein verschil als betekenisvol aanmerken. Daarnaast nemen we bij vmbo BB een stijging waar van het percentage deelnemende leerlingen dat 2F haalt voor lees- en luistervaardigheid. Voor de andere schooltypen en niveaus zijn uit deze meting geen duidelijke trends af te leiden. Dat geldt ook voor het mbo.

Verschillen tussen leerlingen

Naast berekeningen om te bepalen welk percentage leerlingen welk referentieniveau haalt, zijn we ook nagegaan hoe bepaalde subgroepen ten opzichte van elkaar presteren, bijvoorbeeld jongens ten opzichte van meisjes of leerlingen uit de sector techniek ten opzichte van leerlingen uit de andere sectoren. Daarbij hebben we in de eerste plaats gekeken naar de grootte van de verschillen. Daarnaast zijn we nagegaan of de verschillen significant waren.

Basisonderwijs

Door middel van effectschattingen is nagegaan hoe bepaalde subgroepen ten opzichte van elkaar presteren. Uit de analyses blijkt dat de volgende achtergrondvariabelen samenhangen met de prestaties.

- geslacht (Meisjes presteren beter op begrijpend lezen en spelling dan jongens maar jongens presteren daarentegen weer beter op rekenen.)
- leertijd (Vertraagde leerlingen presteren slechter op alle gemeten vaardigheden dan reguliere leerlingen.)
- Stratum (Scholen met relatief veel leerlingen van laagopgeleide ouders scoren slechter voor begrijpend lezen en woordenschat dan overige scholen.)
- Formatiegewicht (Leerlingen van relatief laagopgeleide ouders presteren slechter dan de overige leerlingen. De verschillen zijn groter voor begrijpend lezen en woordenschat dan voor rekenen)
- Thuis taal (Leerlingen waar thuis een andere taal (Turks, Arabisch of Surinaams/Antilliaans) wordt gesproken dan Nederlands presteren slechter op woordenschat (groot verschil) en begrijpend lezen dan leerlingen waar thuis uitsluitend Nederlands wordt gesproken. Dit effect wordt niet gevonden voor leerlingen met een andere West-Europese thuis taal. De verschillen tussen de taalgroepen zijn gecorrigeerd voor de effecten van de opleiding van de ouders.

De gerapporteerde resultaten werden in alle vier de vergeleken jaren (2008 tot en met 2011) gevonden.

VO en mbo

Berekeningen laten zien dat jongens op de Rekentoets VO beter lijken te presteren voor rekenen dan meisjes. Dit beeld komt ook naar voren uit de analyses van de Eindtoets Basisonderwijs. Bij het mbo valt vooral op dat de deelnemende leerlingen uit de sector techniek beter presteren voor rekenen dan de onderzochte leerlingen uit de andere sectoren.

1 Inleiding

In opdracht van OCW beschrijft Cito de taal- en rekenprestaties van leerlingen in het basis en voortgezet onderwijs en het mbo voor het schooljaar 2010-2011. Dit is gebeurd in termen van het referentiekader van de Expertgroep Doorlopende Leerlijnen (EGDLL). De beschrijving bouwt voort op rapporten over hetzelfde onderwerp die in 2010 en 2011 zijn verschenen. In dit verslag rapporteren we de resultaten van vervolgonderzoek dat in 2011 is uitgevoerd. Daarnaast brengen we de uitkomsten in verband met de resultaten van 2009 en 2010.

In 2010 is de wet referentieniveaus Nederlandse taal en rekenen van kracht geworden. Doel van deze wet is om een betere aansluiting te bewerkstelligen tussen het taal- en rekenonderwijs in de verschillende sectoren. Daarnaast beoogt de wet ook om de taal- en rekenvaardigheden van leerlingen te verbeteren. Dit komt de doorlopende leerlijn van leerlingen ten goede. De wet is gebaseerd op het werk van de EGDLL: *Over drempels met Taal en Rekenen (2008)* en *Een nadere beschouwing (2009)*.

In 2010 zijn eveneens per sector referentieniveaus vastgelegd. Dit zijn beschrijvingen van wat leerlingen moeten kennen en kunnen op het gebied van taal en rekenen. De toewijzing van de referentieniveaus aan de sectoren is als volgt:

Taal

- Niveau 1F en 1S: eind basisonderwijs en speciaal onderwijs
- Niveau 1F: praktijkonderwijs
- Niveau 2F: eind vmbo, mbo 2 en mbo 3
- Niveau 3F: eind havo en mbo 4
- Niveau 4F: eind vwo

Rekenen

- Niveau 1F en 1S: eind basisonderwijs en speciaal onderwijs
- Niveau 1F: praktijkonderwijs
- Niveau 2F: eind vmbo, mbo 2 en mbo 3
- Niveau 3F: eind havo, vwo en mbo 4

Om er zeker van te zijn dat leerlingen op drempelmomenten ook inderdaad beschikken over het vereiste niveau voor taal en rekenen is bepaald dat deze vakken verplicht getoetst worden aan het einde van het basisonderwijs, het voortgezet onderwijs en het mbo. Hieronder zetten we de genomen maatregelen in het kader van de toetsing van taal en rekenen voor de verschillende sectoren op een rijtje.

Basisonderwijs

In het basisonderwijs nemen de leerlingen in groep 8 in de nabije toekomst mogelijk deel aan een verplichte eindtoets. Inhoudelijk vindt er een aanpassing plaats zodat er over alle deelnemende leerlingen gerapporteerd kan worden in termen van de referentieniveaus. Omdat het constructietraject van de Eindtoets meerdere jaren omvat, zal deze aanpassing niet meteen al in 2013 gerealiseerd kunnen worden. We verwachten dat op zijn vroegst in 2015 de benodigde wijzigingen zijn doorgevoerd zodat de Eindtoets van dat jaar een indicatie kan geven van het aantal leerlingen dat de referentieniveaus beheerst voor taal en rekenen.

Voortgezet onderwijs

In het schooljaar 2013/2014 krijgen leerlingen in het voortgezet onderwijs voor het eerst te maken met een rekentoets als onderdeel van het eindexamen. Uitgangspunt voor de Rekentoets voor vmbo is het referentieniveau 2F, voor havo/vwo is dat 3F. Aanvullend vindt er in de eerste helft van 2012 een onderzoek plaats om na te gaan of het niveau 3S mogelijk een geschikter niveau is voor het vwo. Voor taal geldt dat de eindexamens Nederlands worden geijkt aan het referentiekader.

Daarnaast worden in het VO de exameneisen aangescherpt:

- Het gemiddeld centraal examencijfer moet zonder afronding een 5,5 of hoger zijn (vanaf 1 augustus 2011);
- Voor de (kern)vakken Nederlands, Engels en wiskunde mag maximaal één vijf behaald worden op havo en vwo (vanaf 1 augustus 2012). De uitslag van de Rekentoets gaat op termijn ook deel uitmaken van deze kernvakkenregeling;
- De afwijkende berekening van het eindcijfer in de basisberoepsgerichte leerweg wordt afgeschaft (vanaf 1 augustus 2011).

Middelbaar beroepsonderwijs

In het mbo worden voor de opleidingen op niveau 4 vanaf 2013 - 2014 centraal ontwikkelde examens (COE's) ingevoerd voor taal en rekenen. Voor taal bestaat dit examen uit een toets voor de onderdelen lezen en luisteren. Daarnaast moeten de mbo's een instellingsexamen afnemen voor de onderdelen spreken, gesprekken voeren, schrijven en begrippenlijst/taalverzorging. Het centraal ontwikkelde examen rekenen omvat alle onderdelen van rekenen (getallen, verhoudingen, meten en meetkunde, verbanden). De centraal ontwikkelde examens Nederlandse taal en rekenen voor mbo niveau 2 en 3 worden een jaar later dan de examens voor mbo niveau 4 ingevoerd. Voor de niveau 1 opleidingen valt de beslissing over wel of geen centrale examinering niet voor 2014.

Volgens de zak-slaagregeling in het mbo is het examen met goed gevolg afgelegd, als:

- voor alle examenonderdelen een eindwaardering van tenminste een 6 dan wel "voldoende" is behaald, of
- voor zover een centraal examen deel uitmaakt van de examinering voor Nederlandse taal of rekenen, voor één van die examenonderdelen een eindwaardering van ten minste het cijfer 5 is behaald, en voor de overige examenonderdelen een eindwaardering van ten minste het cijfer 6 dan wel "voldoende" is behaald.

Inhoud en opzet rapportage

In het licht van bovenstaande maatregelen is het van belang om te monitoren waar leerlingen staan ten opzichte van de referentieniveaus taal en rekenen. Het doel van deze rapportage is om daar duidelijkheid over te verschaffen toegespitst op het jaar 2011. In dit rapport zal worden ingegaan op de volgende onderzoeksvragen:

- 1 In hoeverre kan de populatie die aan de toetsen heeft deelgenomen beschouwd worden als een representatieve steekproef?
- 2 Welk deel van de BaO-, VO- en mbo-leerlingen (per leerjaar) beheerst niveau 1F, 2F en 3F?
- 3 Hoe verhouden de prestaties van verschillende deelpopulaties in de sectoren zich tot elkaar?
- 4 Hoe kunnen eventuele verschillen in resultaten van metingen binnen één en dezelfde sector verklaard worden?
- 5 Is er een bepaalde ontwikkeling in de prestaties van leerlingen ten opzichte van vorige jaren?

In hoofdstuk 2 van dit verslag wordt ingegaan op de meetinstrumenten die in dit onderzoek zijn gebruikt en rapporteren we de methode van onderzoek. In hoofdstuk 3 volgt een beschrijving van de onderzoekspopulatie. De resultaten van het onderzoek worden gepresenteerd in hoofdstuk 4. In hoofdstuk 5 geven we weer hoe bepaalde achtergrondvariabelen van leerlingen samenhangen met de prestaties en in hoofdstuk 6 maken we een voorbehoud ten aanzien van de gepresenteerde resultaten.

Toetsen

De resultaten in dit verslag zijn uitsluitend gebaseerd op toetsen die door Cito zijn afgenomen in 2011. Het gaat om de volgende toetsen:

- Eindtoets Basisonderwijs
- Diagnostische toetsen taal en rekenen (afgenomen in VO en mbo)
- Pilot Rekentoets VO

In deze publicatie wordt niet gerapporteerd over toetsen die in 2011 die onder verantwoordelijkheid van CVE zijn afgenomen zoals de centraal ontwikkelde examens in het mbo. Dat behoort tot het taakgebied van CVE.

Taakverdeling CVE en Cito

Met ingang van het schooljaar 2011-2012 bepaalt CVE de samenstelling en normering van de volgende toetsen taal en rekenen:

- Eindexamens Nederlands in het voorgezet onderwijs (vanaf het schooljaar 2013-2014 gekoppeld aan het Referentiekader taal en rekenen)
- Rekentoets VO (maakt vanaf schooljaar 2013-2014 deel uit van het eindexamen)
- Centraal ontwikkelde examens in het mbo (worden ingevoerd vanaf schooljaar 2013-2014)

Omdat niet alle toetsen vanaf het schooljaar 2011-2012 al wettelijk zijn voorgeschreven, gaat het in sommige gevallen om een pilot. Cito zal de constructie van de toetsen voor zijn rekening nemen.

2 Methode van onderzoek

Voor de rapportage van de taal- en rekenprestaties is gebruikgemaakt van vier verschillende bronnen van informatie.

- Eindtoets Basisonderwijs
- diagnostische toetsen taal en rekenen
- pilot Rekentoets vo

In de volgende paragrafen beschrijven we de opzet van de verschillende toetsen. We vermelden ook het aantal leerlingen dat eraan deelnam. Het gaat hierbij om de ruwe aantallen, dat wil zeggen dat de leerlingen die slechts enkele antwoorden hebben aangestreept en daarna zijn afgehaakt, ook zijn meegeteld. In bijlage 1 vermelden we de aantallen waarbij we leerlingen hebben weggelaten die meer dan 20 procent van de antwoorden niet hebben ingevuld. We gaan ervan uit dat deze leerlingen de toets niet serieus hebben gemaakt. Het opnemen van hun resultaten zou een vertekend beeld geven. Ook leerlingen die een onderdeel zo slecht hebben gemaakt dat ze onder het gokniveau uitkwamen zijn om deze reden uit de gegevensverzameling verwijderd. De aantallen uit bijlage 1 vormen de basis voor berekeningen die zijn uitgevoerd om de resultaten te duiden.

Enkele beschrijvende statistische eigenschappen van de gebruikte toetsen zijn weergegeven in bijlage 2. Deze bijlage bevat ook informatie over de betrouwbaarheid van de toetsen. Aan de hand van de richtlijnen van de COTAN (Evers, Lucassen, Meijer & Sijtsma, 2010) is nagegaan in hoeverre de toetsen voldoende betrouwbaar waren voor het doel waarvoor ze in dit onderzoek zijn gebruikt (dit wil zeggen: meting op groepsniveau). Uit deze analyse kwam naar voren dat de betrouwbaarheid van de toetsen bijna altijd goed is en in een enkel geval voldoende. Dat betekent dat de betrouwbaarheidscoëfficiënt van de hierboven genoemde toetsen vrijwel steeds boven de .70 ligt.

De analyses die we hebben uitgevoerd om antwoord te geven op de onderzoeksvragen houden verband met de klassieke testtheorie en een op de itemresponstheorie (IRT) gebaseerd meetmodel. (OPLM, Verhelst, 1993; Verhelst en Glas, 1995).

2.1 Eindtoets Basisonderwijs

De Eindtoets Basisonderwijs wordt op het merendeel van de scholen ingezet bij het adviseren van een geschikt type voortgezet onderwijs. Leerlingen van wie de prestaties achterblijven kunnen een speciale versie van de Eindtoets maken (de Niveautoets). Deze versie bevat evenveel opgaven als de Eindtoets. Ook worden dezelfde leergebieden bevraagd. Het niveau van de opgaven is echter aangepast, zodat ze beter aansluiten bij de doelgroep.

In 2011 namen 144.408 leerlingen uit het basisonderwijs deel aan de Eindtoets of de Niveautoets. Dit aantal is exclusief leerlingen die de toets op een locatie in het buitenland maakten en leerlingen uit het SO en SBO. De leerlingen bogen zich over 200 opgaven op het gebied van taal, rekenen-wiskunde en studievaardigheden. Van deze leerlingen maakte 82% ook de toetsen wereldoriëntatie (facultatief onderdeel; 90 opgaven). In de onderstaande tabel is het aantal opgaven op het gebied van taal en rekenen aangegeven.

tabel 2.1.1: opzet Eindtoets Basisonderwijs en Niveautoets 2011 voor taal en rekenen

subonderdeel	onderdeel	aantal opgaven
schrijven van teksten	taal	30 opgaven
spelling	taal	20 opgaven
begrijpend lezen	taal	30 opgaven
woordenschat	taal	20 opgaven
taal totaal		100 opgaven

subonderdeel	onderdeel	aantal opgaven
getallen en bewerkingen	rekenen-wiskunde	25 opgaven
verhoudingen, breuken en procenten	rekenen-wiskunde	20 opgaven
meten, meetkunde, tijd en geld	rekenen-wiskunde	15 opgaven
rekenen-wiskunde totaal		60 opgaven

2.2 Diagnostische toetsen

De diagnostische toetsen zijn afgenomen in het voorjaar van 2011. Scholen konden zich vrijwillig opgeven voor deze afname. In totaal deden 117.843 leerlingen mee uit het VO en 30.244 uit het mbo. Van 10.914 leerlingen hebben we de herkomst niet kunnen vaststellen. In totaal hebben 159.001 leerlingen aan minstens één van de onderdelen van de diagnostische toetsen gewerkt, bijna een verdubbeling ten opzichte van 2010. Deze toename werd grotendeels veroorzaakt door stijgende leerlingaantallen in het VO. In het mbo nam de belangstelling voor deelname aan de diagnostische toetsen slechts licht toe. 2011 was het derde en laatste jaar dat scholen zich in konden schrijven voor deze toetsen (Voesten, 2012).

Bij de diagnostische toetsen waren opgavenverzamelingen beschikbaar op drie niveaus: 1/2F, 2F en 3F. Voor elk niveau waren vier onderdelen geconstrueerd te weten: lezen, luisteren, taalverzorging en rekenen. Bij rekenen zijn vier subdomeinen bevroegd, te weten: getallen, verhoudingen, meten en meetkunde en verbanden. Het gebruik van een rekenmachine was niet toegestaan. De afnameduur voor de verschillende onderdelen was variabel en is toegevoegd in onderstaande tabel.

Tussen de taken van verschillend niveau is overlap aangebracht zodat het mogelijk werd om de opgaven van alle niveaus qua moeilijkheidsgraad met elkaar te vergelijken. Deze opzet maakte het ook mogelijk om in te schatten hoe de deelnemende leerlingen zouden presteren op een referentieniveau dat hoger of lager lag dan het referentieniveau dat voor hen van toepassing was. Alle toetsen bevatten meerkeuzevragen die op papier werden voorgelegd. In onderstaande tabel is af te lezen hoeveel opgaven er per onderdeel beschikbaar waren.

tabel 2.2.1: opzet diagnostische toetsen 2011

	aantal opgaven 1/2F	aantal opgaven 2F	aantal opgaven 3F	afnameduur
lezen	35	36	40	60 minuten
luisteren	41	45	40	60 minuten, verdeeld in luisteren 1 en luisteren 2
taalverzorging	49	50	51	50 minuten
rekenen	60	60	60	2 maal 50 minuten

2.3 Pilot Rekentoets VO

De opgaven die gemaakt zijn in het kader van de pilot Rekentoets VO zijn afgenomen in eindexamenklassen van vmbo, havo en vwo in het voorjaar van 2011. Het was het tweede jaar dat een dergelijke afname werd georganiseerd. In totaal hebben 6007 leerlingen aan de pilot meegedaan, ongeveer twee maal zoveel als in 2010. De opgaven werden afgenomen via de computer. De beschikbare afnametijd bedroeg 90 minuten. Er is een opgavenverzameling samengesteld voor niveau 2F en één voor niveau 3F. Een aantal van deze opgaven overlapt met rekenopgaven van de pre-pilot van 2010. Op deze manier konden opgaven uit verschillende jaren met elkaar worden vergeleken (Voesten, 2012).

Ten behoeve van de pilot zijn voor elk niveau twee versies gemaakt. Elke versie bestond uit een deel dat met de rekenmachine gemaakt moest worden en een deel waarbij leerlingen de berekeningen moesten maken zonder rekenmachine. Daarnaast bevatte elke versie ongeveer evenveel kale sommen als contextsommen. Tot de definitie van kale sommen zijn ook sommen met benoemde getallen gerekend. Bij

benoemde getallen kan men bijvoorbeeld denken aan getallen met een euroteken. De subdomeinen die in de pilot bevraagd werden, zijn dezelfde als die van de diagnostische toetsen, te weten: getallen, verhoudingen, meten en meetkunde en verbanden. In elke versie van de toets had een kwart van de scorepunten betrekking op één specifiek subdomein. Om een goede aansluiting bij datgene wat onderwezen is in het basisonderwijs te bewerkstelligen, is gekozen om het merendeel van de opgaven zonder rekenmachine te presenteren. In onderstaande tabel is af te lezen hoeveel opgaven er per onderdeel beschikbaar waren.

tabel 2.3.1: opzet Rekentoets VO

	aantal opgaven 2F, versie 1	aantal opgaven 2F, versie 2	aantal opgaven 3F, versie 1	aantal opgaven 3F, versie 2
opgaven met rekenmachine	5 opgaven	11 opgaven	20 opgaven	22 opgaven
opgaven zonder rekenmachine	54 opgaven	48 opgaven	40 opgaven	38 opgaven
totaal	59 opgaven	59 opgaven	60 opgaven	60 opgaven
ankers naar 2010	32 opgaven	25 opgaven	35 opgaven	35 opgaven
nieuw in 2011	28 opgaven	35 opgaven	25 opgaven	25 opgaven

De inhoudelijke keuzes die gemaakt zijn ten behoeve van de Rekentoets VO zijn tot stand gekomen voordat de Rekentoetswijzers waren gepubliceerd. Deze publicaties vormen het kader voor eisen die in de rekentoetsen gesteld worden in termen van 'kennen' en 'kunnen'. Daarom wijkt de samenstelling van de toetsen enigszins af van wat er in deze documenten wordt aanbevolen. Zo bevatten de toetsen van 2011 meer opgaven zonder rekenmachine dan opgaven met rekenmachine. Daarnaast is er binnen de Rekentoetswijzers een stringenter definitie van kale opgaven gehanteerd. Dit verklaart waarom de benoemde getallen uit de afname van 2011 zijn ondergebracht in de categorie 'contextloos'. Bij het samenstellen van de rekentoetsen voor 2012 hebben de Rekentoetswijzers wel als uitgangspunt kunnen fungeren.

3 Populatiebeschrijving

In dit hoofdstuk beschrijven we de populaties die hebben deelgenomen aan de toetsen. In alle gevallen ging het niet om aselechte steekproeven, maar om groepen die zich vrijwillig hadden ingeschreven voor de toetsen. Als we de prestaties van deze groepen willen opvatten als een indicatie voor 'de stand van het land', moeten we aangeven hoe de steekproeven zich verhouden tot de landelijke populaties. Dit doen we in paragraaf 3.1 tot en met 3.5. Daarbij komt met name de representativiteit met betrekking tot regio en verstedelijking aan bod.

3.1 Populatiebeschrijving Eindtoets Basisonderwijs

Het aantal leerlingen dat deelnam aan de Eindtoets is vergelijkbaar met 2010 en 2009. De verdeling over achtergrondkenmerken is over de jaren heen constant, wat een eerste indicatie is voor representativiteit van de onderzochte populatie.

Om na te kunnen gaan of de populatie die heeft deelgenomen aan de Eindtoets inderdaad wel representatief is voor de gehele basisschoolpopulatie moet worden nagegaan of de verdeling van verschillende achtergrondvariabelen in de Eindtoetspopulatie een afspiegeling is van landelijke verdelingen. Om deze landelijke verdelingen in kaart te brengen wenden we ons over het algemeen tot de databank Statline van het CBS en tot het CFI. Het CBS heeft bepaalde kenmerken echter per leeftijd geïnterviewd en niet per leerjaar. Om bijvoorbeeld een idee te krijgen van het percentage leerlingen uit het basisonderwijs dat meedoet aan de Eindtoets, hebben we het aantal 11-jarige basisschoolleerlingen opgevraagd. Op basis van een vergelijking van dit aantal met het aantal leerlingen uit onze database schatten we dat onze Eindtoetspopulatie 82 procent van het totaal aantal basisschoolleerlingen omvat.

Hieronder presenteren we een tweetal tabellen waarin de spreiding over regio en geslacht bij de Eindtoets wordt afgezet tegen landelijke verdelingen. De landelijke verdelingen betreffen wederom 11-jarigen. We gaan ervan uit dat deze getallen de spreiding van basisschoolleerlingen in groep 8 over regio en geslacht het best benaderen. Overigens hebben we in navolging van het CBS de provincies als volgt over de regio's verdeeld:

- Noord Nederland: Groningen, Friesland, Drenthe;
- Oost Nederland: Gelderland, Overijssel, Flevoland;
- West Nederland: Utrecht, Noord-Holland, Zuid-Holland, Zeeland;
- Zuid Nederland: Brabant, Limburg.

tabel 3.1.1: Verdeling leerlingen basisonderwijs naar regio

	Eindtoets basisonderwijs	CBS (11-jarige kinderen)
Noord	10%	10%
Oost	24%	24%
West	42%	46%
Zuid	24%	21%

tabel 3.1.2: Verdeling leerlingen basisonderwijs naar geslacht

	Eindtoets basisonderwijs	CBS (11-jarige kinderen)
jongens	50%	51%
meisjes	50%	49%

Uit bovenstaande tabellen blijkt dat de deelnemende leerlingen aan de Eindtoets representatief zijn voor geslacht. Het (geringe) verschil voor geslacht is mogelijk te verklaren doordat in het SO/SBO relatief meer jongens zitten. De eindtoetsgegevens betreffen hier alleen het reguliere basisonderwijs. Voor regio worden kleine verschillen gevonden. Leerlingen uit het zuiden van het land zijn licht oververtegenwoordigd en leerlingen uit het westen van het land zijn juist licht ondervertegenwoordigd. Dit is op zich geen probleem voor duiding van de resultaten op de toets, omdat we straks bij het hoofdstuk over verschillen tussen leerlingen zullen zien dat de invloed van regio op de resultaten nihil is. Dat wil zeggen dat er geen verschillen zijn in prestatie tussen de leerlingen uit de verschillende regio's, als er gecorrigeerd wordt voor formatiegewicht. Dit geldt overigens ook voor urbanisatiegraad, een reden waarom we hebben afgezien van het in kaart brengen van verdelingen voor verstedelijking. Een achtergrondvariabele die wel samenhang vertoont met de prestaties is het formatiegewicht van leerlingen. In onderstaande tabel presenteren we

daarom de landelijke verdelingen voor formatiegewicht ten opzichte van die voor de Eindtoets 2011. De landelijke verdelingen voor 2011 zijn overigens alleen bekend voor het gehele basisonderwijs. Uit de tabel komt naar voren dat de leerlingen die deelnamen aan de Eindtoets tezamen een goede afspiegeling vormen van de landelijke populatie. Hoewel de verschillen tussen de steekproef en de populatie gering zijn, is er voor gekozen om zelfs deze kleine verschillen op te heffen door de steekproef geheel representatief aan de populatie te maken.

tabel 3.1.3: Verdeling leerlingen in het basisonderwijs naar formatiegewicht

	Eindtoets basisonderwijs	CFI (BaO algemeen)
0,00	86,8%	87,2%
0,30	7,5%	7,3%
1,20	5,7%	5,5%

3.2 Populatiebeschrijving diagnostische toetsen in VO en MBO

Bij de diagnostische toetsen is in de eerste plaats gekeken in hoeverre de deelnemende scholen voor voortgezet onderwijs representatief waren voor het landelijk gemiddelde. Daarbij hebben we ons met name gericht op de vergelijking met landelijke gegevens voor regio en verstedelijkingsgraad. Voor het bepalen van de landelijke gemiddeldes zijn de gegevens van het CFI uit 2008 gebruikt. Dit betreft de gegevens van alle VO-scholen die door de overheid gefinancierd worden. Daarnaast staan ook de gegevens uit Wolf vermeld. Wolf is een programma dat gebruikt wordt voor het verzamelen van data bij de landelijke eindexamens. Beide populatie-schattingen wijken alleen op afrondingsfouten van elkaar af.

Op basis van de postcodes van alle geregistreerde locaties zijn de landelijke percentages voor regio en verstedelijkingsgraad bepaald. De verdeling van provincies over regio's volgen de gebruikelijke indeling (zie boven). De verstedelijkingsgraad is ingedeeld in vijf categorieën. Deze variëren van 1 (zeer sterk verstedelijkt) tot 5 (niet verstedelijkt).

Op landelijk niveau staan 1293 locaties geregistreerd (1153 volgens Wolf). 66 locaties hebben deelgenomen aan de diagnostische toetsen 2011. De resultaten voor de verdeling van de populatie over regio en verstedelijking zijn weergegeven in onderstaande tabellen. In vergelijking met de landelijke verdeling bleef de deelname van leerlingen uit de grootste vier steden van Nederland wat achter. Aangezien de zeer sterk verstedelijkte gebieden voornamelijk in de G4 liggen, was de deelname in de zeer sterk verstedelijkte gebieden eveneens lager dan verwacht op basis van landelijke spreiding van scholen.

Tabel 3.2.1: Verdeling VO-scholen naar regio voor de diagnostische toetsen taal en rekenen 2011.

Regio	RNGT-V5	CFI 2008	Wolf 2011
Noord	14%	13%	13%
Oost	23%	21%	20%
West (incl. G4)	43%	48%	48%
G4	9%	14%	14%
Zuid	19%	19%	19%
Onbekend	0	1	1
Aantal scholen	436	1293	1153

Tabel 3.2.2: Verdeling VO-scholen naar verstedelijkingsgraad voor de diagnostische toetsen taal en rekenen 2011.

Verstedelijking	RNGT-V5	CFI 2008	Wolf 2011
Zeer sterk	15%	20%	20%
Sterk	33%	33%	33%
Matig	25%	22%	22%
Weinig	19%	18%	17%
Niet	8%	7%	8%
Onbekend	0	1	1
Aantal scholen	436	1293	1153

Populatiebeschrijving diagnostische toetsen in mbo

Bij scholen voor mbo die deelnamen aan de diagnostische toetsen was het niet duidelijk of het om neven- of hoofdlocaties ging. Er is dan ook geen goede populatieverdeling naar regio of verstedelijkingsgraad op te stellen. Duidelijk is wel dat alle regio's hebben deelgenomen. Omdat voor het mbo niet bekend is hoe de landelijke verdelingen zijn over regio en verstedelijkingsgraad, kan niets gezegd worden over de representativiteit van de deelnemers aan de diagnostische toetsen uit het mbo.

Tabel 3.2.3: Verdeling mbo-scholen naar regio voor de diagnostische toetsen taal en rekenen 2011.

Regio	RNGT-V5
Noord	19%
Oost	20%
West (incl. G4)	43%
G4	14%
Zuid	18%
Onbekend	0
Aantal scholen	260

Tabel 3.2.4: Verdeling mbo-scholen naar verstedelijkingsgraad voor de diagnostische toetsen taal en rekenen 2011

Verstedelijking	RNGT-V5
Zeer sterk	22%
Sterk	33%
Matig	22%
Weinig	18%
Niet	6%
Onbekend	0
Aantal scholen	260

3.3 Populatiebeschrijving Rekentoets VO

Voor de Rekentoets VO is eveneens de landelijke spreiding berekend aan de hand van de postcodes van de geregistreerde locaties. Daaruit komt naar voren dat er sprake is van een lichte oververtegenwoordiging van de regio's noord en west. Ook voor urbanisatiegraad komt de steekproef niet helemaal overeen met landelijke gemiddeldes. De deelnemers aan de Rekentoets VO komen relatief iets minder vaak uit sterk verstedelijkte gebieden en juist iets vaker uit weinig of niet verstedelijkte gebieden.

Tabel 3.3.1: Verdeling vo-scholen naar regio voor de Rekentoets VO 2011.

Regio	RVO 2011	CFI 2008	Wolf 2011
Noord	17%	13%	13%
Oost	12%	21%	20%
West (incl. G4)	55%	48%	48%
G4	9%	14%	14%
Zuid	17%	19%	19%
Onbekend	0	1	1
Aantal scholen	66	1293	1153

Tabel 3.3.2: Verdeling VO-scholen naar verstedelijking voor de Rekentoets VO 2011.

Verstedelijking	RVO 2011	CFI 2008	Wolf 2011
Zeer sterk	20%	20%	20%
Sterk	21%	33%	33%
Matig	24%	22%	22%
Weinig	24%	18%	17%
Niet	11%	7%	8%
Onbekend	0	1	1
Aantal scholen	66	1293	1153

4 Resultaten

In dit hoofdstuk rapporteren we hoe leerlingen uit de verschillende sectoren presteren ten aanzien van het Referentiekader taal en rekenen. We leggen eerst uit hoe de standaarden voor de verschillende referentieniveaus zijn bepaald bij de verschillende toetsen. Vervolgens presenteren we de resultaten voor 2011. Ten slotte geven we aan hoe deze resultaten zich verhouden tot voorgaande jaren.

4.1 IJking

Dit verslag heeft tot doel om in kaart te brengen hoe leerlingen in het schooljaar 2010/2011 scoren ten aanzien van de referentieniveaus. Om dat te kunnen doen, is het noodzakelijk dat er bij de gebruikte meetinstrumenten grensscores zijn vastgesteld die aangeven hoe vaardig een leerling moet zijn om een bepaald referentieniveau te halen. De beschrijving van de referentieniveaus geeft aan welke inhoudselementen leerlingen moeten beheersen om een bepaald niveau te halen. Om vast te stellen of een leerling het referentieniveau voldoende beheerst, zijn opgaven ontwikkeld bij de referentieniveaus. Van deze opgaven weten we dan voor welk niveau ze gemaakt zijn (en dus bij welk niveau ze passen). Binnen de ruimte die de referentieniveaus laten, bestaat dan nog steeds de mogelijkheid om moeilijker en gemakkelijker opgaven te maken. Vaak is pas na afname van die opgaven echt bekend hoe moeilijk ze zijn. Het is goed mogelijk om bijvoorbeeld twee toetsen te maken die beide het 2F niveau meten, maar waarvan de ene veel moeilijker opgaven bevat dan de andere. Als er dus een toets met opgaven bij een referentieniveau is ontwikkeld, moet nog worden onderzocht of het om een makkelijke of een moeilijke toets gaat. Er moet dus worden vastgesteld hoeveel opgaven een leerling goed moet hebben op de toets om het referentieniveau te halen. Als de opgaven wat gemakkelijker zijn dan moet een leerling logischerwijs meer opgaven goed hebben dan wanneer de opgaven wat moeilijker zijn (Beguin, 2010).

Deze relatie tussen een toets en de referentieniveaus wordt dus zowel inhoudelijk als in de gestelde prestatie-eis vastgesteld. Dit gebeurt op twee momenten:

- Tijdens de constructie van de toets, door die afzonderlijke opgaven te selecteren die betrekking hebben op de juiste inhoudsdomeinen en waarvan het niveau past bij de te meten referentieniveaus.
- Bij het bepalen van de grensscore die het beste onderscheid maakt tussen kandidaten die een referentieniveau wel of niet hebben bereikt. Dit gebeurt onder andere door het gebruik van ankeritems. Ankeritems zijn items die in meerdere toetsen voorkomen, zodat de score van de ene toets kan worden vergeleken met score van de andere toets (Beguin, 2010).

In 2009 hebben we voor het eerst in kaart gebracht hoe leerlingen presteerden ten aanzien van de referentieniveaus. We hebben daarvoor de afnamegegevens gebruikt van de diagnostische toetsen uit dat jaar. Ten behoeve van de rapportage hebben we een eerste indicatieve ijking uitgevoerd. Dit hield in dat we hebben bepaald hoeveel opgaven van de diagnostische toetsen leerlingen goed moesten hebben om vast te stellen dat zij een bepaald referentieniveau beheersten. Hieronder herhalen we de uitgangspunten van de ijking in 2009 nog eens:

- Het ijkpunt voor 1F wordt vastgelegd op de toetsscore waarbij 75 procent van de leerlingen van groep 8 uit het basisonderwijs 2009 dit niveau ook daadwerkelijk beheerst.
- Het ijkpunt voor 2F wordt vastgelegd op de toetsscore waarbij 50 procent van de leerlingen van klas 4 vmbo KB 2009 dit niveau ook daadwerkelijk beheerst.
- Het ijkpunt voor 3F wordt vastgelegd op de toetsscore waarbij 75 procent van de leerlingen van klas 5 havo 2009 dit niveau ook daadwerkelijk beheerst.

Voor alle uitgangspunten geldt dat hierbij aansluiting is gezocht met het advies van de commissie EGDLL (Monitor taal en rekenen. Eerste meting: een indicatie van leerprestaties in termen van het referentiekader, 2010).

In 2011 is een nieuwe versie van de diagnostische toetsen afgenomen. Daarnaast is in het BaO, VO en mbo een aantal andere toetsen afgenomen die ook taal- en rekenopgaven bevatten, te weten:

- Eindtoets Basisonderwijs
- Pilot Rekentoets VO

Om de ijkpunten uit 2009 over te kunnen brengen op de diagnostische toetsen die in 2011 zijn afgenomen is er bij de afnames in het VO en mbo voor gezorgd dat er een overlap was tussen de toetsen. Opgaven die waren afgenomen in 2009, maakten ook deel uit van de afnames van 2011, zodat de ijkpunten van de eerste indicatieve ijking uit 2009 konden worden overgebracht op de toetsen van 2011.

4.2. Verdeling van de referentieniveaus in de populaties

Resultaten Eindtoets Basisonderwijs

De Eindtoets Basisonderwijs was in de periode 2008 tot en met 2011 inhoudelijk nog niet afgestemd op de referentieniveaus. Dat komt onder andere tot uiting in de samenstelling van de toets. Voor rekenen onderscheidde we bijvoorbeeld in de Eindtoets de subdomeinen: 1) getallen en bewerkingen, 2) verhoudingen, breuken, procenten, 3) meten, meetkunde, tijd en geld terwijl er in het referentiekader sprake is van vier subdomeinen te weten: 1) getallen, 2) verhoudingen, 3) meten en meetkunde en 4) verbanden. Dit heeft tot gevolg dat we met behulp van de Eindtoets Basisonderwijs 2008-2011 geen volledig inzicht kunnen geven in de prestaties van de basisschoolpopulatie ten aanzien van 1F. In de tweede plaats is het ijkpunt 1F nog niet definitief vastgelegd. Dit kan pas plaatsvinden nadat de Eindtoets inhoudelijk op de Referentieniveau is afgestemd. Aangezien er nog geen harde koppeling gemaakt kan worden tussen prestaties op de Eindtoets enerzijds en de referentieniveaus anderzijds, maken we gebruik van gegevens uit het project JPON (jaarlijks peilingsonderzoek van het onderwijsniveau).

In dit project is in 2008 al een begin gemaakt met het monitoren van taal- en rekenvaardigheid van basisschoolleerlingen op basis van de gegevens uit de Eindtoets Basisonderwijs. Er zijn toen vijf prestatieniveaus gedefinieerd. Het eerste niveau, dat van de zeer laag vaardige leerling, is vastgelegd op een punt van de meetlat dat door 90 procent van de leerlingen wordt beheerst. Het niveau van de laag vaardige leerling is gefixeerd op een punt dat door 75 procent van de leerlingen wordt beheerst. Nu is 75 procent ook het percentage dat is gebruikt om 1F vast te leggen bij de ijking van de diagnostische toetsen 2009 en daarom hebben we dit aangehouden als referentiepunt. Het is de verwachting dat het niet ver af ligt van niveau 1F.

Hieronder presenteren we een tabel waarin verschillende prestatieniveaus worden weergegeven voor de periode 2008 tot en met 2011. Zoals we hiervoor al hebben aangegeven, worden in de Eindtoets deels andere opgavencategorieën onderscheiden dan in het Referentiekader taal en rekenen. Daarom is de indeling van de tabel voor het basisonderwijs anders dan voor het voortgezet en middelbaar beroepsonderwijs, waarvoor toetsen zijn gemaakt die inhoudelijk wel zijn gebaseerd op de referentieniveaus. Uit de tabel is af te lezen dat er in 2011 meer leerlingen waren die het prestatieniveau van de laag vaardige leerling beheersten dan in 2008. Dit geldt voor alle opgavencategorieën. Het niveau van de laag vaardige leerling voor begrijpend lezen werd in 2011 bijvoorbeeld door 79 procent van de leerlingen beheerst in plaats van door 75 procent in 2008. Dat betekent dat er vier procent meer leerlingen waren die dit niveau onder de knie hadden. Het gaat maar om een klein verschil in prestatie, maar dit verschil is voor alle opgavencategorieën (m.u.v. woordenschat) significant, dus kan niet aan toeval worden toegeschreven. Bovendien kunnen we vanwege het grote aantal leerlingen dat meedoet aan de Eindtoets ook een relatief klein verschil als betekenisvol aanmerken, ook al is er in statistische termen eigenlijk te weinig leerwinst om te kunnen spreken van een effect (Cohen, 1977).

tabel 4.2.1: percentage leerlingen dat een prestatieniveau beheerst in de periode 2008 tot en met 2012 voor taal- en rekenvaardigheid

Taalvaardigheid										
Percentage leerlingen met minstens niveau	2008	woordenschat			spelling			begrijpend lezen		
		2009	2010	2011	2009	2010	2011	2009	2010	2011
zeer laag vaardige leerling	90	90	92	90	90	91	91	91	91	92
laag vaardige leerling	75	74	79	76	75	76	76	76	77	79
standaard leerling	50	49	55	51	50	51	51	52	53	56
hoog vaardige leerling	25	24	29	26	25	26	26	26	27	30
zeer hoog vaardige leerling	10	10	13	10	10	11	11	11	11	13

Rekenvaardigheid										
Percentage leerlingen met minstens niveau	2008	getallen en bewerkingen			verhoudingen, breuken, procenten			meten, meetkunde, tijd en geld		
		2009	2010	2011	2009	2010	2011	2009	2010	2011
zeer laag vaardige leerling	90	90	91	91	90	91	91	90	91	91
laag vaardige leerling	75	75	76	76	75	77	77	74	77	77
standaard leerling	50	50	52	51	50	53	53	49	53	53
hoog vaardige leerling	25	25	26	26	25	27	27	24	28	27
zeer hoog vaardige leerling	10	10	11	11	10	11	11	10	11	11

Resultaten diagnostische toetsen VO

De prestaties van de VO-leerlingen van de verschillende schooltypen zijn hieronder weergegeven in tabellen. Hieruit komt naar voren dat er een grote variatie lijkt te tussen de taal- en rekenprestaties in de verschillende vmbo-leerwegen. Voor vmbo-BB en vmbo-BB+ (Dit omvat het leerwegondersteunend en Praktijkonderwijs) geldt dat het percentage leerlingen dat niveau 2F beheerst voor geen van de vaardigheden boven de 50 procent lijkt te liggen. Voor rekenen en taalverzorging wordt het minst goed gepresteerd. Het percentage BB-leerlingen dat voor deze vaardigheid niveau 2F lijkt te beheersen, is respectievelijk 18 en 21 procent. Daarnaast kunnen we uit de tabellen opmaken dat de prestaties van de deelnemende leerlingen uit leerjaar 4 KB erg uiteenlopen. Voor lees- en luistervaardigheid lijkt ongeveer 75 procent van de leerlingen niveau 2F te beheersen, voor taalverzorging en rekenen ligt het beheersingspercentage rond de 50 procent. Voor de onderzochte leerlingen uit klas 4 vmbo GT geldt dat het percentage leerlingen dat 2F beheerst voor alle vaardigheden boven de 70 procent ligt. Van de deelnemende havo-leerlingen bereikt minstens driekwart van de leerlingen niveau 3F voor leesvaardigheid, luistervaardigheid en taalverzorging. In vergelijking hiermee lijken de prestaties voor rekenen iets achter te blijven, want voor deze vaardigheid beheerst slechts 68 procent van de deelnemende havo-leerlingen niveau 3F.

tabel 4.2.2: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het voortgezet onderwijs voor leesvaardigheid bij de diagnostische toetsen

Leesvaardigheid	1F	2F	3F
vmbo BB+, klas 4*	52	27	0
vmbo BB, klas 4	72	49	1
vmbo KB, klas 4	91	79	9
vmbo GT, klas 4	95	89	26
havo, klas 5	100	99	75
vwo, klas 6*	100	100	95

tabel 4.2.3: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het voortgezet onderwijs voor luistervaardigheid bij de diagnostische toetsen

Luistervaardigheid	1F	2F	3F
vmbo BB+, klas 4*	56	27	0
vmbo BB, klas 4	72	44	1
vmbo KB, klas 4	91	75	9
vmbo GT, klas 4	95	87	36
havo, klas 5	100	100	84
vwo, klas 6**	100	100	100

tabel 4.2.4: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het voortgezet onderwijs voor taalverzorging bij de diagnostische toetsen

Taalverzorging	1F	2F	3F
vmbo BB+, klas 4*	39	9	0
vmbo BB, klas 4	54	21	2
vmbo KB, klas 4	84	50	8
vmbo GT, klas 4	91	74	38
havo, klas 5	99	95	80
vwo, klas 6*	100	100	97

tabel 4.2.5: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het voortgezet onderwijs voor rekenen bij de diagnostische toetsen

Rekenen	1F	2F	3F
vmbo BB+, klas 4	44	6	0
vmbo BB, klas 4	60	18	1
vmbo KB, klas 4	79	47	11
vmbo GT, klas 4	96	80	34
havo, klas 5	98	92	68
vwo, klas 6	100	99	95

* = aantal leerlingen kleiner dan 300, maar groter dan 100,

** = aantal leerlingen kleiner dan 100

Resultaten volgens de diagnostische toetsen voor het mbo

In onderstaande tabellen worden de prestaties van de vier verschillende mbo-opleidingen (niveau 1 tot en met 4) met elkaar vergeleken. De afnameresultaten laten zien dat de mbo-leerlingen het best lijken te presteren op leesvaardigheid. Voor dit niveau haalt 60 tot 94 procent van de deelnemers de betreffende referentieniveaus. De meeste moeite lijken de onderzochte leerlingen te hebben met rekenen en taalverzorging. Het percentage leerlingen dat voor deze leergebieden het betreffende referentieniveau lijkt te halen varieert van 32 tot 65 procent.

tabel 4.2.6: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het mbo voor leesvaardigheid bij de diagnostische toetsen

Leesvaardigheid	1F	2F	3F
mbo, niveau 1, leerjaar 1	58	42	3
mbo, niveau 2, leerjaar 2	90	79	16
mbo, niveau 3, leerjaar 3	98	94	39
mbo, niveau 4, leerjaar 4	99	98	60

tabel 4.2.7: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het mbo voor luistervaardigheid bij de diagnostische toetsen

Luistervaardigheid	1F	2F	3F
mbo, niveau 1, leerjaar 1	56	39	5
mbo, niveau 2, leerjaar 2	81	63	12
mbo, niveau 3, leerjaar 3	90	80	31
mbo, niveau 4, leerjaar 4	97	92	49

tabel 4.2.8: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het mbo voor taalverzorging bij de diagnostische toetsen

Taalverzorging	1F	2F	3F
mbo, niveau 1, leerjaar 1	46	23	5
mbo, niveau 2, leerjaar 2	60	32	7
mbo, niveau 3, leerjaar 3*	75	47	13
mbo, niveau 4, leerjaar 4	87	74	50

tabel 4.2.9: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het mbo voor rekenen bij de diagnostische toetsen

Rekenen	1F	2F	3F
mbo, niveau 1, leerjaar 1	52	24	4
mbo, niveau 2, leerjaar 2	67	36	8
mbo, niveau 3, leerjaar 3*	87	65	26
mbo, niveau 4, leerjaar 4	93	79	50

* = aantal leerlingen is kleiner dan 300, maar groter dan 100

Resultaten Rekentoets VO

De resultaten van de Rekentoets VO staan weergegeven in onderstaande tabel. Voor vmbo klas 4 kon geen uitsplitsing gemaakt worden naar leerweg. Het gevolg is dat het percentage zowel betrekking heeft op BB-, KB- als GT-leerlingen. Dat maakt een vergelijking met de uitkomsten van de diagnostische toetsen lastig. Voor havo en vwo kan deze vergelijking wel worden gemaakt. Daaruit komt naar voren dat het percentage vwo-leerlingen dat 3 F lijkt te halen nagenoeg overeen komt. Voor havo ligt het percentage deelnemende leerlingen dat 3 F haalt volgens de Rekentoets VO zo'n 10 procent lager dan volgens de diagnostische toetsen. Dit verschil zou te wijten kunnen zijn aan een verschil in afnameconditie. De Rekentoets VO is digitaal afgenomen terwijl de diagnostische toetsen een papieren toets is. In het algemeen worden digitale toetsen als iets moeilijker ervaren dan papieren toetsen.

tabel 4.2.10: percentage deelnemende leerlingen dat de verschillende referentieniveaus haalt in het voortgezet onderwijs (meetinstrument is Rekentoets vo)

Rekenen	2F	3F
vmbo BB, KB en GT, klas 4	48	
havo, klas 5		59
vwo, klas 6		92

4.3 Vergelijking resultaten over jaren heen

basisonderwijs

Voor het basisonderwijs zijn in het kader van het project JPON effectschattingen uitgevoerd om na te gaan hoe de leerlingen uit verschillende jaren presteren ten opzichte van elkaar. De resultaten zijn weergegeven in onderstaande tabellen. Daaruit blijkt dat zowel voor rekenen als voor begrijpend lezen de leerlingen in 2011 significant beter presteren dan in 2008. De verschillen zijn echter te klein om in statistische termen te kunnen spreken van een effect (Hemker en van Weerden, 2012).

tabel 4.3.1: Jaareffect Rekenen/wiskunde

Variabele: contrast	Model	Effectgrootte	Effectkwalificatie
2009 – 2008	Alleen jaareffect	-0,01	geen effect
2010 – 2009	Alleen jaareffect	0,07	geen effect ^D
2011 - 2010	Alleen jaareffect	0,00	geen effect
Zoende: 2011 - 2008	Alleen jaareffect	0,05	geen effect ^D

B/C/D:geclassificeerd als “geen effect”, maar wel significant: B: $\square = 0,01-0,001$; C: $\square = 0,001-,0001$; D: $\square < 0,0001$

tabel 4.3.2: Jaareffect Taal – Begrijpend lezen

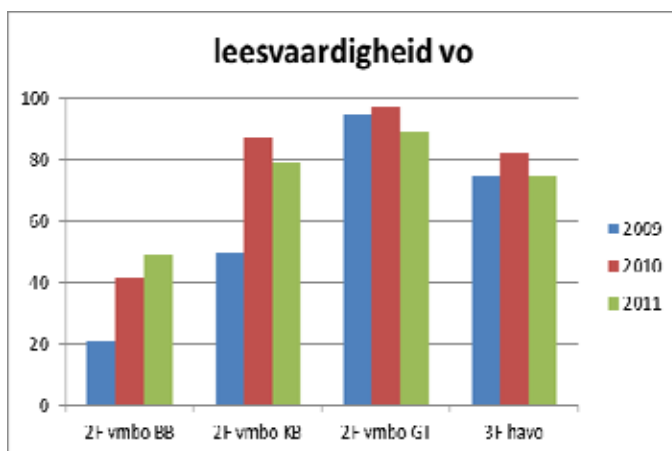
Variabele: contrast	Model	Effectgrootte	Effectkwalificatie
2009 – 2008	Alleen jaareffect	0,04	geen effect ^C
2010 – 2009	Alleen jaareffect	0,04	geen effect ^D
2011 - 2010	Alleen jaareffect	0,07	geen effect ^C
Zoende: 2011 - 2008	Alleen jaareffect	0,15	geen effect ^D

B/C/D:geclassificeerd als “geen effect”, maar wel significant: B: $\square = 0,01-0,001$; C: $\square = 0,001-,0001$; D: $\square < 0,0001$

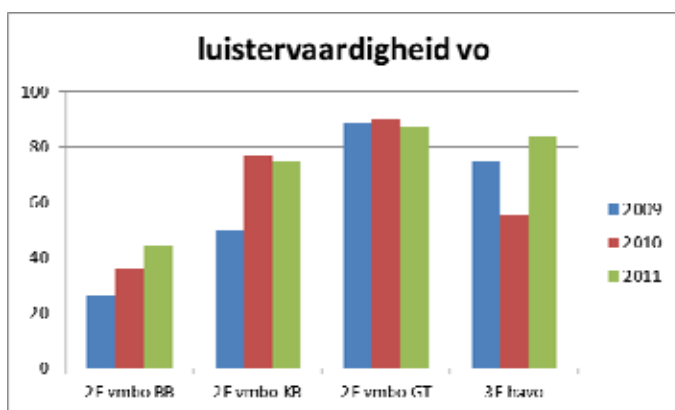
voortgezet onderwijs

Om te bepalen of de leerlingen in 2011 beter of slechter presteren dan in 2009 hebben we de afnamegegevens van de diagnostische toetsen over een periode van drie jaar met elkaar vergeleken. In onderstaande grafieken hebben we het percentage examenleerlingen dat respectievelijk 2F en 3F haalt in 2009, 2010 en 2011 naast elkaar gezet. Uit deze grafieken komt naar voren dat deze percentages per jaar licht variëren. Bij vmbo BB nemen we een stijging waar van het percentage onderzochte leerlingen dat 2 F haalt voor lees- en luistervaardigheid. Voor de andere schooltypen en niveaus zijn geen trends af te leiden. Met uitzondering van vmbo BB, lees- en luistervaardigheid kunnen we dus niet vaststellen dat leerlingen in de periode 2009 – 2011 beter of slechter zijn geworden in taal en rekenen. Datzelfde beeld komt terug in analyses waarbij alle groepen in het VO betrokken zijn, dus niet alleen de leerlingen in de examenjaren. Resultaten van deze analyses zijn opgenomen in bijlage 3. Dat we over het algemeen geen trends hebben kunnen ontdekken is niet verwonderlijk: de wet referentieniveaus is pas per 1 augustus 2010 van kracht gegaan. De effecten van de ingezette intensivering van taal- en rekenonderwijs op taal- en rekenprestaties van laatstejaarsleerlingen zullen pas na verloop van jaren zichtbaar worden.

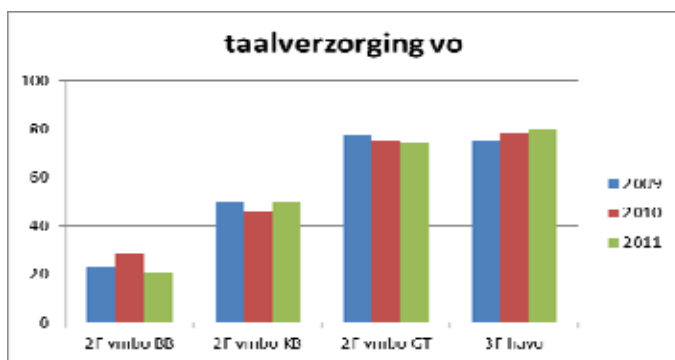
figuur 4.3.3: percentage deelnemende leerlingen uit het VO dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor leesvaardigheid bij de diagnostische toetsen



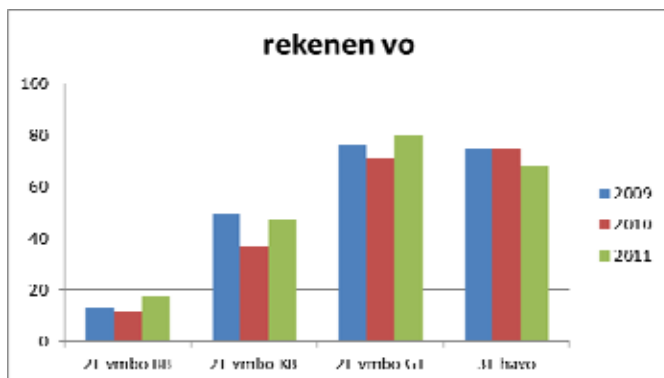
figuur 4.3.4: percentage deelnemende leerlingen uit het VO dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor luistervaardigheid bij de diagnostische toetsen



figuur 4.3.5: percentage deelnemende leerlingen uit het VO dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor taalverzorging bij de diagnostische toetsen



figuur 4.3.6: percentage deelnemende leerlingen uit het VO dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor rekenen bij de diagnostische toetsen

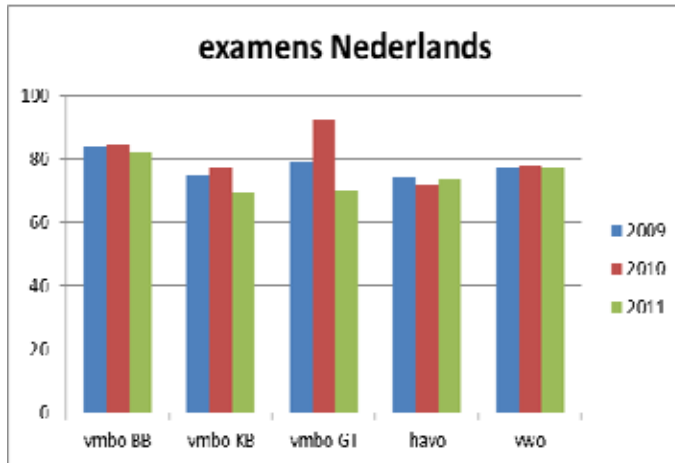


Ter vergelijking hebben we ook op een rijtje gezet wat het percentage geslaagde leerlingen was voor de eindexamens Nederlands in de periode 2009 – 2011. Daarbij zijn we voor vwo, havo, vmbo GT en KB uitgegaan van schriftelijke examens. Voor vmbo BB hebben we de resultaten van het digitale examen gebruikt, omdat de meeste leerlingen deelnamen aan de CBT-variant van het examen en niet aan de papieren variant.

Het percentage geslaagde leerlingen in de periode 2009 – 2011 voor de verschillende schooltypen is afgebeeld in onderstaande grafiek. Deze laat zien dat de percentages geslaagde leerlingen over het algemeen te dicht bij elkaar liggen om te kunnen spreken van een stijgende trend. Verder valt op dat het examen van 2010 voor vmbo GT een afwijkend percentage onvoldoendes vertoont in vergelijking met andere jaren. Dit werd niet veroorzaakt door een uitzonderlijk presterende populatie, maar door het feit dat het examen aanzienlijk makkelijker bleek dan het referentie-examen. Daarvoor kan worden gecompenseerd door een N-term te kiezen tussen de 0,0 en de 2,0. In dit geval heeft CVE voor de laagst mogelijke N-term

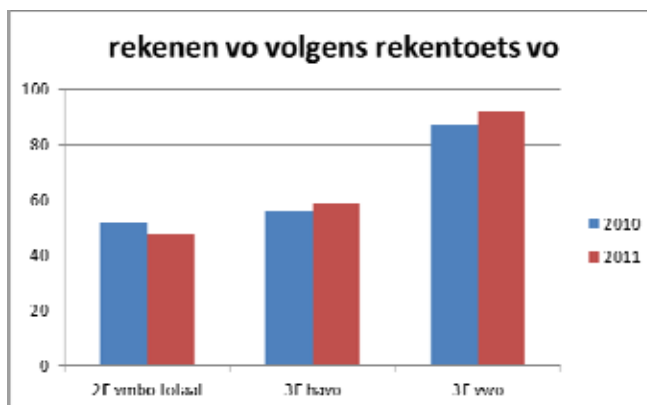
gekozen, namelijk 0,0. Dit was echter niet voldoende om dezelfde eisen te kunnen stellen als in de andere jaren, met als gevolg een uitzonderlijk hoog percentage voldoende. Gegevens over aantallen leerlingen, gemiddelde scores en meetnauwkeurigheid zijn te vinden in bijlage 2.

figuur 4.3.7: percentage examenleerlingen dat een 5,5 haalt of hoger voor een examen Nederlands in de periode 2009 – 2011



Tenslotte hebben we voor het VO ook de prestaties op de Rekentoets VO voor 2010 en 2011 met elkaar vergeleken. De resultaten van die vergelijking zijn te vinden in onderstaande grafiek. Ook uit deze grafiek is geen duidelijke trend af te leiden. De verschillen tussen 2010 en 2011 zijn niet groot en wijken bovendien af van wat we bij de diagnostische toetsen hebben gezien. Dit bevestigt de conclusie dat de rekenprestaties van leerlingen in de periode van 2009 tot en met 2011 min of meer constant lijken.

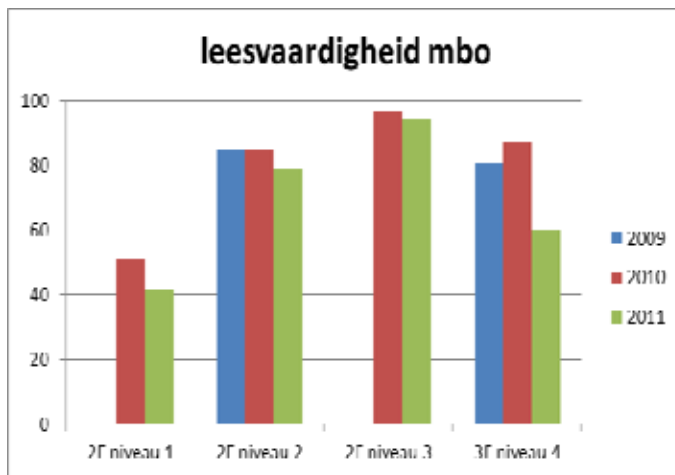
figuur 4.3.8: percentage deelnemende examenleerlingen dat een 5,5 haalt of hoger voor de rekentoets VO in de periode 2010 -2011



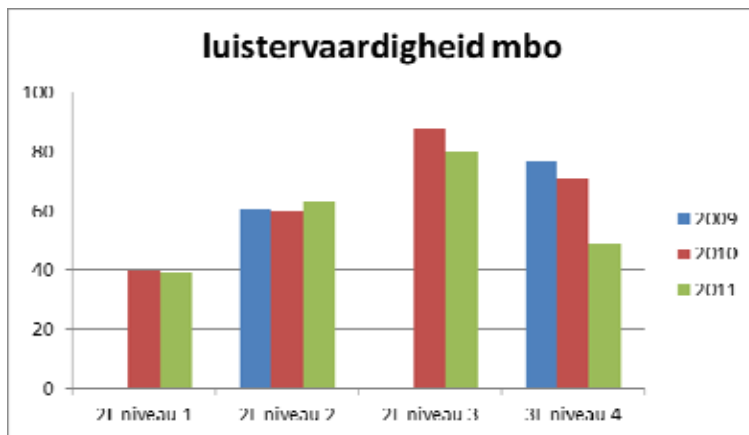
mbo

Ook voor mbo hebben we de prestaties van leerlingen op de diagnostische toetsen in de periode 2009-2011 vergeleken. Omdat in 2009 geen data beschikbaar waren voor de opleidingen op niveau 1 en niveau 3, ontbreken in de grafieken de gegevens voor deze opleidingen. Voor de opleidingen op niveau 2 en 4 zijn de plaatjes wel compleet. In alle gevallen gaat het om leerlingen die in het laatste jaar van hun opleiding zitten. Bekijken we de grafieken globaal, dan kunnen we over het algemeen niet spreken van trends richting verbetering of verslechtering. Daarvoor zijn de verschillen te klein of de patronen te grillig. De analyses die betrekking hebben op alle groepen en niet alleen op de leerlingen in het laatste jaar van hun opleiding bevestigen dit beeld. Deze analyses zijn terug te vinden in bijlage 3.

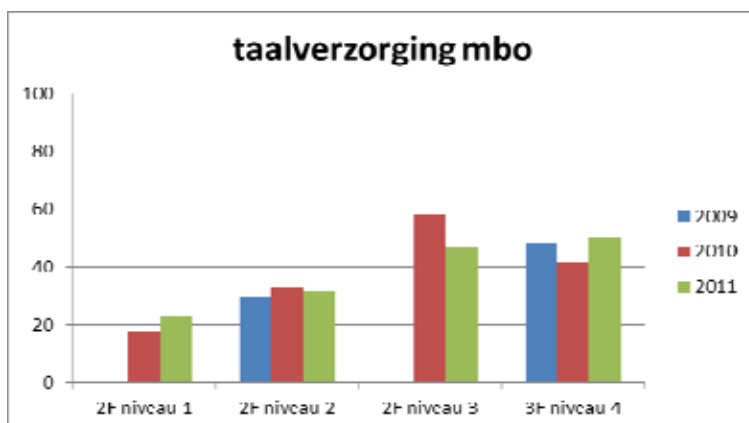
figuur 4.3.9: percentage deelnemende mbo-leerlingen dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor leesvaardigheid bij de diagnostische toetsen



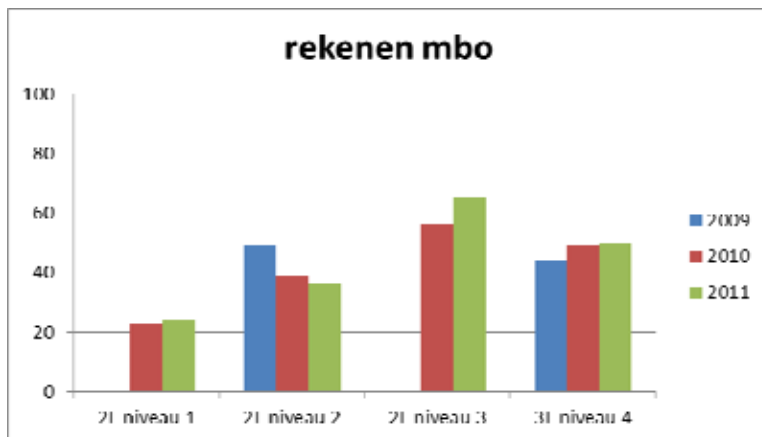
figuur 4.3.10: percentage deelnemende mbo-leerlingen dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor luistervaardigheid bij de diagnostische toetsen



figuur 4.3.11: percentage deelnemende mbo-leerlingen dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor taalverzorging bij de diagnostische toetsen



figuur 4.3.12: percentage deelnemende mbo-leerlingen dat het van toepassing zijnde referentieniveau haalt over de periode 2009 – 2011 voor rekenen bij de diagnostische toetsen



5 Verschillen tussen leerlingen

Naast berekeningen om te bepalen welk percentage leerlingen welk referentieniveau haalt, is ook nagegaan hoe bepaalde subgroepen ten opzichte van elkaar presteren, bijvoorbeeld jongens ten opzichte van meisjes of leerlingen uit de sector techniek ten opzichte van leerlingen uit de andere sectoren. Om dit te kunnen bepalen zijn effectschattingen uitgevoerd waarbij andere kenmerken van leerlingen die van grote invloed zijn op de prestaties constant worden gehouden. Leerlingen uit de ene sector presteren bijvoorbeeld gemiddeld beter dan leerlingen uit de andere sector. Bij effectschattingen voor andere achtergrondvariabelen wordt hier rekening mee gehouden. Een eventueel verschil tussen bepaalde groepen met verschillende achtergrondvariabelen kan dan niet meer worden toegeschreven aan een verschil in vaardigheid tussen de sectoren. Mocht het bijvoorbeeld zo zijn dat meisjes slechter presteren dan jongens voor rekenen, dan kan dat niet verklaard worden doordat er meer meisjes in de sector zorg zitten en meer jongens in de sector techniek. Bij de berekening van het effect is daar dan al voor gecorrigeerd. We spreken in dit geval ook wel van toegevoegde of gecorrigeerde effecten.

De verschillen in vaardigheid worden gerapporteerd in termen van effectgrootten. Voor de interpretatie van de effectgrootten worden de in de literatuur gebruikelijke kwalificaties gehanteerd (Cohen, 1977). Deze zijn af te lezen in onderstaande tabel. In dit hoofdstuk geven we per toets een overzicht van de verschillen tussen leerlingen.

tabel 5.1: interpretatie van effectgroottes

Effectgrootte	Kwalificatie
-0,8	groot negatief effect
-0,5	matig negatief effect
-0,2	klein negatief effect
0,0	geen effect
0,2	klein positief effect
0,5	matig positief effect
0,8	groot positief effect

Tenslotte hebben we berekend of de gevonden effectgroottes significant zijn. De overschrijdingskans die voor VO en mbo is aangehouden bij het berekenen van significantie is 0,05. Is een effect in VO of mbo significant dan is de kans dat het verschil tussen subgroepen door toeval tot stand is gekomen kleiner dan 5 procent. Bij de effectschattingen van de Eindtoets is gekozen voor kleinere overschrijdingskansen. De kans dat de gevonden resultaten aan toeval te wijten zijn, is daar dus nog kleiner.

Verschillen tussen leerlingen op basis van de Eindtoets

In onderstaande tabel geven we aan voor welke achtergrondvariabelen effecten zijn gevonden en of deze effecten klein, matig of groot zijn. De richting van effectgrootte is hier overigens niet gegeven. Deze kan worden opgezocht in bijlage 4. De hieronder gepresenteerde tabel is overgenomen uit het verslag van het project JPON (Hemker & van Weerden, 2012).

tabel 5.2: Effectkwalificaties voor verschillende achtergrondvariabelen op de Eindtoets Basisonderwijs

Variabele	Contrast	Effectkwal. Begrijpend lezen	Effectkwal. Rekenen
Geslacht	Meisjes - Jongens	klein	klein
Leertijd	Vertraagd - Regulier	matig	matig
Stratum	Stratum: lovs 2 – lovs 1	[D]	[D]
	Stratum: lovs 3 – lovs 2	klein	[C]
	Stratum: lovs 3 – lovs 1	klein	[D]
Formatiegewicht	F 0.30 – F 0.00	matig	klein
	F 1.20 – F 0.30	x	[C]
	F 1.20 – F 0.00	matig	klein
Soort toets	NT - EB	groot	groot
Advies VO	vmbo-KB - vmbo-BB	groot	groot
	vmbo-GT - vmbo-KB	matig	matig
	havo - vmbo-GT	groot	groot
	vwo - havo	groot	groot
Regio	Oost - Noord	x	x
	West - Noord	x	[B]
	Zuid - Noord	x	[D]
	West - Oost	x	x
	Zuid - Oost	x	[D]
	Zuid - West	x	x
Urbanisatiegraad	Sterk - Zeer sterk	x	x
	Matig - Sterk	x	x
	Weinig - Matig	x	x
	Niet - Weinig	x	x
Thuis taal	NLs+Ander - Alleen NLs	klein	x
	Alleen ander - Alleen NLs	klein	x
	Alleen ander - NLs+Ander	x	x
	Turks - Alleen NL	matig	x
	Arabisch - Alleen NLs	klein	x
	Surinaams/Antilliaans - Alleen NLs	klein	matig
West-Europees - Alleen NLs	x	x	
Schoolgrootte	Groot - Klein	x	[D]

x = geen effect, niet significant

[B] = geen effect, maar significant met $\alpha < ,01$

[C] = geen effect, maar significant met $\alpha < ,001$

[D] = geen effect, maar significant met $\alpha < ,0001$

Uit de tabel komt naar voren dat er geen toegevoegde effecten worden gevonden voor regio, urbanisatiegraad en schoolgrootte. Deze achtergrondvariabelen hangen dus niet direct samen met de prestaties. Kleine tot matige effecten vonden we voor geslacht, leertijd, stratum, formatiegewicht en thuistaal. Deze achtergrondvariabelen vertonen dus wel samenhang met de prestaties. Hieronder hebben we per achtergrondvariabele aangegeven wat dit betekent.

- Meisjes presteren beter op begrijpend lezen dan jongens maar jongens presteren daarentegen weer beter op rekenen.
- Vertraagde leerlingen presteren slechter op begrijpend lezen en rekenen dan reguliere leerlingen.
- Leerlingen van scholen uit stratum 3 scoren slechter voor begrijpend lezen dan leerlingen van scholen uit stratum 1 of 2. Voor rekenen is er geen effect van stratum. De stratumindeling is een indeling van scholen op basis van het formatiegewicht van de leerlingen en geeft informatie over de sociaal-economische achtergrond van de schoolpopulatie. We hebben de indeling in strata overgenomen van de indeling die gebruikt is bij het leerlingvolgsysteem van Cito. Deze is als volgt:
 - Stratum 1 LOVS: percentage leerlingen met formatiegewicht 0,30 of 1,20 van 0 tot 10 procent (0 – 9,99%);
 - Stratum 2 LOVS: percentage leerlingen met formatiegewicht 0,30 of 1,20 van 10 tot 25 procent (10 – 24,99%);
 - Stratum 3 LOVS: percentage leerlingen met formatiegewicht 0,30 of 1,20 van 25 tot en met 100 procent (25 – 100%).

- Leerlingen met formatiegewicht 0,3 en 1,2 presteren slechter dan leerlingen zonder formatiegewicht. De verschillen tussen beide deelpopulaties zijn groter voor begrijpend lezen dan voor rekenen. Er zijn echter geen significante verschillen tussen leerlingen met formatiegewicht 0,3 en 1,2. Leerlinggewichten zijn getallen die een indicatie geven van het opleidingsniveau van de ouders.
 - Het gewicht 0,00 wordt toegekend aan leerlingen van wie één van de ouders of beide ouders voortgezet onderwijs heeft afgerond boven het niveau vmbo basis- of kaderberoepsgerichte leerweg.
 - Het gewicht 0,30 wordt toegekend aan leerlingen van wie beide ouders of de ouder die belast is met de dagelijkse verzorging maximaal lbo/vbo, praktijkonderwijs of de vmbo basis- of kaderberoepsgerichte leerweg heeft afgerond.
 - Het gewicht 1,20 wordt toegekend aan leerlingen van wie één van de ouders maximaal basisonderwijs heeft afgerond of (v)so-zmlk en van wie de andere ouder hetzelfde opleidingsniveau heeft óf maximaal lbo/vbo, praktijkonderwijs of vmbo basis- of kaderberoepsgerichte leerweg.
- Leerlingen waar thuis een andere taal (Turks, Arabisch of Surinaams/Antilliaans) wordt gesproken dan Nederlands presteren slechter op begrijpend lezen dan leerlingen waar thuis uitsluitend Nederlands wordt gesproken. Dit toegevoegde effect wordt niet gevonden voor leerlingen met een andere West-Europese thuistaal. Het maakt overigens niet uit of in de gezinnen waarin Turks, Arabisch of Surinaams/Antilliaans wordt gesproken, daarnaast ook nog Nederlands wordt gepraat of niet. Voor rekenen vonden we geen toegevoegde effecten voor thuistaal met uitzondering van leerlingen die thuis Surinaams/Antilliaans spreken. Zij presteren slechter dan leerlingen bij wie thuis uitsluitend Nederlands wordt gesproken.

Voor de achtergrondvariabelen Advies VO en Soort toets hebben we tenslotte matige tot grote toegevoegde effecten gevonden. Dit betekent dat de prestaties van leerlingen in hoge mate samenhangen met het soort toets dat de leerling heeft gemaakt en met het advies voor het VO dat uit de toets naar voren komt. Deze resultaten liggen in de lijn der verwachting. De Eindtoets is tenslotte bedoeld om te adviseren bij de schoolkeuze van leerlingen. Daarnaast is de Niveautoets bedoeld voor kinderen bij wie de prestaties achterblijven ten opzichte van leeftijdgenoten en dus ligt het ook voor de hand dat de prestaties van deze leerlingen op de Niveautoets verschillen van de prestaties van leerlingen op de reguliere Eindtoets. In bijlage 4 hebben we een tabel opgenomen waarin de grootte van de effectschattingen staat aangegeven. Daarin kan ook worden afgelezen of het om een positief of negatief effect gaat.

Verschillen tussen leerlingen op basis van de diagnostische toetsen

Bij de diagnostische toetsen hebben we in de eerste plaats gekeken naar verschillen in prestaties tussen de verschillende sectoren in het mbo. Daarbij moesten we de sector natuur buiten beschouwing laten, omdat er onvoldoende leerlingen uit deze sector hadden meegedaan om hun prestaties te kunnen vergelijken met leerlingen uit andere sectoren. Ook hebben we vanwege het geringe aantal leerlingen in hogere leerjaren de analyse moeten beperken tot leerlingen uit het eerste leerjaar. Bij de analyses zijn de leerlingen uit de sector techniek als uitgangspunt genomen. Ze vormen als het ware de referentie en vervolgens is bepaald in hoeverre de prestaties van de leerlingen uit de andere sectoren daarvan afwijken. Dit is af te lezen aan de hand van de effectgroottes. In de tabel staan ook de aantallen leerlingen weergegeven (N). Als de effectgroottes niet significant zijn, dan wordt dit aangegeven met een sterretje.

De resultaten van de effectschattingen laten zien dat er bij rekenen een matig negatief effect is voor de sectoren zorg en economie ten opzichte van de sector techniek. Bij luistervaardigheid zien we hetzelfde, alleen is er daar sprake van een klein negatief effect. Dat betekent dat de leerlingen uit de sector techniek vooral beter zijn in rekenen dan leerlingen uit de andere sectoren. Daarnaast presteren ze ook iets beter voor luisteren. Bij taalverzorging zijn de effecten klein. De leerlingen uit de sector economie lijken voor dat vak iets beter te presteren dan de leerlingen uit de sector techniek. Deze resultaten zijn consistent met die van 2010.

tabel 5.3: effectgroottes van sector, gecorrigeerd voor niveau opleiding (1,2 3 of 4)

	leesvaardigheid		luistervaardigheid		taalverzorging		rekenen	
	N	effect-grootte	N	effect-grootte	N	effect-grootte	N	effect-grootte
techniek	2541	referentie	2170	referentie	2194	referentie	2348	referentie
zorg	3007	-0,09	2344	-0,24	2758	0,11	2577	-0,60
economie	4313	-0,04*	3541	-0,21	4048	0,19	3389	-0,43

* = niet significant

Naast effectschattingen voor sector hebben we dit soort analyses ook uitgevoerd om na te gaan hoe leerlingen uit de beroepsopleidende leerweg (BOL) en de beroepsbegeleidende leerweg (BBL) presteerden ten opzichte van elkaar. Daarbij hebben we ons vanwege de leerlingaantallen weer geconcentreerd op leerjaar 1. De resultaten van de analyses zijn weergegeven in onderstaande tabel. Daaruit komt naar voren dat er voor taal weinig verschillen zijn tussen de leerlingen uit de BOL en de BBL. Voor rekenen ligt dat anders. Voor dit vak presteren de leerlingen uit de BBL beter dan de leerlingen uit de BOL. In termen van effectgroottes is er sprake van een klein negatief effect voor de leerlingen uit de BOL. Ook deze resultaten zijn consistent met die van 2010.

tabel 5.4: effectgrootte van leerweg, gecorrigeerd voor niveau opleiding (1, 2, 3 of 4)

	leesvaardigheid		luistervaardigheid		taalverzorging		rekenen	
	N	effect-grootte	N	effect-grootte	N	effect-grootte	N	effect-grootte
BBL	1783	referentie	1417	referentie	1390	referentie	1164	referentie
BOL	10027	-0,19	8212	-0,18	9184	-0,02*	8726	-0,27

* = niet significant

Verschillen tussen leerlingen op basis van de Rekentoets VO

Voor de Rekentoets VO hebben we de prestaties van de jongens afgezet tegen die van de meisjes. Ook dit is gebeurd met behulp van effectschattingen. De resultaten worden weergegeven in onderstaande tabellen.

tabel 5.5: effectgrootte van geslacht voor eindexamenleerlingen van het vmbo

rekenen vmbo totaal		
	N	effect-grootte
jongens	1060	referentie
meisjes	826	-0.52

* = niet significant

tabel 5.6: effectgrootte van geslacht voor eindexamenleerlingen van het havo

rekenen havo totaal		
	N	effect-grootte
jongens	218	referentie
meisjes	189	-0.41

* = niet significant

tabel 5.7: effectgrootte van geslacht voor eindexamenleerlingen van het vwo

rekenen vwo totaal		
	N	effect-grootte
jongens	208	referentie
meisjes	254	-0.56

* = niet significant

Uit de effectschattingen komt naar voren dat er een klein tot matig negatief effect is voor meisjes ten opzichte van jongens. Dit effect komt tot uiting bij alle schooltypen. Meisjes lijken dus meer moeite te hebben met rekenen dan jongens. Dit blijkt ook uit de resultaten van de Eindtoets Basisonderwijs. Wat het verschil tussen jongens en meisjes betekent ten aanzien van het percentage leerlingen dat het beoogde referentieniveau behaalt, is weergegeven in tabel 4.2 in bijlage 4. Daaruit komt naar voren dat er bij vmbo en havo respectievelijk 27 en 13 procent meer jongens het beoogde referentieniveau bereikt dan meisjes.

Bij vwo bedraagt het verschil tussen jongens en meisjes slechts 6 procent. Dat komt omdat de toets 3F relatief makkelijk is voor de vwo-populatie.

6 Voorbehoud

Hoewel het onderzoek dat aan dit verslag ten grondslag ligt zorgvuldig is opgezet, moeten de resultaten om de volgende redenen met voorzichtigheid worden geïnterpreteerd:

- De gerapporteerde prestaties hangen samen met de normering die op de verschillende toetsen is toegepast. Pas als toetsen taal en rekenen die in de nabije toekomst een verplichtend karakter hebben een aantal jaren operationeel zijn, kan CVE een definitieve normering vaststellen. Als deze definitieve normering strenger of soepeler uitvalt dan de voorlopige normering, zal dit van invloed zijn op de gerapporteerde prestaties.
- De referentieniveaus beschrijven gewenste kennis- en vaardigheidsniveaus aan het einde van een onderwijsperiode. Alle toetsresultaten zijn echter al in het voorjaar verzameld. Mogelijk zijn de prestaties daarom enigszins onderschat.
- Met uitzondering van de Eindtoets basisonderwijs zijn aan de toetsen voor leerlingen in deze implementatiefase geen belangrijke beslissingen verbonden. Dit kan een negatief effect hebben op de motivatie en de prestaties van leerlingen, omdat er dikwijls ook nog geen gericht onderwijsaanbod heeft plaatsgevonden.
- Er was niet bij alle meetinstrumenten sprake van een volledig representatieve steekproef.
- De Eindtoets Basisonderwijs van 2011 is inhoudelijk niet gebaseerd op de referentieniveaus. Daarnaast zijn er nog geen officiële standaarden voor 1F bepaald.

Literatuur

Beguïn, A. (2010) *Concept tijdpad en globale opzet voor de standaardbepaling van de referentieniveaus*. Arnhem: Cito

Cito (2009) *Monitor taal en rekenen. Eerste meting: een indicatie van leerprestaties in termen van het referentiekader*. Arnhem: Cito

Cohen, J., (1977) *Statistical power analysis for the behavioral sciences*. Londen: Academic press, inc.

Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010) *COTAN Beoordelingssysteem voor de kwaliteit van tests*. Zaandam: Heijnis & Schipper

Expertgroep Doorlopende Leerlijnen (1) (2008) *Over de drempels met rekenen, Consolideren, onderhouden, gebruiken en verdiepen; onderdeel van de eindrapportage van de Expertgroep Doorlopende leerlijnen Taal en Rekenen*. Enschede

Expertgroep Doorlopende Leerlijnen (2) (2008) *Over de drempels met taal, de niveaus voor de taalvaardigheid; onderdeel van de eindrapportage van de Expertgroep Doorlopende leerlijnen Taal en Rekenen*. Enschede

Expertgroep Doorlopende Leerlijnen (3) (2008) *Over de drempels met taal en rekenen; hoofdrapport van de Expertgroep Doorlopende leerlijnen Taal en Rekenen*. Enschede

Expertgroep Doorlopende Leerlijnen (2009) *Een nadere beschouwing, over de drempels met taal en rekenen*. Enschede

Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2009). *Referentiekader taal en rekenen. De referentieniveaus*. Enschede: SLO

Hemker, B. T. & Weerden, J. van (2012) *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2011 (in voorbereiding)*. Arnhem: Cito

Til, A. van (2011) *Verantwoording meting taal en rekenen 2010*. Arnhem: Cito

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

Verhelst, N.D. & Glas, C.A.W. (1995) The one parameter logistic model. In: G.H. Fischer & I.W. Molenaar (Eds.). *Rasch models: Foundations, recent developments and applications* (pp. 215-239). New York: Springer

Voesten, R. (red.) (2011) *Inhoudsverantwoording diagnostische toetsen taal en rekenen*. Arnhem: Cito

Voesten, R. (2011) *Verslag van de pilotafname rekenen 2011*. Arnhem: Cito

Bijlage 1: Aantallen leerlingen diagnostische toetsen

tabel 1.1: aantal leerlingen voor de diagnostische toetsen taal en rekenen 2011 totaal

	1/2F	2F	3F	totaal
lezen	34.462	44.353	18.075	96.890
luisteren	17.462	24.807	8.187	50.456
taalverzorging	34.462	44.353	18.075	96.890
rekenen	43.134	50.390	23.190	116.714

tabel 1.2: aantal leerlingen VO voor de diagnostische toetsen taal en rekenen 2011

	1/2F	2F	3F	totaal
lezen	28642	30498	12461	71.601
luisteren	14628	13835	3849	32.312
taalverzorging	28192	31022	12716	71.930
rekenen	35530	35994	18177	89.701

tabel 1.3: aantal leerlingen mbo voor de diagnostische toetsen taal en rekenen 2011

	1/2F	2F	3F	totaal
lezen	3915	12261	5220	21.396
luisteren	1782	9695	4158	15.635
taalverzorging	3903	10964	4888	19.755
rekenen	3836	10664	3948	18.448

tabel 1.4: aantal leerlingen voortgezet onderwijs voor het onderdeel leesvaardigheid, uitgesplitst per leerjaar

schooltype	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen
	klas 1	klas 2	klas 3	klas 4	klas 5	klas 6
praktijkonderwijs	71	117	142	161	37	3
vmbo BB+	1690	913	874	274	0	0
vmbo BB	691	410	1824	909	4	1
vmbo KB	2454	1594	4087	1474	1	0
vmbo GL/TL	5492	3322	5612	2881	1	0
havo	5843	2795	5235	4380	1331	1
atheneum	3439	2078	3651	2287	1036	159
gymnasium	1434	950	841	567	152	26

tabel 1.5: aantal leerlingen mbo voor het onderdeel leesvaardigheid, uitgesplitst per leerjaar

schooltype	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen
	klas 1	klas 2	klas 3	klas 4
mbo, niveau 1	574	104	73	42
mbo, niveau 2	3571	731	131	161
mbo, niveau 3	2772	525	404	201
mbo, niveau 4	5489	939	853	649

tabel 1.6: aantal leerlingen mbo voor het onderdeel leesvaardigheid, uitgesplitst per leerweg en leerjaar

	BBL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	181	72	38	40
mbo, niveau 2	946	265	71	44
mbo, niveau 3	478	130	119	55
mbo, niveau 4	178	78	58	101
	BOL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	314	25	28	2
mbo, niveau 2	2474	431	49	104
mbo, niveau 3	2166	354	271	111
mbo, niveau 4	5074	818	772	519

tabel 1.7: aantal leerlingen mbo voor het onderdeel leesvaardigheid, uitgesplitst per sector en leerjaar

	techniek			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	111	8	0	1
mbo, niveau 2	945	118	3	1
mbo, niveau 3	525	73	41	2
mbo, niveau 4	961	66	256	132
	zorg en welzijn			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	98	12	20	1
mbo, niveau 2	752	158	75	55
mbo, niveau 3	601	138	140	136
mbo, niveau 4	1556	529	214	257
	economie			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	90	9	10	0
mbo, niveau 2	1140	95	1	69
mbo, niveau 3	993	173	34	10
mbo, niveau 4	2090	241	239	17
	natuur			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	55	55	37	40
mbo, niveau 2	119	146	39	7
mbo, niveau 3	73	17	106	19
mbo, niveau 4	36	5	3	74

tabel 1.8: aantal leerlingen voortgezet onderwijs voor het onderdeel luistervaardigheid, uitgesplitst per leerjaar

schooltype	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen
	klas 1	klas 2	klas 3	klas 4	klas 5	klas 6
praktijkonderwijs	49	61	86	103	23	1
vmbo BB+	990	567	687	185	0	0
vmbo BB	343	279	1333	698	4	0
vmbo KB	1252	940	2738	1023	1	0
vmbo GL/TL	2683	1762	3439	1342	1	1
havo	1981	835	1899	1573	473	0
atheneum	1091	660	1238	433	283	70
gymnasium	388	195	237	163	31	1

tabel 1.9: aantal leerlingen mbo voor het onderdeel luistervaardigheid, uitgesplitst per leerjaar

schooltype	aantal leerlingen	aantal leerlingen	aantal leerlingen	aantal leerlingen
	klas 1	klas 2	klas 3	klas 4
mbo, niveau 1	465	94	51	39
mbo, niveau 2	2632	614	78	75
mbo, niveau 3	2166	378	351	207
mbo, niveau 4	4823	538	715	544

tabel 1.10: aantal leerlingen mbo voor het onderdeel luistervaardigheid, uitgesplitst per leerweg en leerjaar

	BBL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	157	70	36	37
mbo, niveau 2	728	217	37	26
mbo, niveau 3	391	94	112	55
mbo, niveau 4	141	59	50	77
	BOL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	263	20	15	2
mbo, niveau 2	1790	365	32	47
mbo, niveau 3	1667	251	224	119
mbo, niveau 4	4492	432	646	443

tabel 1.11: aantal leerlingen mbo voor het onderdeel luistervaardigheid, uitgesplitst per sector en leerjaar

	techniek			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	100	6	0	1
mbo, niveau 2	756	85	1	1
mbo, niveau 3	454	43	35	2
mbo, niveau 4	860	81	188	111
	zorg en welzijn			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	70	5	2	1
mbo, niveau 2	474	107	25	1
mbo, niveau 3	457	86	121	140
mbo, niveau 4	1343	324	206	181
	economie			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	82	9	10	0
mbo, niveau 2	825	77	1	60
mbo, niveau 3	730	108	34	10
mbo, niveau 4	1904	68	196	16
	natuur			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	57	54	35	37
mbo, niveau 2	74	150	39	7
mbo, niveau 3	47	17	106	19
mbo, niveau 4	7	2	6	82

tabel 1.12: aantal leerlingen voortgezet onderwijs voor het onderdeel taalverzorging, uitgesplitst per leerjaar

schooltype	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4	aantal leerlingen klas 5	aantal leerlingen klas 6
praktijkonderwijs	68	116	136	152	34	1
vmbo BB+	1701	888	867	282	0	0
vmbo BB	706	378	1859	927	4	1
vmbo KB	2360	1509	4192	1484	1	0
vmbo GL/TL	5416	3238	5622	3014	1	1
havo	5638	3221	5220	4282	1769	1
atheneum	3355	2169	3753	2066	1092	222
gymnasium	1300	853	868	558	189	31

tabel 1.13: aantal leerlingen mbo voor het onderdeel taalverzorging, uitgesplitst per leerjaar

schooltype	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	444	103	72	40
mbo, niveau 2	2930	515	124	143
mbo, niveau 3	2343	509	215	215
mbo, niveau 4	5387	1029	841	447

tabel 1.14: aantal leerlingen mbo voor het onderdeel taalverzorging, uitgesplitst per leerweg en leerjaar

	BBL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	128	72	38	39
mbo, niveau 2	696	118	67	35
mbo, niveau 3	357	103	48	55
mbo, niveau 4	209	97	61	63
	BOL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	243	21	28	1
mbo, niveau 2	2108	370	44	97
mbo, niveau 3	1875	365	159	126
mbo, niveau 4	4958	889	755	364

tabel 1.15: aantal leerlingen mbo voor het onderdeel taalverzorging, uitgesplitst per sector en leerjaar

techniek				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	63	6	0	0
mbo, niveau 2	707	98	2	1
mbo, niveau 3	463	75	19	0
mbo, niveau 4	961	69	240	81
zorg en welzijn				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	109	16	21	1
mbo, niveau 2	517	148	66	51
mbo, niveau 3	530	157	88	151
mbo, niveau 4	1602	590	225	196
economie				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	81	7	10	0
mbo, niveau 2	1101	95	2	69
mbo, niveau 3	843	158	36	13
mbo, niveau 4	2023	277	235	20
natuur				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	45	56	37	39
mbo, niveau 2	116	34	35	7
mbo, niveau 3	59	13	18	17
mbo, niveau 4	33	13	3	8

tabel 1.16: aantal leerlingen voortgezet onderwijs voor het onderdeel rekenen, uitgesplitst per leerjaar

schooltype	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4	aantal leerlingen klas 5	aantal leerlingen klas 6
praktijkonderwijs	46	87	112	120	22	1
vmbo BB+	1908	941	833	251	0	0
vmbo BB	733	441	1666	721	9	0
vmbo KB	2789	1894	3742	1230	4	0
vmbo GL/TL	7206	4556	6558	2846	2	1
havo	7269	5015	7101	5342	1181	1
atheneum	4663	3183	6142	2592	1359	306
gymnasium	2299	1419	1758	570	229	58

tabel 1.17: aantal leerlingen mbo voor het onderdeel rekenen, uitgesplitst per leerjaar

schooltype	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	387	76	35	33
mbo, niveau 2	2421	475	78	96
mbo, niveau 3	2101	474	220	185
mbo, niveau 4	5438	985	608	439

tabel 1.18: aantal leerlingen mbo voor het onderdeel rekenen, uitgesplitst per leerweg en leerjaar

	BBL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	117	54	31	33
mbo, niveau 2	662	127	43	29
mbo, niveau 3	218	77	56	48
mbo, niveau 4	167	75	48	74
	BOL			
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	206	18	3	0
mbo, niveau 2	1665	313	29	61
mbo, niveau 3	1813	365	155	99
mbo, niveau 4	5042	865	541	344

tabel 1.19: aantal leerlingen mbo voor het onderdeel rekenen, uitgesplitst per sector en leerjaar

techniek				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	69	8	0	0
mbo, niveau 2	709	67	1	1
mbo, niveau 3	398	31	11	2
mbo, niveau 4	1172	239	134	109
zorg en welzijn				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	51	4	2	0
mbo, niveau 2	450	98	27	13
mbo, niveau 3	413	145	64	136
mbo, niveau 4	1663	430	212	172
economie				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	67	9	0	0
mbo, niveau 2	710	103	2	50
mbo, niveau 3	768	194	37	12
mbo, niveau 4	1844	195	170	63
natuur				
	aantal leerlingen klas 1	aantal leerlingen klas 2	aantal leerlingen klas 3	aantal leerlingen klas 4
mbo, niveau 1	42	42	30	33
mbo, niveau 2	65	67	38	7
mbo, niveau 3	27	11	46	18
mbo, niveau 4	2	5	3	32

tabel 1.20: aantal leerlingen Rekentoets VO 2011

	vmbo	havo	vwo	overig	totaal
2F-versies	3017	13	125	405	3560
3F-versies	42	871	852	682	2447
totaal	3059	884	977	1087	6007

Bijlage 2: Eigenschappen van de gebruikte toetsen

Deze bijlage geeft een aantal statistische eigenschappen van de gebruikte meetinstrumenten. Het gaat hierbij om het aantal leerlingen dat de toets gemaakt heeft, het aantal opgaven waaruit de toets bestaat, de gemiddelde score, de standaarddeviatie en de betrouwbaarheid.

tabel 2.1: beschrijvende gegevens van de reguliere Eindtoets basisonderwijs 2011 exclusief wereldoriëntatie

	Maximum score	gemiddelde score	standaarddeviatie	betrouwbaarheid
Totaal	200	148,36	25,27	0,95
Taal	100	75,78	11,61	0,89
rekenen-wiskunde	60	42,74	10,68	0,91
studievaardigheden	40	29,84	5,99	0,82

tabel 2.2: beschrijvende gegevens bij het onderdeel lezen van de diagnostische toetsen taal en rekenen 2011

Toets	Maximumscore	Gemiddelde score	Standaarddeviatie	Standaardmeetfout	Betrouwbaarheid
1/2F	35	23,31	5,16	2,40	0,78
2F	36	22,98	5,65	1,91	0,80
3F	40	27,53	5,77	2,59	0,80

tabel 2.3: beschrijvende gegevens bij het onderdeel luisteren van de diagnostische toetsen taal en rekenen 2011

Toets	Maximumscore	Gemiddelde score	Standaarddeviatie	Standaardmeetfout	Betrouwbaarheid
1/2F	41	28,91	5,54	2,63	0,77
2F	45	31,36	6,40	2,72	0,82
3F	40	27,13	5,02	2,63	0,73

tabel 2.4: beschrijvende gegevens bij het onderdeel taalverzorging van de diagnostische toetsen taal en rekenen 2011

Toets	Maximumscore	Gemiddelde score	Standaarddeviatie	Standaardmeetfout	Betrouwbaarheid
1/2F	49	34,85	7,81	2,73	0,88
2F	50	34,59	8,07	2,84	0,88
3F	51	36,31	8,68	2,82	0,89

tabel 2.5: beschrijvende gegevens bij het onderdeel rekenen van de diagnostische toetsen taal en rekenen 2011

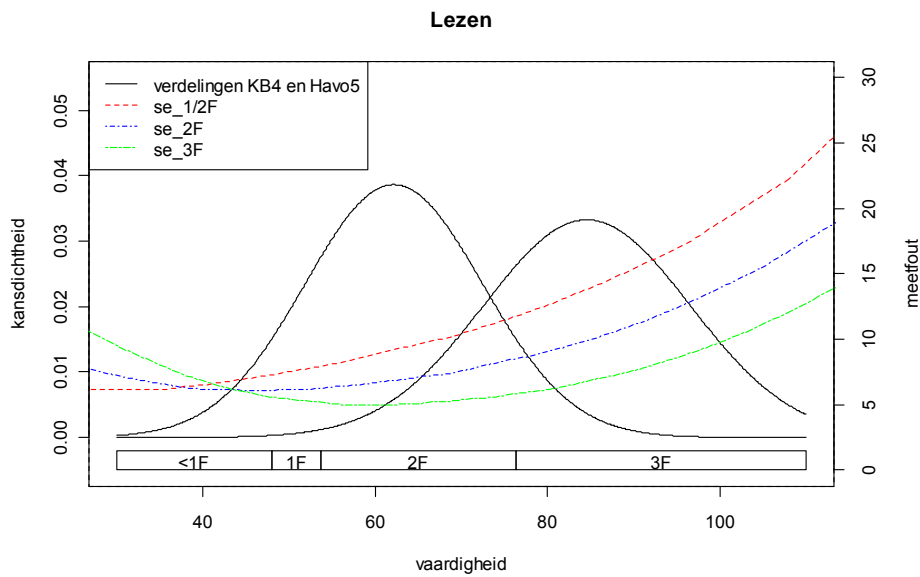
Toets	Maximumscore	Gemiddelde score	Standaarddeviatie	Standaardmeetfout	Betrouwbaarheid
1/2F	60	38,58	11,92	3,25	0,93
2F	60	36,26	11,60	3,31	0,92
3F	60	40,83	9,72	3,13	0,90

De figuren 2.1 tot en met 2.4 geven grafisch weer hoe het gesteld is met de meetnauwkeurigheid van de verschillende toetsen. In elk van deze figuren staat voor iedere toets telkens de grootte van de meetfout afgebeeld.

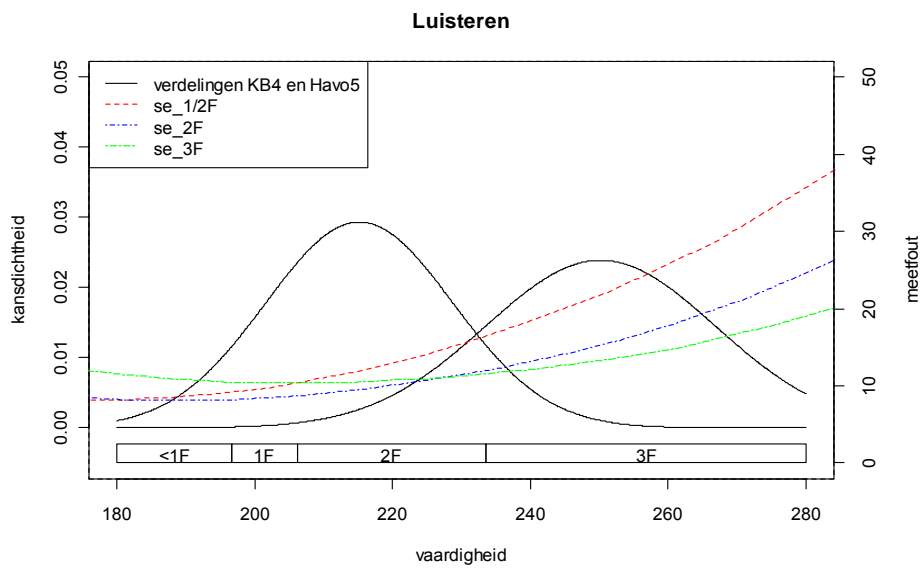
In de onderstaande figuren is de meetfout ofwel standard error (se) op de verticale as getekend bij de geschatte vaardigheden op de horizontale as. Aangezien er steeds drie toetsen voor ieder onderdeel waren, namelijk 1/2F, 2F en 3F, zijn er drie van dergelijke lijnen getekend. In de figuur is bijvoorbeeld te zien dat de meetfout van de 3F-toets kleiner is dan die van de andere toetsen bij hoge vaardigheid. Tevens zijn er twee vaardigheidsverdelingen in iedere figuur getekend, van de KB4-deelnemers en van de Havo5-deelnemers. De Havo5-deelnemers hebben in alle gevallen gemiddeld een hogere vaardigheid dan de

KB4-deelnemers. Tot slot zijn er in de balk onderin de referentieniveaus, zoals toegekend, aan de deelnemers ingetekend.

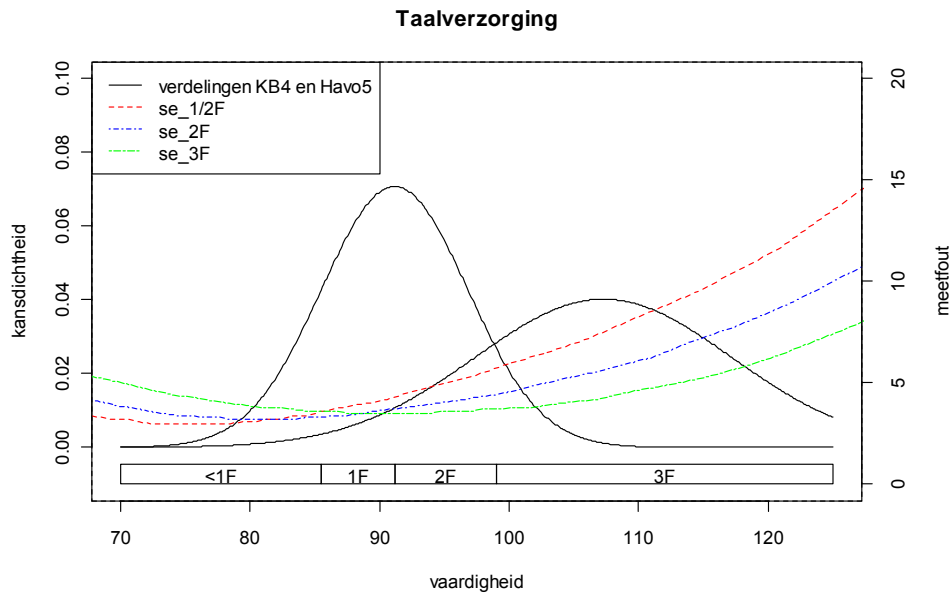
figuur 2.1: meetfout ten opzichte van vaardigheid voor leesvaardigheid



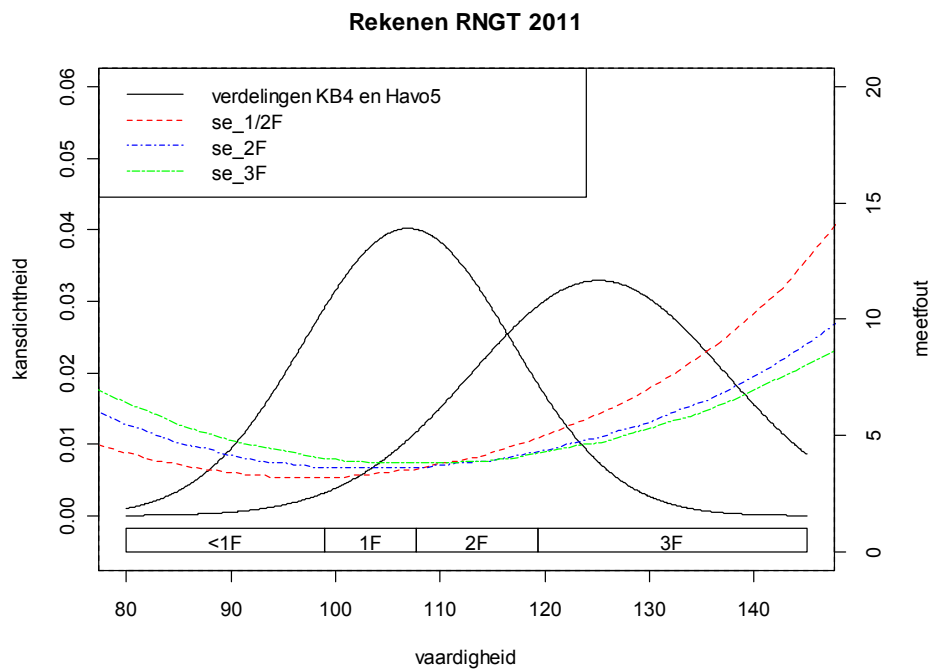
figuur 2.2: meetfout ten opzichte van vaardigheid voor luistervaardigheid



figuur 2.3: meetfout ten opzichte van vaardigheid voor taalverzorging



figuur 2.4: meetfout ten opzichte van vaardigheid voor rekenvaardigheid



tabel 2.6: beschrijvende gegevens bij de Rekentoets VO 2011

Toets	maximum score	gemiddelde score	standaard-deviatie	standaard-meetfout	betrouwbaarheid
2F versie 1	59	28,71	8,42	3,35	0,84
2F versie 2	59	31,95	8,34	3,32	0,84
3F versie 1	60	33,12	9,26	3,04	0,89
3F versie 2	60	35,72	9,42	3,06	0,89

tabel 2.7: beschrijvende gegevens examens Nederlands 2009

Schooltype/Leerweg	VWO	HAVO	VMBO GL/TL	VMBO KB	VMBO BB	VMBO BB
Soort examens	CSE	CSE	CSE	CSE	CSE	CBT
Vaknaam	Nederlands	Nederlands	Nederlands	Nederlands	Nederlands	Nederlands
Totaal aantal VO-kandidaten	37798	47647	51168	28956	3105	23181
Steekproefgrootte	2369	2387	3909	2468	564	14573
Aantal vragen	32	29	36	35	34	-
P'-waarde totale steekproef	0,58	0,53	0,69	0,65	0,63	-
P'-waarde, alleen jongens	0,58	0,51	0,67	0,64	0,63	-
P'-waarde, alleen meisjes	0,59	0,54	0,70	0,67	0,64	-
Maximumscore	50	47	50	50	49	39
Gemiddelde score	29,1	24,8	34,4	32,6	31,1	-
Normeringsterm	1,1	1,3	0,0	0,2	0,6	0,6 - 1,1
Gemiddeld cijfer	6,3	6,0	6,2	6,1	6,3	6,4
Percentage onvoldoendes (<5.5)	23	26	21	25	17	16
Standaardafwijking	6,2	5,1	5,1	5,5	4,7	-
Betrouwbaarheid (asyp. GLB)	0,64	0,51	0,57	0,60	0,68	-
Standaardmeetfout	3,7	3,6	3,4	3,5	2,7	-

tabel 2.8: beschrijvende gegevens examens Nederlands 2010

Schooltype/Leerweg	VWO	HAVO	VMBO GL/TL	VMBO KB	VMBO KB	VMBO BB	VMBO BB
Soort examen	CSE	CSE	CSE	CSE	CBT	CSE	CBT
Vaknaam	Neder- lands	Neder- lands	Neder- lands	Neder- lands	Neder- lands	Neder- lands	Neder- lands
Totaal aantal kandidaten	35920	50169	51171	28891	3667	1568	23012
Steekproefgrootte	22608	32443	39315	20632	3101	1146	17193
Aantal vragen	27	28	38	32	-	29	-
P'-waarde totale steek- proef	0,68	0,47	0,73	0,67	-	0,55	-
P'-waarde, alleen jongens	0,67	0,46	0,72	0,65	-	0,55	-
P'-waarde, alleen meisjes	0,69	0,49	0,75	0,68	-	0,55	-
Maximumscore	47	52	49	50	45	39	39
Gemiddelde score	32,0	24,7	36,0	33,3	-	21,5	-
Normeringsterm	0,0	1,7	0,0	0,2	0,0 - 0,4	1,4	0,6 - 1,3
Gemiddeld cijfer	6,1	6,0	6,6	6,2	6,2	6,4	6,4
Percentage onvoldoendes (<5,5)	22	28	8	23	21	16	16
Standaardafwijking	4,8	5,8	4,3	5,3	-	4,1	-
Betrouwbaarheid (asyp. GLB)	0,48	0,43	0,58	0,63	-	0,59	-
Standaardmeetfout	3,4	4,3	2,8	3,3	-	2,6	-

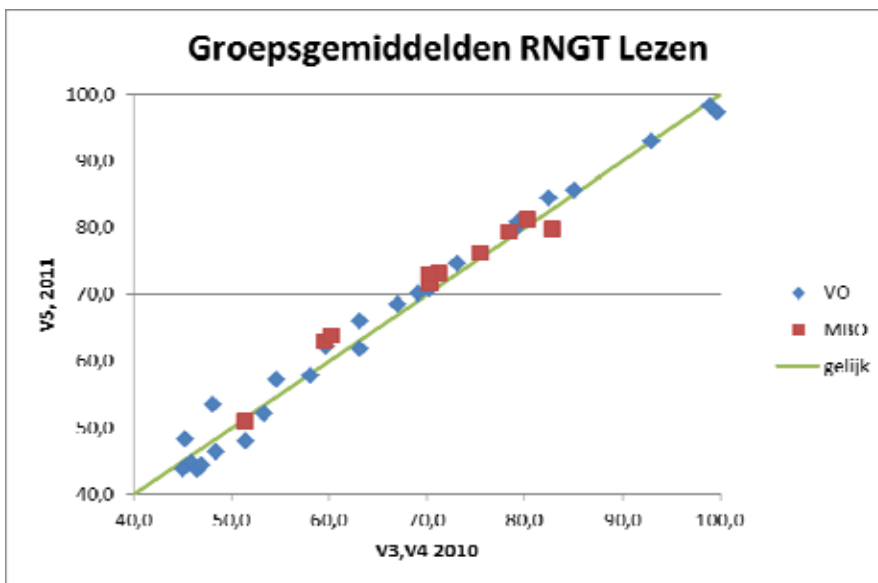
tabel 2.9: beschrijvende gegevens examens Nederlands 2011

Schooltype/Leerweg	VWO	HAVO	VMBO GL/TL	VMBO KB	VMBO KB	VMBO BB	VMBO BB
Soort examen	CSE	CSE	CSE	CSE	CBT	CSE	CBT
Vaknaam	Neder- lands	Neder- lands	Neder- lands	Neder- lands	Neder- lands	Neder- lands	Neder- lands
Totaal aantal kandidaten	37.402	50.771	51.347	28.291	11.187	376	21.935
Steekproefgrootte	26.611	36.800	41.512	13.328	10.418	235	15.659
Aantal vragen	30	29	36	34	-	30	-
P'-waarde totale steek- proef	0,63	0,53	0,65	0,54	-	0,54	-
P'-waarde, alleen jongens	0,62	0,50	0,64	0,53	-	0,54	-
P'-waarde, alleen meisjes	0,64	0,55	0,66	0,55	-	0,56	-
Maximumscore	48	47	49	51	45	39	39
Gemiddelde score	30,3	24,7	31,9	27,4	-	21,2	-
Normeringsterm	0,5	1,3	0,1	1,2	0,0 - 0,4	1,3	0,9 - 1,8
Gemiddeld cijfer	6,2	6,0	6,0	6,0	6,0	6,2	6,3
Percentage onvoldoendes (<5.5)	23,1	26,6	29,8	30,4	26,4	16,6	18,0
Standaardafwijking	5,3	5,4	4,9	5,7	-	3,5	-
Betrouwbaarheid (asyp. GLB)	0,53	0,47	0,63	0,64	-	0,61	-
Standaardmeetfout	3,7	3,9	3,0	3,5	-	2,2	-

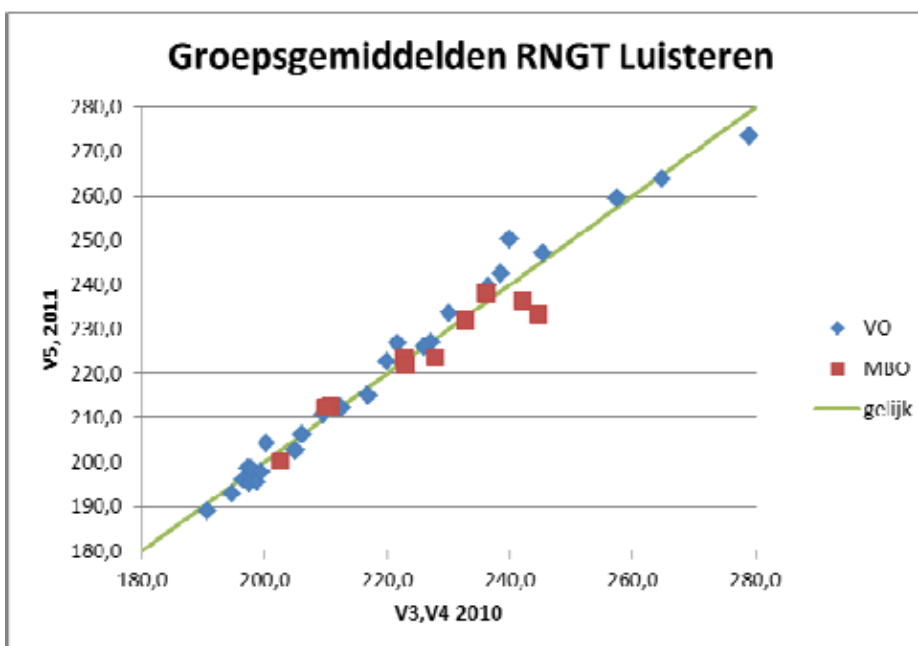
Bijlage 3: Vergelijkingen van prestaties over de jaren heen

Om te bepalen of de leerlingen in 2011 beter of slechter presteerden dan de leerlingen van 2010 zijn de afnamegegevens met elkaar vergeleken. Daarbij is het gemiddelde prestatieniveau van alle leerjaren per opleiding in kaart gebracht, op voorwaarde dat we over voldoende data beschikken. De resultaten betreffen dus niet alleen de leerlingen die in het laatste jaar van de opleiding zitten, maar ook de andere leerjaren. In onderstaande grafieken zijn de resultaten af te lezen. Op de horizontale as staan de groepsgemiddelden van 2010 op de vaardigheidsschaal aangegeven, op de verticale as die van 2011. Indien de groepen in 2010 en in 2011 precies even goed zouden presteren, zouden alle stippen op de diagonale lijn liggen die ter illustratie in de grafiek is opgenomen. Als de stippen vooral boven de diagonale lijn liggen, dan hebben de leerlingen van 2011 beter gepresteerd dan de leerlingen van 2010. Is het omgekeerde het geval, dan hebben de leerlingen van 2010 het beter gedaan. De grafiek geeft een globaal beeld van de groepsgemiddeldes. Men kan aan de stippen niet zien om welke groepen het gaat.

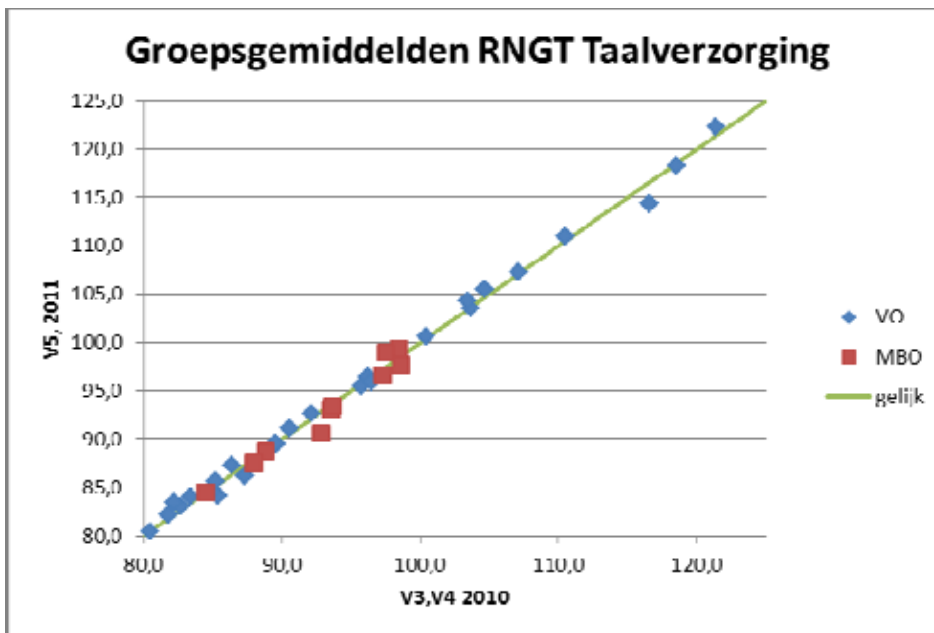
figuur 3.1: vergelijking van groepsgemiddelden op de vaardigheidsschaal voor lezen op de diagnostische toetsen



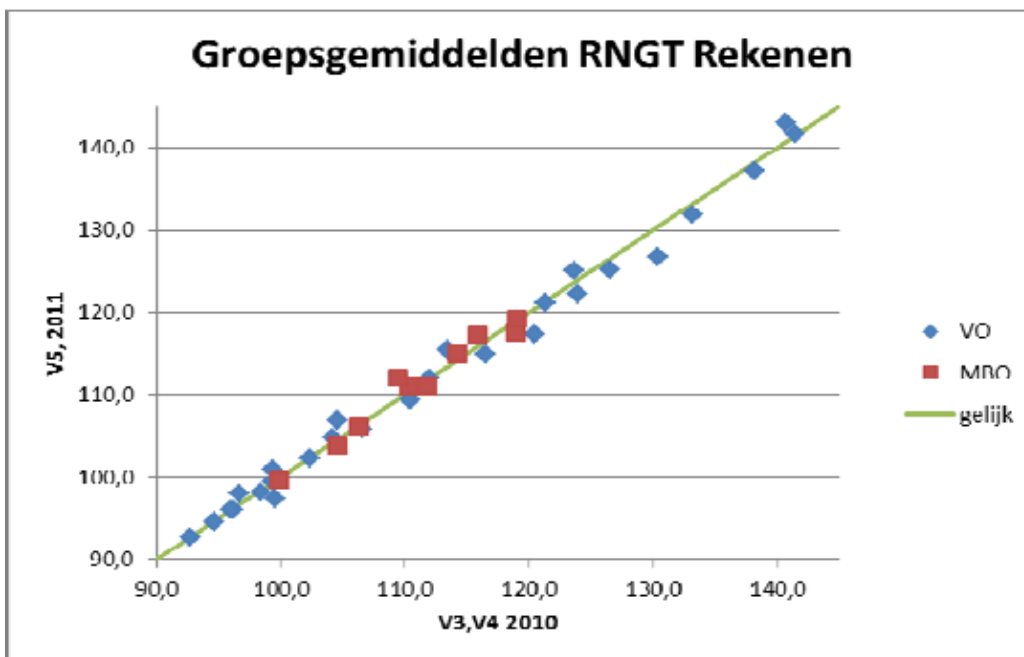
figuur 3.2: vergelijking van groepsgemiddelden op de vaardigheidsschaal voor luisteren op de diagnostische toetsen



figuur 3.3: vergelijking van groepsgemiddelden op de vaardigheidsschaal voor taalverzorging op de diagnostische toetsen



figuur 3.4: vergelijking van groepsgemiddelden op de vaardigheidsschaal voor rekenen op de diagnostische toetsen



Bijlage 4: Achtergrondvariabelen en deelpopulaties

tabel 4.1: effectgroottes van achtergrondvariabelen voor de Eindtoets 2011

Variabele	Contrast	Effectgrootte BL	Effectgrootte rekenen
Geslacht	Meisjes - Jongens	0,20	-0,47
Leertijd	Vertraagd - Regulier	-0,58	-0,68
Stratum	Stratum: lovs 2 – lovs 1	-0,16	-0,10
	Stratum: lovs 3 – lovs 2	-0,22	-0,08
	Stratum: lovs 3 – lovs 1	-0,38	-0,18
Formatiegewicht	F 0.30 – F 0.00	-0,56	-0,47
	F 1.20 – F 0.30	-0,02	0,13
	F 1.20 – F 0.00	-0,59	-0,35
Soort toets	NT - EB	-1,55	-1,70
Advies VO	vmbo-KB - vmbo-BB	0,87	0,86
	vmbo-GT - vmbo-KB	0,50	0,57
	havo - vmbo-GT	0,94	1,01
	vwo - havo	1,33	1,33
Regio	Oost - Noord	0,04	0,03
	West - Noord	0,08	0,07
	Zuid - Noord	0,08	0,10
	West - Oost	0,04	0,04
	Zuid - Oost	0,04	0,07
	Zuid - West	0,00	0,03
Urbanisatiegraad	Sterk - Zeer sterk	-0,01	0,00
	Matig - Sterk	-0,01	-0,01
	Weinig - Matig	-0,03	0,02
	Niet - Weinig	-0,04	-0,05
Thuis taal	NLs+Ander - Alleen NLs	-0,24	0,00
	Alleen ander - Alleen NLs	-0,31	-0,03
	Alleen ander - NLs+Ander	-0,07	-0,04
	Turks - Alleen NLs	-0,53	-0,03
	Arabisch - - Alleen NLs	-0,27	0,00
	Surinaams/Antilliaans - Alleen NLs	-0,47	-0,55
	West-Europees - Alleen NLs	0,04	-0,02
	Schoolgrootte	Groot - Klein	0,04

tabel 4.2: percentage leerlingen dat een beoogd referentieniveau bereikt volgens de Rekentoets VO

Groep, variant, geslacht	aantal leerlingen	beoogd referentieniveau	percentage leerlingen dat het referentieniveau bereikt
vmbo 4, 2F totaal	3017	2F	48%
vmbo 4, 2F, jongens	1060	2F	60%
vmbo 4, 2F, meisjes	826	2F	33%
Havo 5, 3F	871	3F	59%
Havo 5, 3F, jongens	218	3F	67%
Havo 5, 3F, meisjes	189	3F	54%
vwo 6, 3F	852	3F	92%
vwo 6, 3F, jongens	208	3F	94%
vwo 6, 3F, meisjes	254	3F	88%