

Voorwoord/ korte samenvatting

Voor U ligt de rapportage over de Eindtoets Basisonderwijs 2008. De individuele scholen hebben reeds in een vroeg stadium hun individuele resultaten gegroepeerd naar leerdomeinen ontvangen. Dit stelt hen in staat om gericht bij te sturen en de kwaliteit op reken- en taalgebied binnen het basisonderwijs te verhogen.

In dit rapport komen achtereenvolgens aan bod de context en doelstellingen van de nationale eindtoets basisonderwijs, de gevolgde onderzoeksmethodologie en uiteraard de onderzoeksresultaten samen met enkele aanbevelingen.

De belangrijkste onderzoeksresultaten zijn wel dat er net als in 2007 grote prestatieverschillen bestaan tussen de verschillende basisscholen voor wat betreft de verscheidene taal- en rekentoetsen en dat eenzelfde conclusie getrokken kan worden ten aanzien van prestatieverschillen tussen verschillende klassen van individuele scholen. Dat laatste duidt op een belangrijke invloed van de individuele leerkracht op de toetsresultaten. Daarnaast is ook duidelijk de invloed van schoolexterne factoren zichtbaar. De schoolresultaten hangen voor een belangrijk deel samen met de socio-economische context/omgeving van waaruit de school zijn leerlingen betreft. Het educatieniveau van de ouders speelt hierin een belangrijke rol.

Voorts volgt uit de analyse van de afgenomen toetsen dat het berekende nationale eindniveau voor taal en rekenen nagenoeg gelijk is en ten opzichte van 2007 een half punt hoger ligt. Deze hogere score is naar alle waarschijnlijkheid vooral het gevolg van een toetseffect en in mindere mate van een leereffect. Toch is de goede ontwikkelingsrichting ingeslagen, maar het zal nog wel even duren voordat men met de resultaten tevreden kan zijn.

Kortom, het koppelen van schoolinterne en schoolexterne factoren aan de toetsresultaten geeft wederom een schat aan interessante gegevens en inzichten over het presteren van leerlingen en scholen in de complexe context waarbinnen ze moeten functioneren.

Om de leesbaarheid en de referenties te vergemakkelijken zijn alle hoofdstukken voorzien van doorgenummerde conclusies en aanbevelingen.

Een speciale dank gaat uit naar het Examenbureau voor de uitstekende organisatie en samenwerking.

Inhoudsopgave

Voorwoord/korte samenvatting

	<i>blz.</i>
1. <i>Context en doelstellingen</i>	
- <i>Toetscontext</i>	4
- <i>Doelstellingen Eindtoets Basisonderwijs</i>	4
2. <i>Methodologie</i>	
- <i>Onderzoekspopulatie en algemene onderzoeksopzet</i>	6
- <i>De schoolcontext</i>	7
- <i>Verwerking van de data</i>	9
3. <i>Analyseresultaten</i>	
- <i>Respondenten en non respons</i>	10
- <i>Toetsresultaten per school</i>	12
- <i>Toetsresultaten per schoolklas</i>	14
- <i>Toetsresultaten in vergelijking met 2007</i>	16
- <i>Leereffect of toetseffect?</i>	19
- <i>Toetsresultaten en schoolcontext</i>	
• <i>educatieniveau ouders</i>	24
• <i>geboorteland en voertaal thuis</i>	26
• <i>aantal keren doubleren</i>	28
• <i>aantal jaren kleuteronderwijs</i>	29
• <i>gebroken gezinnen</i>	30
• <i>schoolperceptie</i>	31
- <i>Toetsresultaten naar geslacht</i>	32
- <i>Schoolresultaten en doorstroom naar Colegio Arubano(2007)</i>	33
4. <i>Aanbevelingen</i>	35
 <i>Bijlagen</i>	 37

1. Context en doelstellingen

Toetscontext

Systematisch verzameld feitenmateriaal draagt in belangrijke mate bij aan inzichtelijke kennis ten aanzien van ingewikkelde maatschappelijke processen en verschijnselen en mogelijke ontwikkelingen en trends hierbinnen. Objectieve meetgegevens zijn een voorwaarde voor verantwoorde beleidsontwikkeling en mede daarom doet de vraag naar die gegevens zich des te sterker gelden naarmate het desbetreffende maatschappelijk belang groter is. Goed onderwijs wordt wereldwijd gezien als een van de grootste maatschappelijke belangen en kwaliteitsbewaking van de opbrengsten is dan ook een constant zorgpunt.

Dat geldt in bijzondere mate voor Aruba omdat in het Arubaanse onderwijs op alle niveaus sterke inhoudelijke, didactische en organisatorische vernieuwingstendenzen ingezet zijn en worden. Een van de belangrijkste, meest omvangrijke onderwijsvernieuwingen betreft die van het basisonderwijs.

Het behoeft dan ook geen verwondering dat in toenemende mate de behoefte gevoeld wordt om op nationaal niveau over gestandaardiseerde meetgegevens te beschikken betreffende de prestaties van basisschoolleerlingen; in ieder geval met betrekking tot de hoofdleergebieden, zijnde rekenen en taal. Directe aanleiding waren in eerste instantie in het voortgezet onderwijs slecht presterende groepen leerlingen die met relatief goede cijfers het basisonderwijs verlaten hebben en basisschoolleerlingen die vanwege hun relatief hoge leeftijd (door eerder, vroegtijdig doubleren) en niet zozeer vanwege hun prestaties, bevorderd worden naar een hogere klas in het basisonderwijs. Het bredere kader is de kwaliteitsbewaking van de onderwijsuitkomsten aan de basis, het basisonderwijs dus.

Vandaar dat aan het einde van het schooljaar 06/07, op initiatief van de Minister van Onderwijs, Sociale Zaken en Infrastructuur, drs. Marisol Tromp voor alle 6de klassers van het openbare en door de overheid gesubsidieerde basisonderwijs voor het eerst een nationale eindtoets taal en rekenen is gehouden. Inmiddels heeft deze toets een vervolg voor het schooljaar 07/08 gehad en is men druk bezig met de voorbereidingen van de eindtoets voor het schooljaar 08/09.

Doelstellingen Eindtoets Basisonderwijs

1. Inzicht verkrijgen in het huidige nationale niveau voor wat betreft rekenen en taal aan het einde van de 6^{de} klas basisonderwijs.
2. Inzicht verkrijgen in de prestatieverschillen tussen en binnen de reguliere basisscholen (openbaar en gesubsidieerd) voor wat betreft rekenen en taal aan het einde van de 6^{de} klas basisonderwijs.
3. Hulp bieden aan initiatieven tot kwaliteitsverbeteringen voor wat betreft het reken- en taalonderwijs op school- en leerlingniveau in het basisonderwijs.

Bovenstaande doelen zijn nagenoeg ongewijzigd overgenomen uit het rapport "Eindtoets Primair Onderwijs 2007". Daar waar mogelijk, zullen de toetsresultaten uit het schooljaar 07/08 expliciet te vergeleken worden met die van het voorafgaande schooljaar. Herhaalde metingen bij valide toetsen leveren immers meer betrouwbare inzichten op waardoor mogelijkheden geschapen worden de nagestreefde hulp efficiënter en effectiever gestalte te geven.

De vierde doelstelling betreffende 'hulp bieden bij het vaststellen van nationale eindtermen voor taal en rekenen in het basisonderwijs' is weggelaten. Dat is gebeurd op grond van twee rededenen. Enerzijds is duidelijk aangegeven door het Examenbureau dat hoe dan ook de toetsresultaten op geen enkele manier gebruikt zullen worden bij het opstellen van de desbetreffende kerndoelen. Daarnaast is het zo dat de nationale kerndoelen voor rekenen en Nederlands inmiddels door de afdeling Curriculumontwikkeling van Directie Onderwijs zijn ontwikkeld en dat deze binnenkort vastgesteld zullen worden door de minister van Onderwijs, Sociale Zaken en Infrastructuur. De methodeoverstijgende nationale kerndoelen voor rekenen en Nederlands zullen dan ook op korte termijn aan het veld kenbaar gemaakt worden.

Doelstelling 3 kan breed gezien worden. De invoering van een elektronisch leerlingvolgsysteem in het basisonderwijs is inmiddels de planningsfase voorbij. Ook liggen al plannen klaar om naast de eindtoets basisonderwijs voortgangstoetsen af te nemen op klassen lager dan de zesde klas. De bedoeling hiervan is om leerlingen individueel bij te sturen, dit in tegenstelling tot de eindtoets die een eindniveau op school- of landelijk niveau probeert te meten. Toch zal de eindtoets naar verwachting ook op individueel leerlingniveau in de nabije toekomst een rol gaan spelen namelijk bij de doorstroombeslissingen naar het Algemeen Voortgezet Onderwijs. Hij zal dan één van de meerdere evaluatieonderdelen zijn waarnaar gekeken wordt alvorens een doorstroomadvies te formuleren. Eenzelfde functie kan de toets dienen op schoolniveau, bijvoorbeeld bij adviezen of beleidsbeslissingen ten aanzien van onderwijsverbeterplannen op reken- en taalgebied. Voorwaarde hiervoor is wel dat men spreken kan van een valide eindtoets die een eerlijk, betrouwbaar en volledig beeld geeft van de leerresultaten. Een uitgebreide toetsevaluatie zal hier uiteindelijk uitsluitsel over moeten geven.

Helaas heeft die toetsevaluatie nog niet plaatsgevonden. Vandaar dat het voorliggende rapport, net als dat van 2007, zich vooral concentreert op de eerste twee onderzoeksdoelen.

2. Methodologie

Onderzoekspopulatie en algemene onderzoeksopzet

De onderzoekspopulatie bestaat uit de zesde klas leerlingen van alle gesubsidieerde en openbare reguliere basisscholen. Uitgezonderd zijn dus: de International School of Aruba (ISA), de Schakel en het incorporerend basisonderwijs (basisonderwijs voor anderstaligen). Getoetst zijn in totaal 1517 leerlingen variërend in de leeftijd tussen 11 en 15 jaar en verdeeld over 34 scholen behorende tot 4 verschillende schoolbesturen.

Het gehanteerde meetinstrument waarmee kennis en vaardigheden op reken- en taalgebied in kaart gebracht zijn, is ook dit jaar weer door het Examenbureau van Directie Onderwijs ontwikkeld. Getracht is om diverse toetsaspecten als toetsvorm en -omvang, moeilijkheidsgraad, moment en manier van afname, etc. op dezelfde manier gestalte te geven als in het vorig schooljaar. Niet in de laatste plaats om de toetsresultaten onderling zo eerlijk mogelijk te kunnen vergelijken. Het toetsinstrument bestaat daarom weer uit een pakket van een drietal rekentoetsen van elk 20 opgaven en een drietal taaltoetsen van respectievelijk 30, 35 en 35 opgaven. Deze in totaal 160 opgaven zijn zonder uitzondering multiple choice vragen met 4 antwoordmogelijkheden waarvan slechts één correct is.

De toetsontwikkeling is gebeurd binnen twee commissies (één voor taal en één voor rekenen) elk bestaande uit vier docenten uit het basisonderwijs bijgestaan door enkele klankbordleden (waarvan vier uit het voortgezet onderwijs en één uit de inspectie), aangevuld met een toetsconstructiedeskundige en/of onderwijskundige medewerker onder voorzitterschap van het Hoofd Examenbureau. Leidraad bij de ontwikkeling waren de eindtermen, doelen en leerlijnen van de meest gebruikte reken- en taalmethodes in het basisonderwijs en uiteraard het geschatte kennisniveau en de belevingswereld van de Arubaanse basisschoolleerling. Daarnaast is gebruik gemaakt van de ervaringen opgedaan uit de eindtoets 06/07.

Dezelfde hoofddomeinen zijn getoetst, hoewel de keuze uit de subonderdelen niet altijd dezelfde gebleven is. Dat komt omdat er meer reken- en taal-subonderdelen zijn dan dat er in één keer getoetst kunnen worden. Het streven is echter om op termijn alle subonderdelen minstens één keer getoetst te hebben.

Veranderingen in de logistiek van de toetsafnames hebben alleen plaatsgevonden ten aanzien van de toetsduur. In tegenstelling tot het voorafgaande jaar is de toetsduur nu voor alle afzonderlijke reken- en taaltoetsen gelijk aan 60 minuten. De totale toetsduur is daarmee met 15 minuten gereduceerd, de meer omvangrijke taaltoetsen 2 en 3 zijn beide in toetsduur elk met een kwartier ingekort, terwijl voor de tweede rekentoets een kwartier extra bedenktijd uitgetrokken is. Verder zijn de zes toetsen opnieuw, samen met gebruiksaanwijzingen voor de leerling, in twee aparte opgavenboekjes gebundeld en afgenomen in de maand mei, vlak voor het einde van het schooljaar. Een overzicht van de toetsdata, toetsvolgorde en toetsduur is gegeven in tabel 1.

Tabel 1

Organisatie van de reken- en taaltoetsen						
28 mei	taaltoets 1	60 min.	rekentoets 1	60 min.	taaltoets 2	60 min.
29 mei	rekentoets 2	60 min.	taaltoets 3	60 min.	rekentoets 3	60 min.

De antwoorden op de verschillende toetsen zijn door de leerlingen ingevuld op aparte door het CBS ontwikkelde antwoordformulieren.

Een eveneens door het CBS ontwikkelde enquête met betrekking tot achtergrondgegevens van vooral de leerling is samen met de eigenlijke toets aan de onderzoekspopulatie voorgelegd en afgenomen. Deze enquête houdt direct verband met het tweede onderzoeksdoel zijnde het vergelijken van schoolprestaties onderling. Schoolprestaties kunnen namelijk niet los gezien worden van de context waarbinnen ze geleverd worden.

De schoolcontext

De omgeving waarbinnen de scholen opereren en moeten presteren is veelzijdig en complex van natuur. Op de eerste plaats hebben we te maken met verschillende “actoren”, met als hoofdactor uiteraard de leerling zelf als producent van de schoolprestaties. Zijn leerprestaties hangen enerzijds af van individuele gegevenheden als aanleg (IQ), leerstijl en leerhouding (motivatie) en anderzijds van de hulp, begeleiding, sturing (feedback) en controle die hij/zij tijdens het leerproces ontvangt. Hier komen de ander actoren in zicht: de leraar en de school, maar evenzeer het gezin (vooral de ouders) en de vrienden of medeleerlingen.

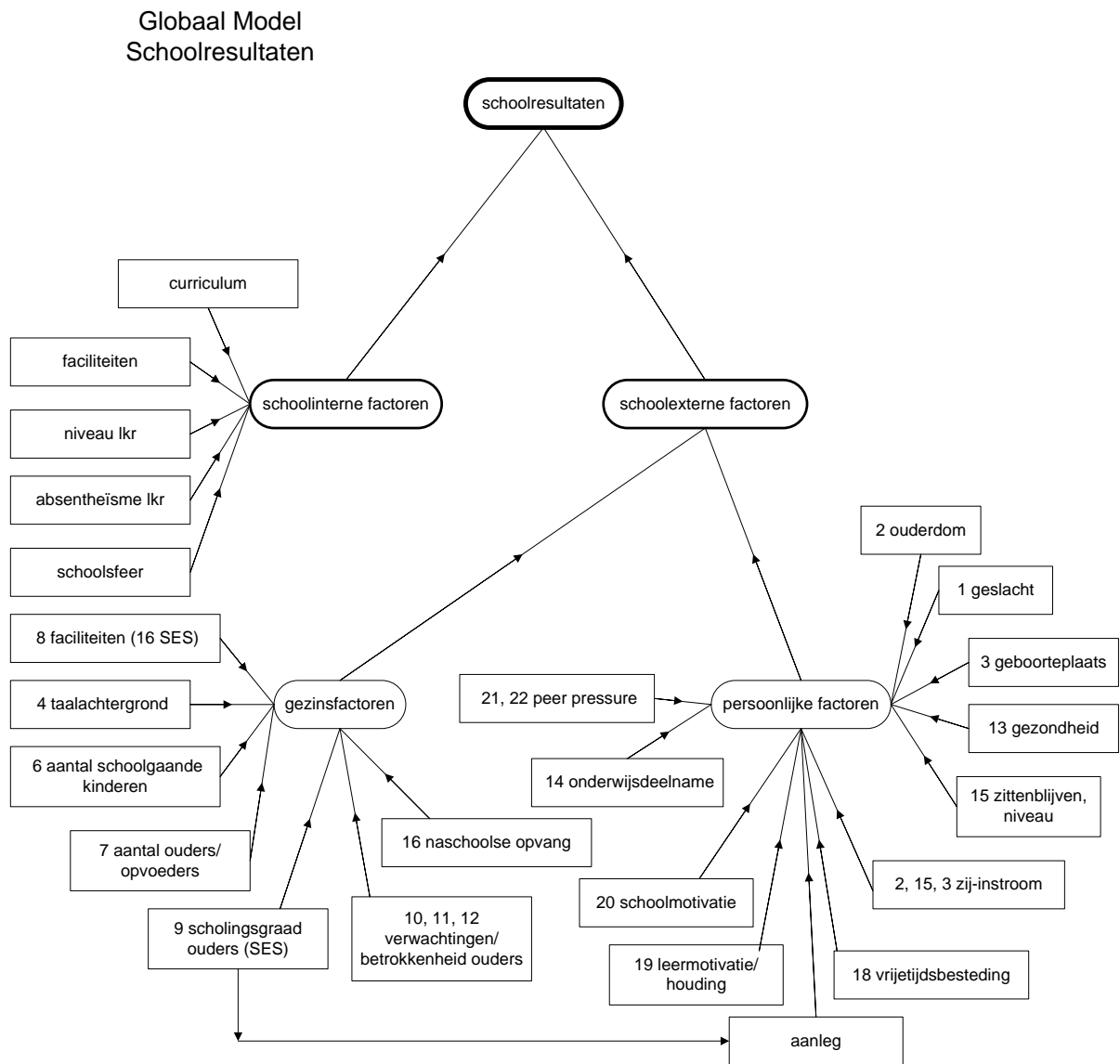
Deze actoren oefenen in het algemeen op tweërlei aard invloed uit op de prestaties van de leerling. Enerzijds via hun voorbeeldfunctie, betrokkenheid en verwachtingspatroon die ze hebben ten aanzien van het leerproces en de leerresultaten van de leerling. Anderzijds door de mogelijkheden die zij bezitten om dat leerproces te sturen, begeleiden, controleren of anderszins hulp te bieden. Dat gebeurt in hun functie als opvoeder en leraar/ begeleider of ‘facilitator’.

Daarnaast hebben we te maken met de socio-economische (financiële) en culturele (normen en waarden) context op diverse niveaus. Op mesoniveau kunnen scholen belangrijke verschillen vertonen t.a.v. de mogelijkheden die zij bezitten om hun onderwijstaak te vervullen. Dat kan te maken hebben met verschillen in personele bezetting of een onevenredige verdeling van de middelen die op macroniveau verkregen worden. Minstens even belangrijk in deze is echter dat de beginsituatie per school al naar gelang de regionale situering of historie (reputatie) sterk kan verschillen. Scholen betrekken hun leerlingen i.h.a. uit een bepaalde regio en die zijn qua sociaaleconomische gezinsstatus (SES) niet altijd homogeen verdeeld. Ook zijn er scholen die minder regionaal gebonden zijn maar hun leerlingen selecteren naar taalachtergrond, religie of SES. Verder zullen scholen verschillen doordat ze verschillende keuzes maken ten aanzien van de gebruikte leermethodes en de gehanteerde didactische/ pedagogische aanpak.

Uit diverse onderzoeksrapporten blijkt steeds weer dat de achtergrondfactoren op microniveau en daarbinnen speciaal de gezinssituatie een belangrijke rol spelen bij de uiteindelijke leerprestaties. De SES is een belangrijke veelvermelde indicator op gezinsniveau, maar niet de enige. Ook de gerelateerde cultureel bepaalde normen en waarden (seksegebonden verwachtingspatronen), de gezinssamenstelling (aantal kinderen, aantal opvoeders) en specifieke gezinssituaties (taalachtergrond, zorgbehoevende gezinsleden, etc.) spelen een rol.

Één manier om de beschreven complexe context in kaart te brengen is de indeling tussen schoolinterne en schoolexterne factoren. Onder schoolinterne contextfactoren vallen zaken als de beschikbare faciliteiten, het curriculum, de kwaliteit van de onderwijs- en leidinggevenden. Schoolexterne contextfactoren behelzen o.a. de gezinsomstandigheden en persoonsgebonden factoren van de leerlingen. Figuur 1 geeft een schematisch overzicht van deze indeling. Bij de verschillende indicatoren voor de schoolexterne factoren staan nummers vermeld die verwijzen naar de overeenkomstige vragen in de enquête vragenlijst. Deze vragenlijst is integraal als bijlage 1 opgenomen. De schoolinterne factoren zijn gemeten aan de hand van een door de inspectie opgestelde uitgebreide vragenlijst.

Een wat algemener model voor schoolprestaties dat niet uitgaat van het verschil tussen schoolinterne en schoolexterne factoren, maar waarin de centrale rol van het gezin meer tot uitdrukking komt, is opgenomen als bijlage 5.



Figuur 1

Tot zover de algemene onderzoeksopzet. De meer specifieke aanpak van de verwerking van de schoolcontextfactoren wordt beschreven in het hoofdstuk "Toetsresultaten en schoolcontext".

Verwerking van de data

De enquêteformulieren, evenals de antwoordformulieren behorende bij de verschillende reken- en taaltoetsen, zijn door het CBS in een aantal fases verwerkt.

De eerste fase bestond uit de visuele controle op volledigheid, zowel voor wat betreft het aantal formulieren per respondent als van het aantal deelnemende scholen, klassen per school en leerlingen per klas. Tevens is hierbij per formulier gecontroleerd op het correct invullen van het 4 cijferige unieke ID nummer. Dit nummer, waarvan de eerste twee cijfers staan voor een schoolidentificatienummer en de laatste twee cijfers voor de respondentidentificatie, is cruciaal om de verschillende persoons- en schoolgegevens correct aan elkaar te koppelen.

De tweede fase betrof het digitaliseren van de formuliergegevens. Hiertoe is een professioneel optisch scanapparaat ingezet met bijbehorende software ('Eyes and Hands'). Daarbij is die software zodanig geprogrammeerd dat het aankruisen van een antwoordoptie bij de formulieren in de daarvoor gereserveerde plaatsen (hokjes) omgezet wordt in een cijfermatige code (bijvoorbeeld 0 voor 'niet aangekruist' en 1 voor 'aangekruist') en dat ingevulde numerieke waardes (bijvoorbeeld de id code) correct ingelezen worden.

Eventuele fouten bij het scanproces zijn met behulp van een speciale module binnen de scansoftware ('Eyes and Hand Verify') via visuele terugkoppeling handmatig gecorrigeerd in de derde fase van de dataverwerking. Ook zijn eventueel dubbel aangekruiste antwoordmogelijkheden daar waar slechts één alternatief aangekruist had moeten worden, of achteraf gecorrigeerde antwoordmogelijkheden, gecontroleerd en correct verwerkt.

In de voorlaatste fase wordt een definitief databestand in het SPSS formaat weggeschreven. SPSS is een softwarepakket, speciaal ontworpen voor het doen van statistische analyse bij sociaal onderzoek. Doordat binnen SPSS geprogrammeerd kan worden, leent dit programma zich ook voor een softwarematige controle op volledigheid, consistentie en geldigheid van de verzamelde data. Die controle is in de vorige fases in principe reeds gebeurd, maar wordt in de afsluitende controlefase nog eens via SPSS programmatuur overgedaan. Fouten die tijdens het verifiëren (fase 3) niet opgemerkt zijn, zoals niet bestaande coderingen (bijvoorbeeld het getal 7 als antwoord op een gesloten vraag met 4 antwoordmogelijkheden) of tijdens het scanproces verkeerd ingelezen numerieke gegevens, kunnen zodoende alsnog opgespoord en gecorrigeerd worden. Voorts bestaat de mogelijkheid om eventueel ontbrekende gegevens te imputeren. Daar is in dit onderzoek, behalve bij ontbrekende id nummers, geen gebruik van gemaakt.

De analysefase is de laatste fase van de dataverwerking. Hier worden de gegevens met behulp van het statistische softwarepakket SPSS geanalyseerd en voor publicatie voorbereid (met behulp van MS Excel).

De analysesresultaten zijn in het volgende hoofdstuk 3 opgenomen.

Bij de rapportage is het CBS gehouden aan de statistiekverordening. Die verordening stelt dat statistische gegevens niet in zodanige vorm openbaar gemaakt mogen worden dat daaruit onmiddellijk gegevens ontleend kunnen worden omtrent afzonderlijke personen (in onderhavig geval de leerlingen). Dat houdt automatisch in dat deze toets op dit moment niet gebruikt kan worden voor eventuele doorstroombeslissingen van individuele leerlingen. In het kader van een uitgebreid leerlingvolgsysteem en de kwaliteitsbewaking binnen het onderwijs kan dit in de toekomst eventueel wel gebeuren.

De scholen zijn wederom geanonimiseerd. Het oorspronkelijk toegekende schoolnummer is in deze publicatie vervangen door een ander willekeurig school id-nummer. Dat is gedaan in de wetenschap dat het vergelijken van schoolresultaten onderling (nog) niet helemaal eerlijk gebeuren kan en om eventuele stigmatisering van scholen te voorkomen. Wel zijn dezelfde school id nummers als in de publicatie van 2007 aangehouden, waardoor resultaten van individuele scholen in de tijd vergeleken kunnen worden.

3. Analyseresultaten

Respondenten en non respons

Tabel 2

afkorting schoolnaam	Respondentenlijst			gemiddelde klasgrootte	aantal deelnemers	
	klas1	klas2	kl. 3/4			
Aloysius	21	19		40	20	40
Bon Bini	21	19		40	20	39
Cac. Macuarima	21	21		42	21	41
Conrado Coronel	30	31		61	30.5	59
Cristo Rey	32			32	32	29
Dominicus	24	24		48	24	47
Faith Revival	12			12	12	11
Fatima	20	19		39	29.5	36
Felipe B Tromp	23	24		47	23.5	45
Filomena	23	26		49	24.5	46
Fransiscus	25			25	25	24
Frere Bonifacius	30			30	30	30
Graf v. Zinzendorf	23			23	23	23
Hilario Angela	21			21	21	20
Kukwisa	24	24		48	24	47
Laura Wernet Paskel	20	18		38	19	38
Maria	23	22		45	22.5	44
Maria Goretti	32	30		62	31	61
Maria Regina	14			14	14	14
Mon Plaisir	28	29		57	28.5	56
Ora Ubao	21	21		42	21	42
Pastor Kraanwinkel	22	21		43	21.5	42
Pieter Boer	25	24		49	24.5	45
Pius X	21	22		43	21.5	38
Reina Beatrix	24	27	21	72	24	68
Rosa	32			32	32	32
Rosario	26	26		52	26	45
Sagrado Curason	23	24	23	70	23.3	68
San Hose	23	25		48	24	48
Santa Teresita	31			31	31	31
St Michael	20			20	20	19
St. Anna	28	25		53	26.5	52
St. Paulus	18	20		38	19	39
Washington	26	25	23/24	98	22.7	97
totaal				1464		1416

Tabel 2 geeft een overzicht van de scholen die aan de eindtoets basisonderwijs hebben deelgenomen. Tevens is per school vermeld het aantal 6^{de} klassen, het aantal leerlingen per klas (de klasgrootte), het totaal aantal 6^{de} klassers (kolom "tot") en de gemiddelde klasgrootte.

In de laatste kolom kan het aantal deelnemende leerlingen aan de eindtoets afgelezen worden.

De verhouding van het aantal niet deelnemende leerlingen tot de doelpopulatie (het totale aantal aanwezige zesde klassers in de aan de eindtoets deelnemende 34 scholen) geeft de grootte van de algemene non respons aan. Deze bedraagt, uitgedrukt in procenten, 3,3 %. Dat is weliswaar bijna twee keer zoveel als tijdens de vorige toets, maar is nog altijd laag te noemen. De instructie naar het veld en de verdere organisatorische afhandeling is goed verlopen waardoor praktisch alle antwoordformulieren in goede staat (niet gekreukeld, vervuild of bewerkt met tipp-ex) en correct ingevuld ter verwerking aan het CBS werden aangeboden. Behalve de 3.3% leerlingen die beide toetsdagen afwezig waren, waren ook enkelen alleen één toetsdag afwezig: voor de eerste toetsdag bedroeg het afwezigheidspercentage 0.3%, voor de tweede 1,0%.

Tabel 3

Percentage afwezig per toets		
	rekenen	taal
toets 1	0.28	0.28
toets 2	0.99	0.35
toets 3	1.06	1.06

Uit tabel 2 valt af te lezen dat in totaal 1416 leerlingen aan de eindtoets hebben deelgenomen. Die hebben niet allemaal aan *alle* toetsen deelgenomen. De specifieke non respons per toets (het aandeel leerlingen dat niet heeft deelgenomen aan één bepaalde toets) is bijzonder laag en stijgt tijdens de toetsperiode. Zie voor nadere gegevens de tabel 3.

Naast de algemene en toetsspecifieke non respons is de item non respons van belang. Dat is de non respons per vraag. Deze is per school en per toets uitgewerkt en kan apart opgevraagd worden. Deze waarde kan dan vergeleken worden met het landelijk gemiddelde. Uit de analyse van deze item non respons volgt dat het percentage leerlingen dat één of meerdere keren binnen een bepaalde toets het antwoord schuldig blijft, voor alle toetsen niet bijzonder hoog is. Opvallend is dat in vergelijking met vorig jaar alle rekentoetsen op alle onderzochte item non respons percentages duidelijk lager scoren terwijl dat niet het geval is met de desbetreffende percentages voor de taaltoetsen. Ervan uitgaande dat de omvang van de item non respons indicatief is voor de moeilijkheidsgraad van een toets, betekent dit dat de rekentoetsen in vergelijking met vorig jaar als gemakkelijker ervaren zijn en dat vooral de derde taaltoets relatief moeilijk was. Zie tabel 4.

Tabel 4

Percentage leerlingen zonder antwoord per toets						
	rekenen 1	rekenen 2	rekenen 3	taal 1	taal 2	taal 3
één keer	2.61	1.91	3.03	3.11	3.88	8.26
twee keer	0.42	0.35	0.64	0.21	0.85	0.56
drie keer of meer	0.21	0.07	0.21	0.78	0.64	0.35
totaal	3.25	2.33	3.88	4.09	5.36	9.17
totalen voor 2007	4.23	2.62	4.76	5.10	5.57	6.10

Het percentage leerlingen dat antwoorden verkeerd ingevuld heeft (meer dan één antwoordmogelijkheid aangekruist) is net als bij de vorige toets bijzonder laag te noemen. Zie tabel 5.

Tabel 5

Percentage leerlingen met dubbele antwoorden per toets						
	rekenen 1	rekenen 2	rekenen 3	taal 1	taal 2	taal 3
één keer	1.20	0.71	0.99	1.34	1.83	2.68
twee keer	0.07		0.14		0.07	0.14
totaal	1.27	0.71	1.13	1.34	1.91	2.82

CONCLUSIE 1.

Het percentage van alle 6^{de} klas leerlingen dat niet deelgenomen heeft aan de eindtoets basisonderwijs bedraagt 3.3%. Dat is een afwezigheidspercentage dat weliswaar bijna dubbel zo hoog ligt als dat van de eindtoets in 2007 maar is nog altijd laag te noemen. Ook de non respons voor de afzonderlijke reken- en taaltoetsen en in mindere mate voor de item (vraag) non respons zijn in het algemeen laag. Zeer laag is verder het percentage leerlingen dat meer dan één antwoordmogelijkheid bij de multiple choice vragen aankruist.

Toetsresultaten per school

Tabel 6

school- id	Toetsgemiddelden per school		
	rekenen	taal	totaal
24	4.2	4.3	4.3
30	3.6	5.1	4.3
22	4.5	4.6	4.5
28	4.8	4.9	4.8
20	5.0	4.7	4.8
26	4.9	4.8	4.9
16	4.9	4.8	4.9
12	4.8	5.1	4.9
14	5.4	4.7	5.1
7	5.1	5.1	5.1
8	5.2	5.2	5.2
6	5.1	5.3	5.2
3	5.5	5.0	5.2
2	5.2	5.4	5.3
17	5.2	5.4	5.3
9	5.4	5.2	5.3
4	5.3	5.4	5.3
5	5.7	5.0	5.4
33	5.6	5.2	5.4
34	5.5	5.4	5.5
18	5.9	5.2	5.5
1	5.7	5.3	5.5
13	5.6	5.5	5.6
27	6.0	5.5	5.7
31	6.2	5.5	5.8
32	5.9	5.9	5.9
15	6.2	5.7	5.9
19	6.0	5.9	6.0
29	6.0	6.5	6.3
21	6.6	6.1	6.3
11	6.2	6.7	6.4
25	6.3	6.6	6.5
23	7.1	6.7	6.9
10	7.6	6.9	7.3

De toetsresultaten per school voor de eindtoets basisonderwijs 2008 zijn in tabel 6 weergegeven. Ze zijn gerangschikt naar oplopende totaalscore voor de schoolgemiddelden van rekenen en taal samen. De toetsgemiddelden zijn genormeerd op een schaal van 0-10 en er is géén rekening gehouden met het kansaspect van de gebruikte multiple choice vragen. De afzonderlijke scores voor rekenen en taal zijn tevens opgenomen. De schaal van 0-10 is gekozen i.v.m. de makkelijke interpreteerbaarheid en onderlinge vergelijkingsmogelijkheden. De schaal kan niet gebruikt worden om een voldoende-/onvoldoende oordeel te vellen.

Dat houdt voornamelijk verband met:

1. het ontbreken van een normering vooraf
2. het ontbreken van een validiteits- en betrouwbaarheids-toetsing van de meetinstrumenten
3. het (nog) niet kunnen betrekken van de *totale* schoolcontext bij het onderling vergelijken

Tabel 6 geeft een eerste indicatie over de onderlinge prestatieverschillen m.b.t. taal- en rekentoetsen van bijna alle Arubaanse basisscholen. Dat vergelijken dient met de grootste voorzichtigheid te gebeuren omdat onderwijsprestaties nu eenmaal niet los gezien kunnen worden van de veelzijdige en complexe context waarbinnen ze geleverd worden. Met andere woorden, een school die in dit score-overzicht laag scoort is niet per definitie een kwalitatief slechte school.

De meest in het oog springende resultaten zijn net als in 2007 de relatief grote onderlinge scoreverschillen tussen de scholen. Voor de rekentoetsen is de spreiding in schoolgemiddelden liefst 4 volle punten. Bij de taaltoetsen bedraagt het verschil tussen de hoogst en laagst scorende school 2.6 punten. Dat is in beide gevallen erg hoog, zeker als men in acht neemt dat het hier om schoolgemiddelden gaat.

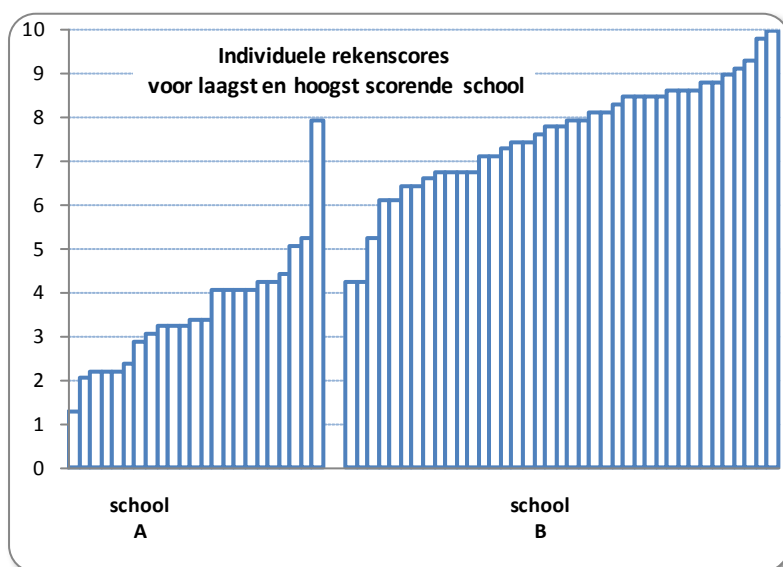
Het nationale gemiddelde voor de rekentoetsen samen laat zich berekenen op 5.5 en dat voor de taaltoetsen op 5.4. In vergelijking met de eindtoets in 2007 zijn de scores voor rekenen en taal dus weer nagenoeg gelijk. Absoluut gezien echter, liggen de scores in 2008 belangrijk hoger dan het gemiddelde van 5.0 uit 2007 voor zowel de reken- als de taaltoetsen.

CONCLUSIE 2.

De landelijk gemiddelde eindscore is voor de reken- en taaltoetsen nagenoeg gelijk en bedraagt respectievelijk 5.5 en 5.4.

Een normerende uitspraak over dit nationale gemiddelde is gezien bovenstaande, in absolute zin niet te geven. Wel kunnen de resultaten vergeleken worden met die van het vorige schooljaar. Dat zal later, in de paragraaf “Toetsresultaten in vergelijking met 2007” afzonderlijk gebeuren. Afgezien van het nationale niveau op rekenen en taalgebied zullen de grootste zorgen vooral betrekking hebben op de relatieve grote onderlinge scoreverschillen tussen de scholen.

Zo scoren praktisch alle leerlingen (op drie na) van de school met de hoogste gemiddelde rekenscore, hoger dan de op één na hoogst scorende leerling van de school met de laagste gemiddelde eindscore voor rekenen. Zie figuur 2. Bij school B scoren drie leerlingen een onvoldoende en vijf een negen of meer. Bij school A scoort slechts één leerling een voldoende en 5 om en nabij de twee. Dat is bijzonder laag als men bedenkt dat de gemiddelde score bij het volledig gokken van de antwoorden op 2,5 ligt¹. Het verschil in rekenscore tussen de twee scholen bedraagt gemiddeld, zoals eerder aangegeven, $7.6 - 3.6 = 4.0$ punten. Dat komt overeen met 40% meer foutief beantwoorde vragen bij school A in vergelijking met school B.



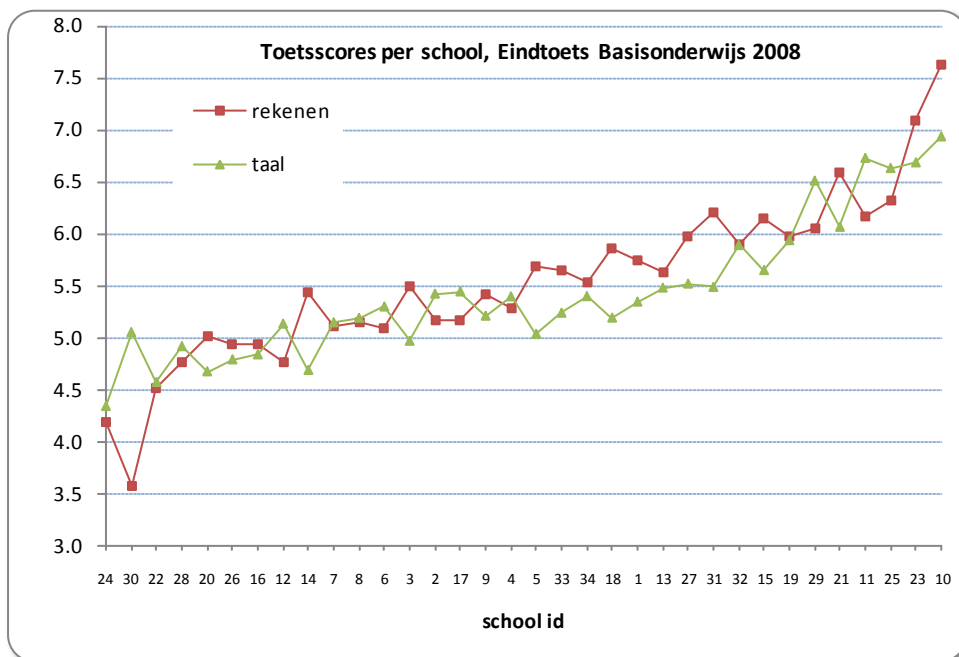
Figuur 2

CONCLUSIE 3.

Naast het nationale eindniveau blijft een groot zorgpunt de grote onderlinge verschillen in toetsscores tussen de scholen. Zo verschilt het laagste schoolgemiddelde voor rekenen 4 volle punten met de hoogste gemiddelde schoolscore voor dit toetsonderdeel.

In figuur 3 zijn de toetsscores uit tabel 6 grafisch weergegeven. Behalve de hierboven beschreven verschillen tussen de scholen kan uit deze figuur ook afgelezen worden de verschillen in toetsscores voor rekenen en taal per school afzonderlijk. Wat opvalt is dat deze verschillen in toetsscores binnen de scholen minder sterk variëren dan de overeenkomstige verschillen tussen de scholen onderling. Het grootste en tevens meest opvallende verschil tussen de reken- en taalscore ligt bij school 30 en bedraagt 1.5 punten. Alle andere verschillen bedragen duidelijk minder dan 1 punt.

¹ gebruikt zijn multiple choice vragen met 4 antwoordmogelijkheden



Figuur 3

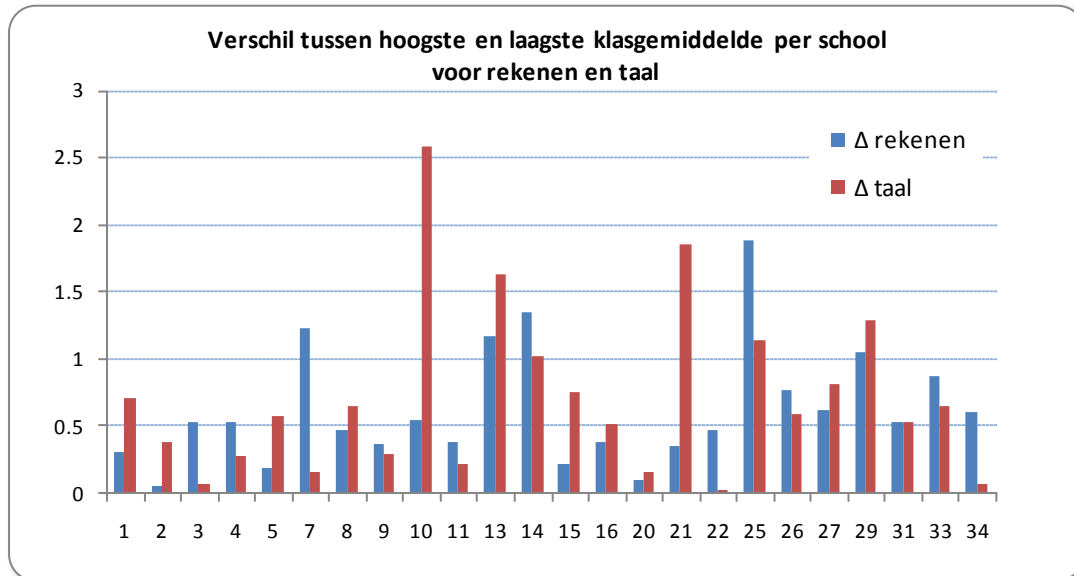
CONCLUSIE 4.

De schoolgemiddelden voor rekenen en taal verschillen voor de meeste scholen relatief weinig van elkaar. Met ander woorden, scholen die relatief goed presteren voor taal, presteren in het algemeen ook goed voor rekenen en scholen met relatief lage rekenresultaten scoren veelal ook laag voor taal. De variatie tussen de toetsgemiddelden voor taal en rekenen binnen de scholen, is kleiner dan de variatie in schoolgemiddelden voor deze toetsen tussen de scholen onderling.

Toetsresultaten per schoolklas

Behalve op schoolniveau kunnen de toetsresultaten ook bekeken worden op klasniveau. Een interessante onderzoeksvraag is in hoeverre de onderzochte scholen intern, tussen de verschillende 6^{de} klassen, zich onderscheiden in gemiddelde toetscore.

Voor alle scholen met meer dan één zesde klas (24 van de 34) is daartoe het maximale verschil in toetsscore tussen de klassen voor zowel rekenen als taal in kaart gebracht. Door alleen te kijken naar het maximale verschil blijven de scholen anoniem omdat het aantal 6^{de} klassen dat voor sommige scholen een unieke identificatie is, onbekend blijft. Bovendien kan zodoende de effectgrootte voor alle betrokken scholen in één figuur overzichtelijk weergegeven worden. Zie figuur 4.



Figuur 4

Af te lezen valt dat maximale prestatieverschillen tussen verschillende 6^{de} klassen binnen dezelfde school van 0.5 punten of meer eerder de regel dan uitzondering zijn. Zo zijn er slechts vijf scholen te vinden waarbij de gemiddelde scores voor zowel rekenen als taal tussen de klassen niet meer verschillen dan 0.5 punten. Voor veel scholen zijn de maximale klasverschillen bovendien opvallend groot. Bijna 30% van de onderzochte scholen bezitten op één of beide toetsen een maximaal verschil van meer dan één punt tussen hun zesde klassen. Binnen een school is de kans dus vrij groot dat één klas zich relatief dicht bevindt bij de laagst scorende scholen uit de schoolprestatielijst, terwijl een andere klas op nationaal niveau dicht aansluit bij de best presterende scholen. (Vergelijk bijvoorbeeld in figuur 3 op de vorige pagina, de ligging van scholen met een score van 5.0 voor rekenen met die scholen die een 6.0 scoren voor dezelfde toets.)

Dit zijn opvallend grote verschillen, omdat aangenomen mag worden dat binnen één en dezelfde school op één en hetzelfde niveau geen grote verschillen te vinden zullen zijn in groot deel van de schoolexterne en schoolinterne factoren. Deze leerlingen van verschillende 6^{de} klassen gebruiken waarschijnlijk dezelfde leermethodes, hebben grotendeels op dezelfde school 5 eerdere klassen doorlopen, zijn doorgaans afkomstig uit dezelfde schoolbuurt, etc. etc. Wel doet zich het verschijnsel voor dat leerlingen uit eenzelfde sociale milieu of met dezelfde taalachtergrond elkaar opzoeken en hierdoor een homogene verdeling van desbetreffende schoolexterne contextvariabelen over de verschillende klassen kunnen voorkomen. Er is echter geen enkele significantie correlatie gevonden tussen de maximale prestatieverschillen tussen de klassen en eventuele bijbehorende verschillen ten aanzien van deze schoolexterne factoren. Het meest belangrijke verschil op klasniveau binnen een school is een voor de hand liggende schoolinterne en menselijke factor: de leerkrachten van de school zelf. De verschilmetingen uit figuur 4 weerspiegelen daarom waarschijnlijk prestatieverschillen op leerkrachtniveau.

CONCLUSIE 5.

Behalve tussenschoolse zijn ook aanzienlijke binnenschoolse verschillen op toetsresultaten te vinden. Deze zijn aanzienlijk zowel in de mate van voorkomen als voor wat betreft het maximale scoreverschil tussen klassen van dezelfde school. Deze verschillen weerspiegelen waarschijnlijk prestatieverschillen op leerkrachtniveau.

Toetsresultaten in vergelijking met 2007

Tabel 7

Landelijke toetsscores per schooljaar

		2007	2008
rekenen	toets 1	5.2	5.5
	toets 2	5.1	6.3
	toets 3	4.8	4.9
	gemiddeld	5.0	5.5
taal	toets 1	4.7	5.1
	toets 2	5.0	5.7
	toets 3	5.2	5.6
	gemiddeld	5.0	5.4

Tabel 7 geeft een overzicht van de afzonderlijke toetsscores voor rekenen en taal in vergelijking met de scores behaald in 2007. Opmerkelijk is dat in 2008 het landelijk gemiddelde voor *alle* toetsen hoger ligt dan in 2007. Gebleven zijn de sterke verschillen in gemiddelde toetsscores tussen scholen onderling en de minder sterke verschillen tussen de reken- en taalscores van individuele scholen. Bovendien zijn de reken- en taalscores op nationaal niveau weer nagenoeg gelijk. (Zie conclusies 3 en 4).

Wat anders is, is de hogere nationale eindscore voor rekenen en taal en de gewijzigde prestatievolgorde van de scholen. Deze volgorde zal hieronder nader onder de loep genomen worden, terwijl een beschouwing over de hogere toetsscores in de volgende paragraaf "Leereffect of toetseffect?" aan bod komt.

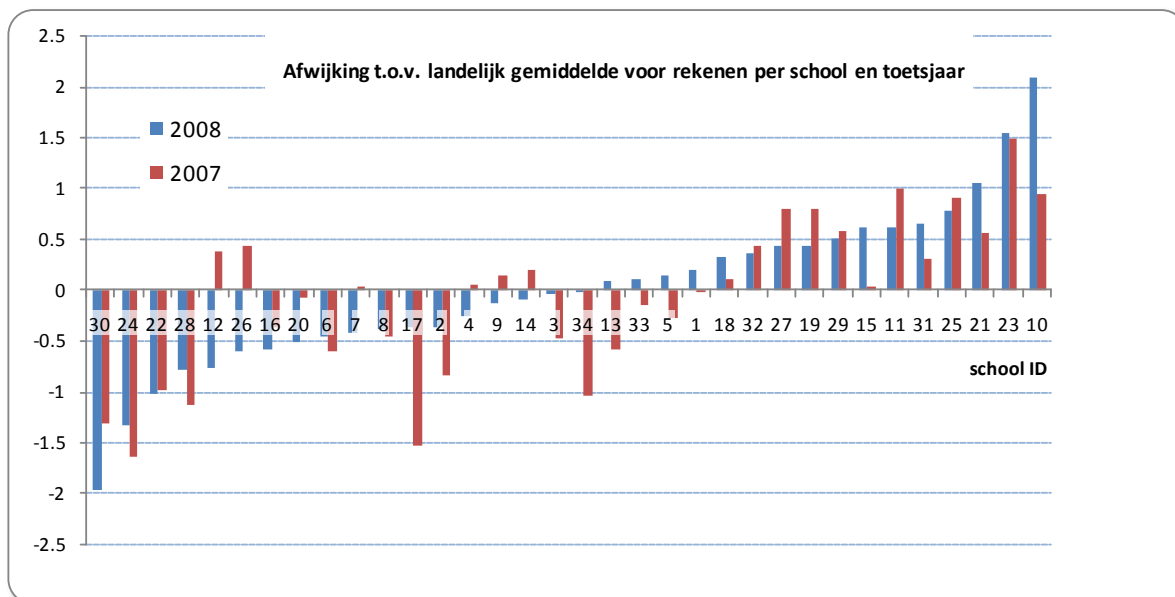
Uit tabel 7 volgt verder dat het opvallende diachronische scoreverloop uit 2007 (oplopende taalscores en aflopende reken scores in de tijd) niet meer opgaat voor de desbetreffende toetsscores in 2008.

CONCLUSIE 6.

De gemiddelde scores voor de afzonderlijke reken- en taaltoetsen liggen voor 2008 duidelijk hoger dan voor 2007. Daardoor is het landelijke gemiddelde voor rekenen en taal met respectievelijk 0.5 en 0.4 punten (op een schaal van 10) toegenomen t.o.v. 2007. Opvallend is het grote scoreverschil voor de tweede rekentoets indien men de resultaten van 2008 vergelijkt met die van 2007. Leerlingen scoren voor deze toets in 2008 gemiddeld 1.2 punten hoger dan in 2007.

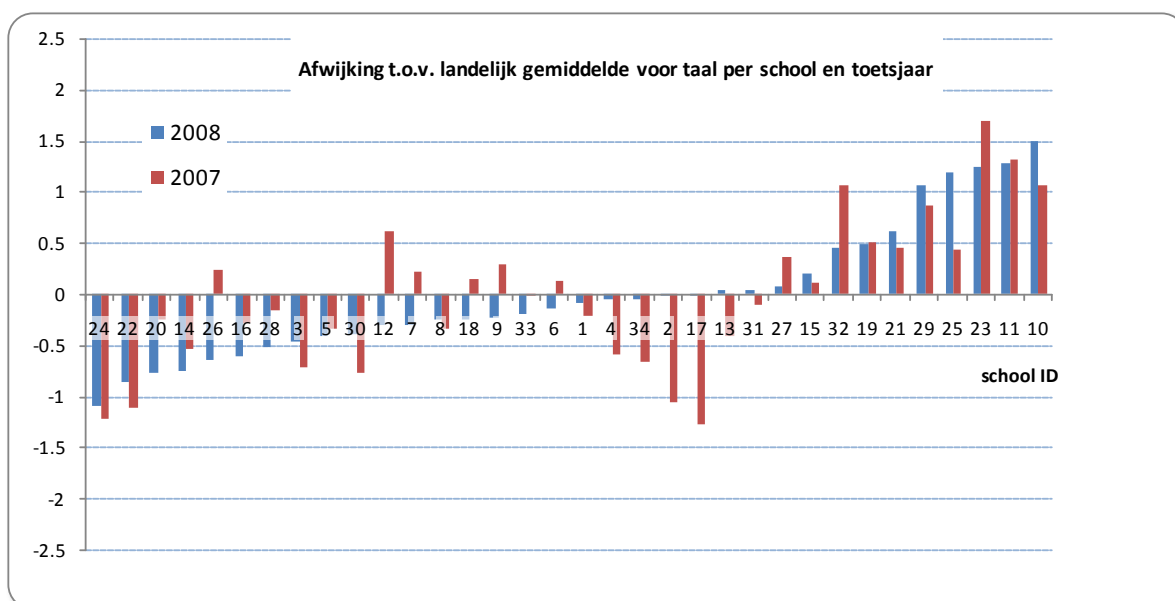
In figuur 5 zijn net als in figuur 4 de individuele schoolgemiddelden voor rekenen weergegeven. Deze keer echter niet als absolute score maar als afwijking ten opzichte van het landelijke gemiddelde. Voor 2008 bedroeg dat landelijke gemiddelde voor rekenen 5.5. Links in de figuur staan de scholen die lager scoren dan dit gemiddelde en rechts de scholen die boven dit gemiddelde scoren. De lengte van de blauwe staven geeft de afwijking ten opzichte van dit gemiddelde aan. School 30 scoort in 2008 dus gemiddeld bijna 2 punten onder het gemiddelde 5.5 (zijnde een 3.6 als absolute score). De scholen zijn weergegeven in de volgorde van oplopende reken score voor 2008. De laagst scorende school uit 2008 staat dus helemaal links en de hoogst scorende school uit 2008 helemaal rechts.

Deze manier van weergeven vergemakkelijkt het vergelijken van de scores uit 2008 met die van 2007 (die als rode staven zijn weergegeven). Wat meteen opvalt, is dat het merendeel van de scholen die het in 2008 relatief goed doen (een naar boven gerichte blauwe staaf) het ook goed deden in 2007 (een naar boven gerichte rode staaf) en dat veel scholen die in 2008 onder het landelijk gemiddelde scoorden dat ook al deden in 2007.



Figuur 5

Eenzelfde conclusie kan getrokken worden met betrekking tot de taalscores uit 2007 en 2008. Zie figuur 6.



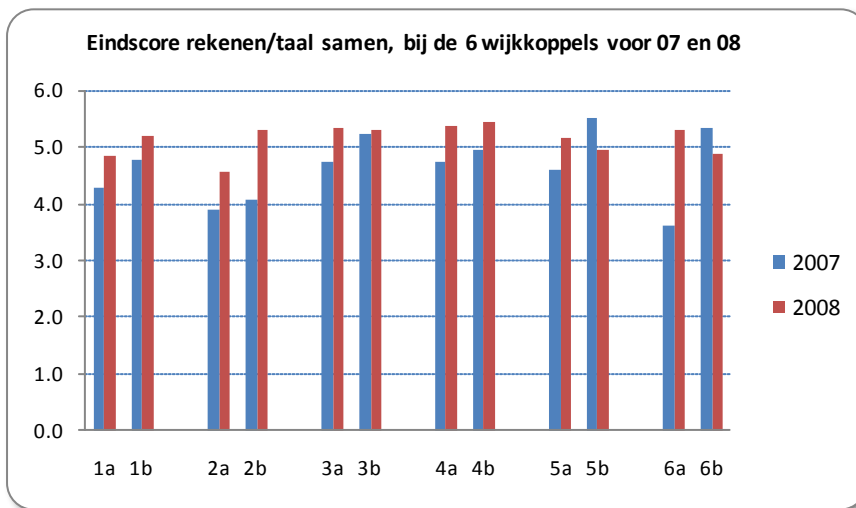
Figuur 6

CONCLUSIE 7.

Hoewel de rangvolgorde voor zowel de reken- als taalscore in 2008 t.o.v. 2007 veranderd is kan gesteld worden dat het merendeel van de scholen die in 2007 duidelijk boven/onder het landelijk gemiddelde voor rekenen of taal scoorden dat ook weer doen in 2008.

Uit beide bovenstaande figuren 5 en 6 kan bovendien relatief eenvoudig afgelezen worden welke scholen ten opzichte van vorig jaar de grootste winst of het grootste verlies in gemiddelde score vertonen. Zo valt meteen op dat bijvoorbeeld school 17 die in 2007 zowel voor rekenen als taal tot de slechtst presterende scholen behoorde in 2008 voor beide onderdelen heel dicht bij het landelijk gemiddelde ligt. Daarentegen vertoont school 26 op beide toetsonderdelen een daling van bovengemiddeld (in 2007) naar 'ondergemiddeld' (in 2008). Dat is opmerkelijk omdat beide scholen een wijkkoppel vormen.

Onder wijkkoppel wordt verstaan een tweetal scholen die in de onmiddellijke nabijheid van elkaar gelegen zijn en doorgaans leerlingen vanuit dezelfde schoolomgeving(wijk) betrekken. Wijkkoppels zijn interessante onderzoeksobjecten omdat wellicht een aanzienlijk deel van de onderlinge verschillen in schoolexterne invloedsfactoren geëlimineerd wordt. Aruba bezit 12 van zulke scholen, verdeeld in 6 koppels in de wijken Brasil, St. Cruz, Madiki/Rancho, San Nicolas, Noord en Bloemond/Paradera (afgezien van het kwartet in Oranjestad).



Figuur 7

Tien van de twaalf wijkkoppelscholen presteren in 2008 beter dan in 2007. Twee scholen doen het in 2008 minder goed dan in 2007. In beide gevallen wordt de in 2007 nog hoogst scorende school in het koppel in 2008 voorbijgestreefd door zijn desbetreffende koppelschool. Dat wisselen in prestatievolgorde doet zich verder bij geen enkel ander koppel voor.

Toetseffect of leereffect?

De onderlinge prestatieverschillen tussen de scholen blijven groot en de relatief goed of slecht presterende scholen handhaven in het algemeen hun plaats op de rangvolgordelijst bij het vergelijken van de toetsresultaten uit 2007 met die van 2008. In 2008 echter ligt het landelijk gemiddelde voor rekenen en taal respectievelijk 0.5 en 0.4 punten hoger en dat is een aanzienlijke stijging voor een gemiddelde op landelijk niveau.

Anders weergegeven, als gekeken wordt naar de eindscore voor alle reken- en taaltoetsen samen dan:

- haalden in 2007 “slechts” 2 van de 34 scholen een eindscore hoger dan 6.0. Meer dan de helft (18 van de 34) scoort onder de 5.0!
- zijn er in 2008 “reeds” 6 scholen die boven de 6.0 scoren en bedraagt het aantal scholen dat onder de 5.0 scoort: 8 van de 34.

Dat zou op zich een reden tot tevredenheid moeten zijn, ware het zo dat deze landelijke toetsresultaten voornamelijk bepaald worden door de effectiviteit van leerprocessen. Betere toetsresultaten weerspiegelen dan toegenomen leereffecten. Dat is in algemene zin helaas niet zo. Toetsresultaten meten niet alleen leereffecten maar ook toetseffecten en ook andere (omgevings)effecten. Met andere woorden betere toetsresultaten kunnen ook het gevolg zijn van gemakkelijkere toetsen. Al snel dient zich dan ook de vraag aan hebben we in het onderhavige geval te maken met een leereffect of een toetseffect?

Het beantwoorden van deze vraag is bij het (nog) niet voorhanden zijn van een genormeerd toetsinstrument moeilijk, men heeft immers ook te maken met twee verschillende leerlingpopulaties.

Doordat de resultaten van de nationale eindtoets basisonderwijs relatief snel bekend waren (nog voor het einde van het schooljaar 07/08) bood zich de mogelijkheid aan om een deel van de leerlingen die deelgenomen hadden aan de eindtoets 2008 ook de toets uit 2007 voor te leggen. Als deze leerlingen ook de toets uit 2007 significant beter zouden maken dan de leerlingpopulatie uit 2007, kan men waarschijnlijk spreken van een leereffect. Indien voor de toets uit 2007 significant lager gescoord wordt dan voor de toets uit 2008 zal dat waarschijnlijk komen omdat beide toetsen qua moeilijkheidsgraad verschillen. Enige voorzichtigheid is ook hier op zijn plaats, vandaar de toevoeging van het bijwoord “waarschijnlijk”. De steekproef onder de leerlingpopulatie dient namelijk representatief te zijn voor de hele populatie en typische problemen zoals die zich voordoen bij herhaald meten dienen zoveel mogelijk vermeden (of gecompenseerd) te worden (tussentijdse leereffecten en toets“moeheid”). Bovendien dient het toetsinstrument betrouwbaar te zijn (consistente resultaten leveren bij toetsherhaling).

Uiteindelijk is gekozen voor het selecteren van een zestal scholen, die duidelijk verschillend presteren voor zowel de reken- als de taaltoetsen uit 2007 en die bij de resultaten uit 2008 in verschillende mate vorderingen maken. Ook is een selectie uit de zes toetsen genomen: toets 2 voor zowel rekenen als taal.

Helaas bleek in de praktijk dat het zeer moeilijk was het gewenste aantal scholen bereid te vinden de toets af te nemen. Dit ondanks het gegeven dat er voor elke school een alternatieve school aangegeven was en dat het schooljaar nog officieel twee weken te gaan had. Meerdere scholen gaven aan de leerlingen onvoldoende te kunnen motiveren voor de twee toetsen, en dat bovendien heel wat schoolverlaters inmiddels reeds met vakantie waren! Uiteindelijk hebben 4 scholen eind juni deelgenomen aan de toetsen uit 2007.

Een tweede tegenslag was dat nadere analyse van de toetsresultaten resulteerde in sterke verdenkingen van fraude. Dat had tot gevolg dat de resultaten van één deelnemende school en verder één klas van een andere school in tweede instantie buiten de analyse gehouden moesten worden. Uiteindelijk zijn de toetsantwoorden

van een kleine honderdtal leerlingen afkomstig uit drie scholen onderzocht. Hun testresultaten zijn vervolgens in tabel 8 onder de kolom “ijking” met die uit de reguliere eindtoets basisonderwijs 2007 en 2008 vergeleken.

Tabel 8

Eindscore Rekentoets-2			
school	ijking	2008	2007
A	4.7	5.1	5.1
B	6.7	7.6	5.5
C	5.7	6.6	5.1
gemiddeld	5.9	6.8	5.4

Eindscore Taaltoets-2			
school-id	ijking	2008	2007
A	5.2	5.7	5.7
B	6.3	6.1	5.7
C	5.2	5.4	5.2
gemiddeld	5.7	5.9	5.5

Uit tabel 8 volgt dat alle scholen beide toetsen uit 2007, op één uitzondering na, slechter maken dan de dienovereenkomstige zelfde toets uit 2008. Dat geldt vooral voor de rekentoetsen (vergelijk kolom “ijking” met kolom “2008”). Dezelfde groep leerlingen scoort voor de rekentoets uit 2007 bijna een vol punt lager dan voor de reguliere rekentoets uit de eindtoets basisonderwijs 2008. Bij de taaltoets is het verschil veel kleiner: slechts 0.2 punten; één school scoort voor de toets uit 2007 bij de ijking zelfs beter dan voor dezelfde toets uit 2008. Dat betekent dat in ieder geval voor rekenen sprake is van een relatief sterk toetseffect. De rekentoets 2 uit 2008 valt gemakkelijker uit dan dezelfde rekentoets uit 2007. Ook in conclusie 6 is reeds gewezen op de opmerkelijke hoge landelijke score voor deze rekentoets. Het is bovendien juist deze toets waarvoor extra bedenktijd in 2008 uitgetrokken is.

Ter vergelijking zijn in tabel 8 ook de schoolresultaten uit 2007 opgenomen. Als de schoolresultaten uit deze laatste kolom vergeleken worden met de prestaties van de leerlingen uit één lichting later (kolom “ijking”) dan is er wellicht ook sprake van een licht leereffect en ook hier zijn de verschillen bij rekenen in sterkere mate aanwezig.

Combineert men deze bevindingen met eerdere conclusies over de item non respons (zie: “Respondenten en non respons”), met het feit dat alle andere reken- en taaltoetsen uit 2008 op nationaal niveau in vergelijking met 2007 hoger scoren en conclusies uit de voorlopige toetsevaluatie van het Examenbureau, dan kan samenvattend geconcludeerd worden:

CONCLUSIE 8.

De betere toetsresultaten uit 2008 ten opzichte van 2007 zijn voor een deel te verklaren uit een toetseffect (gemakkelijkere toetsen), dat is vooral het geval voor de rekentoets-2. Daarnaast is er wellicht ook sprake van een leereffect.

Toetsresultaten en schoolcontext

In het rapport “Eindtoets Primair Onderwijs 2007” zijn twee verschillende methodes beschreven waarmee schooltoetsresultaten gecontroleerd kunnen worden op contextfactoren (zie hoofdstuk 2 onder “Schoolcontext”). Met controleren wordt bedoeld dat de invloed van schoolinterne en schoolexterne factoren op de testresultaten zoveel mogelijk geëlimineerd worden. Deze invloedsfactoren zijn namelijk niet evenredig over alle scholen verdeeld waardoor het vergelijken van scholen onderling met grote voorzichtigheid dient te geschieden. Die twee methodes betroffen de zogenaamde “Multiple Classification Analyses”, kortweg MCA-analyse en de “Structural Equation Modeling”, kortweg SEM-analyse.

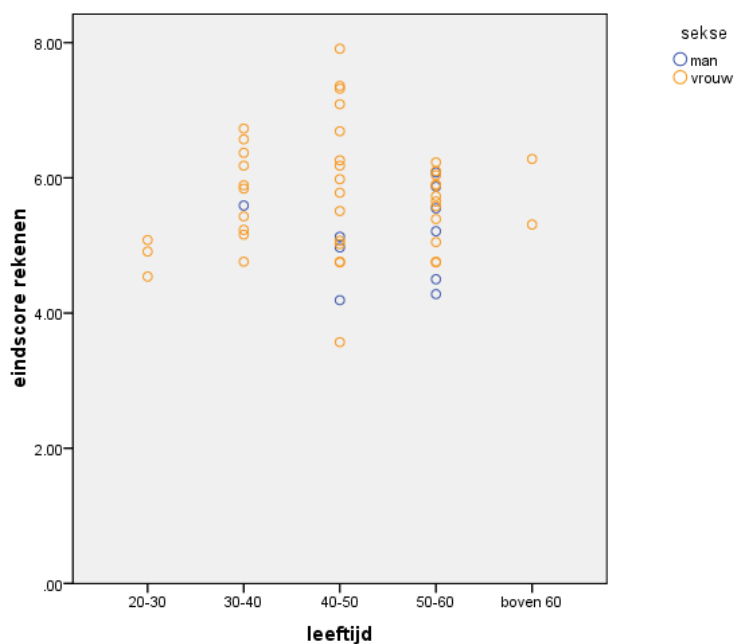
Met behulp van het programma SPSS zijn tijdens de analyse van de toetsresultaten uit 2007 verschillende MCA controles uitgevoerd, telkens met een andere set van schoolcontextfactoren. De sterkste invloedsfactoren zijn vervolgens meegenomen in een SEM analyse met behulp van het programma AMOS (een uitbreiding van SPSS, zie voorbeeld bijlage 4). Ten slotte is met de set sterkste invloedsfactoren een laatste MCA analyse uitgevoerd en een aangepaste lijst met schoolresultaten gepresenteerd. Naar verwachting verminderde hierdoor de maximale variatie van de schoolresultaten en veranderde ook de schoolvolgorde op de prestatielijst. De onderlinge verschillen in schoolprestaties bleven echter aanzienlijk.

Eenzelfde werkwijze is ook dit jaar weer gevolgd. Het voornaamste verschil met de analyse uit 2007 is, dat naast een lichtelijk aangepaste enquête over de schoolexterne factoren, in 2008 ook enkele schoolinterne factoren meegenomen kunnen worden omdat die nu voor het eerst gemeten zijn. Dit is gebeurd met behulp van een zeer uitgebreide door de inspectie ontworpen vragenlijst die aan de schoolhoofden is voorgelegd. Van de 64 6^{de} klassen, verdeeld over 34 basisscholen, zijn de gegevens van 52 (exclusief 2 vervangers) klassen/leerkrachten ontvangen. Na het invoeren en controleren van de data bleek al snel dat de verwachtingen ten aanzien van de bruikbaarheid van de data voor de voorliggende publicatie bijgesteld moesten worden. De schoolinterne variabelen zijn op school/klasniveau gemeten en dat het aantal deelnemende scholen/klassen is relatief klein. Hoe kleiner de populatie echter, hoe sterker de onderzoeksresultaten bepaald worden door individuele scores/meningen. Extra voorzichtigheid (bescheidenheid) bij het trekken van conclusies of presenteren van resultaten is dan op zijn plaats ook al omdat men snel aan de grenzen van individuele anonimiteit zit.

CONCLUSIE 9.

De “small sample variability” binnen de datafile over de schoolinterne variabelen maakt het bijzonder moeilijk om betrouwbare algemene uitspraken of significante (en relevante) verschillen tussen de verschillende subgroepen te vinden.

Als voorbeeld van de relatief grote variabiliteit in uitkomsten over verschillende subgroepen is in figuur 8 weergegeven het scatterplot van de rekenresultaten per klas, uitgesplitst naar geslacht en enkele leeftijdsgroepen van de desbetreffende leerkrachten.



Figuur 8

De belangrijkste resultaten uit het meenemen van schoolexterne factoren worden hieronder besproken. In tegenstelling tot de vorige rapportage uit 2007 wordt deze keer nauwelijks aandacht besteed aan het opstellen van een gecontroleerde schoolrangvolgorde lijst. Wel is onderzocht hoe groot verschuivingen in een bepaalde schoolrangvolgordelijst zijn na controle middels MCA.

Een betrouwbare gedetailleerde absolute schoolrangvolgorde is niet mogelijk en ook niet wenselijk.

Niet wenselijk omdat het uiteindelijke doel van kwaliteitsverbetering in het reken- en taalonderwijs in hoofdzaak gerealiseerd dient te worden via de werkvloer, via de individuele leerkrachten voor de klas. Motivatie, begeleiding, openheid, onderling vertrouwen en samenwerken zijn sleutelwoorden op de weg naar succes. Bij een schoolrangvolgorde ligt de nadruk te sterk op onderlinge prestatieverschillen, op externe motivatie in plaats van interne motivatie en op beoordelen in plaats van begeleiden.

Niet mogelijk omdat:

- nooit alle relevante contextfactoren betrouwbaar genoeg gemeten kunnen worden en bovendien nooit in hun totaliteit in de analyse kunnen worden meegenomen
- van datgene wat wel gemeten is, alleen globale modellen opgesteld kunnen worden; sterk vereenvoudigde weergaves van de werkelijkheid dus,
- de prestatieverschillen tussen de klassen binnen de scholen zo groot kunnen zijn dat de plaats binnen een schoolrangvolgorde zijn betekenis grotendeels verliest (zie conclusie 5)

Wel kan na het uitvoeren van verschillende MCA analyses geconcludeerd worden:

CONCLUSIE 10.

Ook na controle op de schoolexterne factoren blijven aanzienlijke verschillen bestaan in prestaties tussen de verschillende scholen.

Tabel 9

**Verschuivingen in rangvolgorde schoolresultaten
 vóór en na controle op achtergrondfactoren**

school id ²	rangvolgorde		verschuiving
	ongecontroleerd	gecontroleerd	na controle
24	1	2	-1
30	2	1	1
22	3	3	0
26	4	6	-2
16	5	5	0
28	6	7	-1
20	7	4	3
14	8	17	-9
12	9	10	-1
7	10	19	-9
6	11	9	2
8	12	12	0
3	13	20	-7
2	14	8	6
17	15	15	0
9	16	11	5
4	17	13	4
5	18	16	2
33	19	23	-4
34	20	14	6
18	21	18	3
1	22	24	-2
13	23	21	2
31	24	22	2
27	25	27	-2
32	26	33	-7
19	27	28	-1
15	28	25	3
29	29	29	0
21	30	31	-1
11	31	26	5
25	32	30	2
23	33	32	1
10	34	34	0

Ondanks het hiervoor gestelde is het controleren op externe factoren geen overbodige exercitie, al was het alleen maar om aan te tonen hoe sterk de rangvolgorde van tabel 6 na controle op enkele factoren verandert.

Indien gekozen wordt om te controleren op hoogste educatieniveau van de ouders, doubleren, aantal uren/week besteed aan huiswerk en aantal jaren kleuteronderwijs (allemaal significantie invloedsfactoren) ziet men de schoolvolgorde in totaalscore (rekenen en taal samen) uit tabel 6 reeds aardig verschuiven.

Tabel 9 geeft een indruk van die verschuivingen. Een positief getal in de kolom "verschuiving na controle" betekent dat de desbetreffende school de getalswaarde in rangvolgorde stijgt (relatief beter presteert), een negatief getal staat voor een overeenkomstige daling in de rangvolgorde.

Twee scholen verschuiven 9 plaatsen, twee andere 7 plaatsen en weer twee andere 6 plaatsen in rangvolgorde, terwijl 5 scholen dezelfde rangvolgordeplaats bekleden.

Indien gekozen wordt voor een andere set aan variabelen, bijvoorbeeld als "aantal keren doubleren"³ vervangen wordt door de variabele die gekoppeld is aan de vraag "vind je de school leuk?", verandert de gecontroleerde rangvolgorde telkens. Aangezien om technische redenen niet alle schoolexterne variabelen in de MCA analyse meegenomen kunnen worden is dit gegeven een extra reden om uiterst voorzichtig te zijn met rangvolgordelijsten.

De hierboven beschreven relatief forse verschuivingen in schoolvolgorde binnen een schoolprestatielijst na controle op enkele schoolexterne factoren geeft aan dat deze factoren binnen de schoolcontext een belangrijke rol spelen. Dat gegeven wordt ook op een andere manier direct zichtbaar, namelijk indien men de schoolresultaten koppelt aan de geografische (GAC)

² Deze volgorde wijkt lichtelijk af van tabel 6 omdat de MCA analyse alleen die respondenten meeneemt waarop gecontroleerd kan worden voor ALLE factoren.

³ De variabelen "educatieniveau ouders" en "aantal keren doubleren" zijn niet geheel ongecorrleerd, dat kan bij multiple regressie aanleiding geven tot ongewenste multicolineariteit effecten.

ligging van de scholen (zie CBS publicatie “Mapping Census 2000: social-demographic diversity on Aruba”). De groep scholen met de laagste score voor zowel de taal- als rekentoetsen bevindt zich in of zeer nabij wijken die ook laag scoren op indicatoren voor de sociaaleconomische klasse (gemiddeld (huishoud)inkomen, percentage werklozen, criminaliteitscijfers, etc.).

CONCLUSIE 11.

Er bestaat een duidelijke correlatie tussen schoolprestaties en de socio-economische context/omgeving van waaruit de school zijn leerlingen betreft.

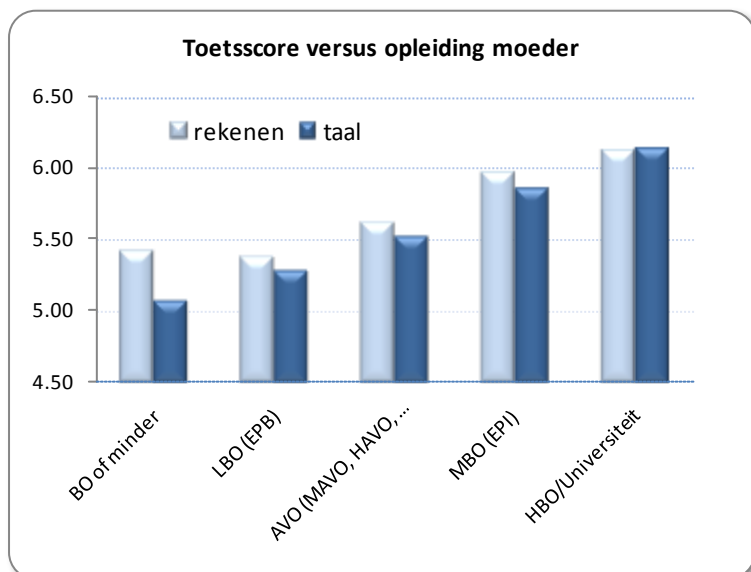
Educationiveau ouders

Wellicht de belangrijkste schoolexterne factor is net als bij de vorige meting, de educatiegraad van de ouders. Het opleidingsniveau van de ouders speelt op velerlei manieren een rol in de leerprestaties van hun kinderen. Ouders met een hoger opleidingsniveau bezitten meer mogelijkheden dan ouders met een lage opleiding om het leerproces van hun kinderen te sturen, te begeleiden of anderszins te faciliteren (o.a. via betaalde bijlessen). Daarnaast zullen voorbeeldfunctie, verwachtingspatroon en betrokkenheid verschillen al naar gelang de sociaaleconomische status (SES). En een belangrijke indicator voor de SES is weer de educatiegraad van de ouders. Een andere factor naast het beschreven ‘nurture’-effect is het doorgeven van aanleg via de genen (het ‘nature’-effect) en ook die factor correleert met het opleidingsniveau van zowel ouders als kinderen. Een schematisch overzicht waaruit de belangrijke ouderrol in de schoolprestaties van kinderen wordt weergegeven is te vinden in de bijlagen 4 en 5. Bijlage 5 is overgenomen uit de CBS publicatie “Het probleem van de overjarige basisschoolleerlingen in nader perspectief; de invloed van de taalachtergrond op de schoolprestaties van basisschoolleerlingen gerelateerd aan de sociaaleconomische en culturele context.” Bijlage 4 bevat een afbeelding uit het programma AMOS, een SEM⁴-analyse instrument waarmee correlatie-coëfficiënten tussen modelvariabelen berekend kunnen worden.

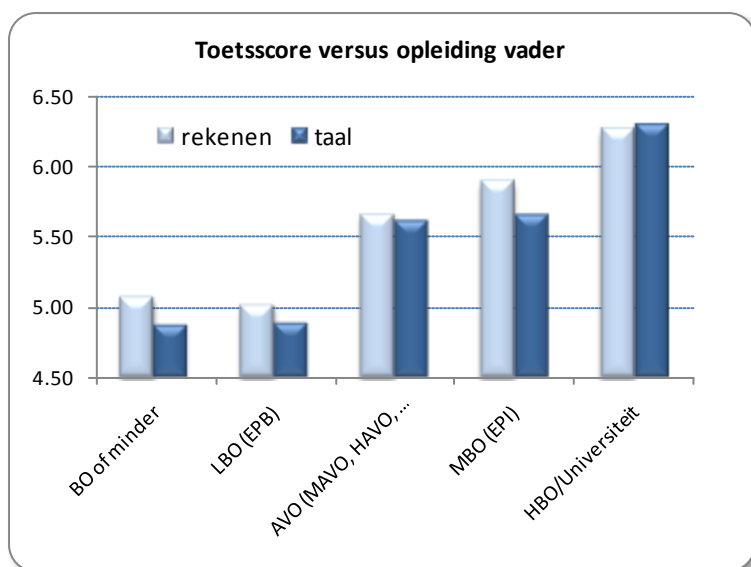
De beschreven ouderlijke invloed is in overeenstemming met bestaande literatuur⁵ en valt in duidelijk uit de figuren 9 en 10 af te lezen.

⁴ “Structural Equation Modeling”

⁵ o.a.: “Parenting and Children’s Educational Achievement, myths and facts about parental involvement”, the liz library (www.thelizlibrary.org/liz/021.htm); Ludger Wößmann, “The Impact of Family Background on Educational Achievement: comparing Europe and the United States”



Figuur 9



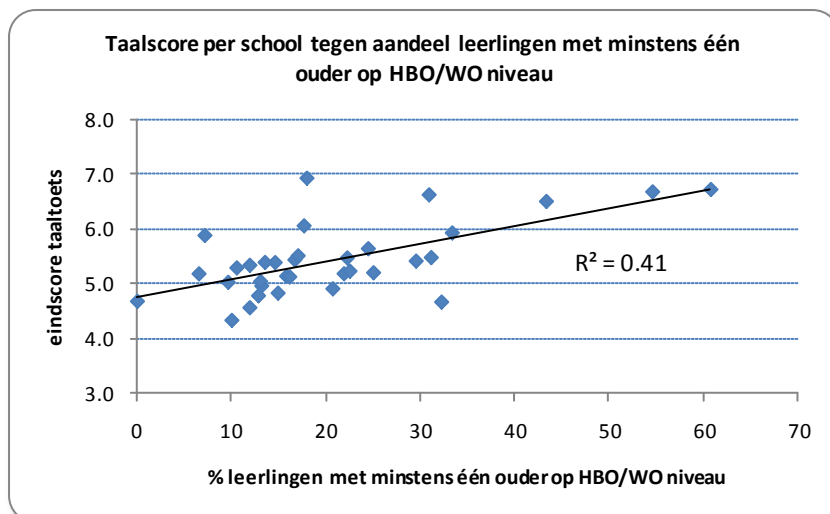
Figuur 10

CONCLUSIE 12.

Een belangrijke schoolexterne factor is het educatieniveau van de ouders, hoe hoger dit educatieniveau hoe hoger de gemiddelde toetsscore van bijbehorende kinderen.

Reeds eerder is gesteld dat vele schoolcontextfactoren niet evenredig over alle scholen verdeeld zijn. Dat geldt ook voor het educatieniveau van de ouders. Figuur 11 laat zien dat scholen nogal verschillen in het percentage leerlingen die minstens één ouder bezitten met een opleiding op HBO of universitair (=WO) niveau (tussen 0 en 62%).

Ook valt af te lezen dat er een redelijk sterke correlatie bestaat tussen deze variabele en de schooleindscore voor de taaltoetsen. Een correlatie die echter bijna uitsluitend gedragen wordt door de drie scholen aan de rechterkant van de grafiek (en de linker extreem). Dat zijn alle drie scholen met een of andere vorm van selectie (religie, taal, geografische ligging, etc.).



Figuur 11

Geboorteland en voertaal thuis

De toetsresultaten uitgesplitst naar geboorteland van de leerling kunnen in tabel 10 afgelezen worden. De weergegeven scores in deze tabel zijn reeds gecontroleerd voor enkele contextfactoren⁶.

Tabel 10

Toetscores naar land van geboorte

land van geboorte	rekenen	taal	eindscore
Aruba	5.6	5.5	5.6
ontwikkelingsland in directe omgeving	6.0	5.4	5.7
Ned. Antillen, Suriname	5.6	5.9	5.7
West Europa, Noord Amerika	5.7	5.8	5.8
elders	6.2	5.8	6.0

⁶ maximaal educatieniveau ouders, aantal jaren kleuteronderwijs en aantal uren per week aan huiswerk

Wat opvalt, is dat het autochtone deel van de getoetste leerlingpopulatie voor de eindscore slechter presteert dan welke allochtone groep dan ook (dat geldt overigens ook indien niet gecontroleerd zou worden voor contextfactoren). Dat komt vooral door de relatief lage score voor de rekentoetsen. Immigranten uit de directe omgeving van Aruba scoren weliswaar lager voor de taaltoetsen, maar compenseren deze lagere score voor taal met een duidelijk hogere score voor rekenen dan de Arubaanse leerlingen.

Indien gekeken wordt naar de taalachtergrond zien de toetsscores er als volgt uit, zie tabel 11. De toetsresultaten zijn wederom gecontroleerd op vermelde contextfactoren.

Tabel 11

Toetsscores naar meest gesproken taal thuis

thuis taal	rekenen	taal	eindscore
Papiamentó	5.5	5.4	5.5
Nederlands	6.2	6.5	6.3
Spaans	5.9	5.5	5.7
Engels	5.3	5.7	5.5
anders	6.9	6.2	6.6

Saillant in tabel 11 is dat leerlingen met als meest gesproken thuistaal het Papiamentó gemiddeld slechter voor Nederlandstalige taaltoetsen scoren dan leerlingen die thuis vooral Spaans praten (en ook deze uitspraak blijft staan indien niet gecontroleerd zou worden). Opvallend des te meer omdat onder de laatste groep ook recente immigranten zitten met een achterstand op taalgebied die wel jaren bedragen kan. Daar staat weer tegenover dat de onderzochte Spaanstalige groep reeds een niveauselectie ondergaan heeft, het taalzwakste deel bevindt zich namelijk in het (overigens sterk krimpende) incorporerend basisonderwijs.

De rekentoetsen zijn natuurlijk niet geheel vrij van taalinvloed omdat het grootste deel van de opgaven in een taalcontext ingebed is. Toch zal de taalinvloed een stuk minder zijn dan bij de specifieke taaltoetsen. Het zal nauwelijks iemand verbazen dat Nederlandstalige leerlingen beter scoren op een taaltoets voor Nederlands dan bijvoorbeeld Spaanstalige leerlingen.

Indien men er van uit gaat dat kinderen met verschillende taalachtergrond (maar met verder gelijke andere contextfactoren) in principe dezelfde rekenaanleg bezitten, zou men verwachten dat de spreiding in gemiddelde scores voor de rekentoetsen geringer zou zijn dan die bij de taaltoetsen (bij uitsplitsing naar taalachtergrond). Men ziet echter een ander beeld: de gemiddelde scores bij rekenen variëren meer dan die van taal tussen de verschillende taalgroepen. Dat komt vooral door de markant hoge score voor rekenopgaves (in een Nederlandstalige context!) bij de groep anderstaligen. Dat lijkt geen toevalstreffer te zijn omdat deze groep een redelijke omvang heeft en ook bij de eindtoets basisonderwijs uit 2007 op rekengebied duidelijk het hoogst scoorde. Waarschijnlijk betekent dit, dat deze leerlingen zich ten opzichte van de andere taalgroepen onderscheiden door andere, niet gemeten cultuureigenschappen met een significante invloed op de rekenprestaties. Een of andere schoolinterne factor kan uitgesloten worden omdat deze relatief grote groep kinderen die in het verre oosten geboren zijn (China, Filippijnen) redelijk homogeen over de verschillende scholen verdeeld is.

CONCLUSIE 13.

Het autochtone deel van de leerlingpopulatie scoort gemiddeld lager voor de eindscore van de nationale eindtoets basisonderwijs dan welk allochtoon deel dan ook. De grootste scoreverschillen tussen leerlingen uit verschillende geboortelanden of met verschillende taalachtergrond, zijn te vinden bij de rekentoetsen. Opvallend is dat de Papiamentstaligen slechter presteren voor de taaltoetsen dan bijvoorbeeld de Spaans-taligen en –net als in 2007- de bijzonder hoge rekenscore voor de groep anderstaligen.

Bovenstaande conclusies lijken coherent te zijn met het doubleergedrag van de verschillende taalgroepen. Daarover meer in de volgende paragraaf.

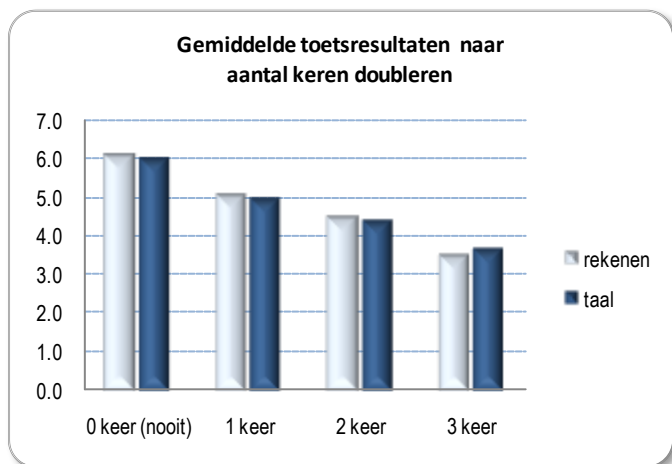
Aantal keren doubleren

Doubleren in het basisonderwijs is een omvangrijk en daarmee bijzonder kostbaar probleem. Het is de hoofdoorzaak waarom ca 40% van de basisschoolpopulatie overjaars is, dus onderwezen wordt op een onderwijsniveau waar ze op grond van hun leeftijd te oud voor zijn. Een conclusie die reeds in 2003 door het CBS met betrekking tot de interne onderwijsefficiëntie genomen is in de publicatie “Onderwijs op Aruba, Context en Output”. Het probleem wordt voornamelijk veroorzaakt door ondermaats presterende Arubaanse leerlingen; zij vormen niet alleen verreweg de grootste deelgroep, maar doubleren bovendien volgens de gegevens uit tabel 12 van alle deelgroepen het meest.

Tabel 12

	Percentage zittenblijvers			
	0 keer (nooit)	1 keer	2 keer	3 keer
Aruba	56	30	13	1
ontwikkelingsland in directe omgeving	59	31	10	0
Ned. Antillen, Suriname	61	19	17	3
West Europa, Noord Amerika	67	29	4	0
elders	63	32	5	0

Door leerlingen een leerjaar te laten overdoen hoopt men dat hun desbetreffende leerprestaties na zittenblijven alsnog aan de minimale vereisten voor overgang naar een opvolgend leerjaar zullen gaan voldoen. Er bestaat echter een duidelijke relatie tussen de gemiddelde eindscore voor zowel de taal- als rekentoetsen en het aantal keren dat een leerling gedoubleerd heeft. Hoe vaker een leerling doubleert, hoe lager zijn gemiddelde toetscore. Zie figuur 12. De resultaten uit 2007 toonden een vergelijkbaar verloop.



Figuur 12

Blijkbaar lukt het de doorsnee leerling na doubleren niet om de toetscore van niet gedoubleerde medeleerlingen te evenaren, laat staan te overtreffen. Of de desbetreffende leerlingen na het zitten blijven er desondanks gemiddeld significant op vooruit gegaan zijn valt uit het onderzoek niet op te maken. Uitsluitel hierover kan alleen verkregen worden via een cohort onderzoek. Dat kan niet met het voorliggende cijfermateriaal, aangezien het bij de eindtoets basisonderwijs om geanonimiseerde toetsresultaten gaat, afgenomen op één bepaald moment.

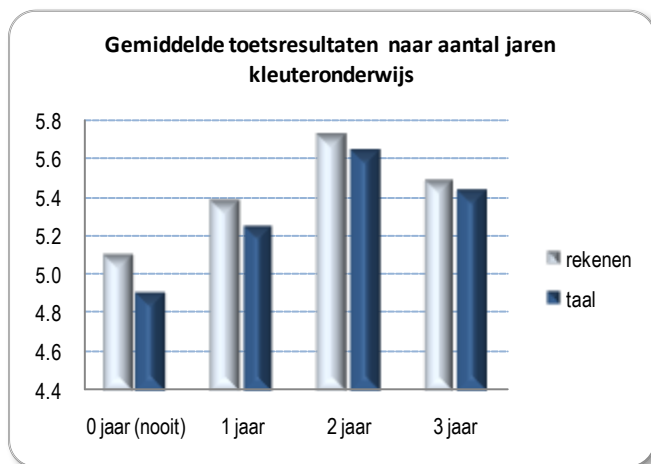
Bezitten doublerende leerlingen min of meer ernstige cognitieve deficiënties of heeft het doubleren meer te maken met persistente gedragsproblemen van deze leerlingen of spelen andere meer externe factoren een voorname rol? Deze belangrijke vraag werd reeds in het vorige verslag van de eindtoets geformuleerd, maar kent tot op heden geen antwoord. Nader onderzoek wordt wederom aanbevolen.

CONCLUSIE 14.

Doubleren en gemiddelde toetsscore voor zowel taal als rekenen zijn relatief sterk gecorreleerd. Gedoubleerde leerlingen presteren lager dan niet gedoubleerde medeleerlingen. Hoe vaker een leerling doubleert, hoe lager zijn gemiddelde toetsscore.

Aantal jaren kleuteronderwijs

In diverse Britse en Amerikaanse publicaties over “early childhood education” wordt gewag gemaakt van het belang van een vroege start in het al dan niet verplichte onderwijs traject (veelal in de “preschool age”). Deze notities vormden de directe aanleiding om in de vragenlijst over de schoolexterne factoren ook het aantal jaren doorlopen kleuteronderwijs op te nemen. Dat bleek uiteindelijk geen loze greep te zijn zoals uit figuur 13 volgt.



Figuur 13

CONCLUSIE 15.

De gemiddelde toetscore voor zowel rekenen als taal neemt toe naarmate het kleuteronderwijs meer volledig is doorlopen. Leerlingen die langer dan twee jaar over het kleuteronderwijs doen⁷, scoren nog altijd hoger dan degenen die slechts één jaar of helemaal geen kleuteronderwijs genoten hebben.

Gebroken gezinnen

In de enquête naar achtergrondfactoren zijn de leerlingen ook gevraagd of ze met hun ouder(s) samenwonen. De antwoordmogelijkheden op deze vraag vindt men in tabel 13, samen met de bijbehorende toetsresultaten voor de reken- en taaltoetsen.

Tabel 13

Toetsscores en gebroken gezinnen

Woon je met je ouder(s) samen?	gemiddelde score		
	rekenen	taal	aantal
Ja, samen met mijn vader en moeder	5.8	5.7	874
Ja, met alleen mijn moeder	5.4	5.3	362
Ja, met alleen mijn vader	5.8	5.8	32
Ja, soms bij vader en soms bij moeder	5.2	5.1	100
Nee, ik woon bij andere familieleden	5.2	5.1	44
Nee, ik woon niet bij familie	5.1	5.5	4

⁷ Deze groep is bijna even groot als de groep die 1 jaar kleuteronderwijs doorlopen heeft (149 tegen 170 leerlingen)

Uit tabel 13 volgt dat aan de eindtoets deelnemende kinderen die samen met hun vader en moeder wonen, gemiddeld hogere toetsresultaten behalen dan kinderen uit gebroken gezinnen. Verder dat kinderen die vast bij één partner wonen, gemiddeld hoger scoren dan kinderen die soms bij de ene dan weer bij de andere partner leven of die niet bij de ouder(s) wonen.

Opmerkelijke uitzondering op de geobserveerde verschillen zijn de kinderen die alleen bij de vader wonen, deze groep scoort voor de toetsen gemiddeld nagenoeg gelijk als de groep leerlingen uit ongebroken gezinnen.

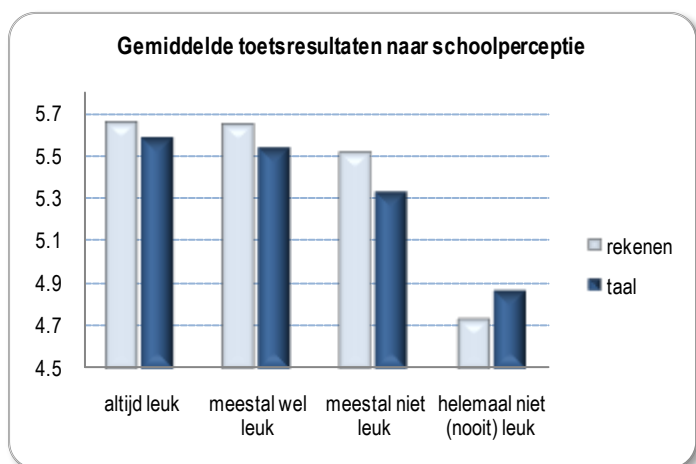
CONCLUSIE 16.

Leerlingen uit gebroken gezinnen scoren gemiddeld duidelijk lager voor zowel de reken- als de taaltoetsen dan leerlingen die met beide ouders samenwonen, uitzondering hierop is de relatief kleine groep leerlingen die alleen met hun vader samen wonen.

Schoolperceptie

Naar verwachting is er een correlatie tussen de schoolresultaten en het al dan niet met plezier naar school gaan. Daarbij wordt voorbijgegaan aan de vraag in hoeverre deze schoolperceptie het gevolg dan wel de oorzaak is van die schoolresultaten. In figuur 14 wordt deze correlatie weergegeven. De figuur geeft de gemiddelde toetsresultaten voor rekenen en taal weer voor de verschillende antwoordmogelijkheden op de vraag “Vind je de school leuk?”. Het percentage van de leerlingen die de school meestal niet of nooit leuk vinden, bedraagt 12% en dat is belangrijk minder dan de 19% die hiervoor in 2007 gemeten werd. Verder is het getoonde verloop praktisch gelijk aan dat voor de resultaten in 2007.

Naast de vraag “Vind je de school leuk?” zijn de leerlingen ook bevraagd in hoeverre ze “Leren en huiswerk maken” leuk vinden en hoe hun beste vrienden over beide vragen denken. De gegeven antwoorden geven echter geen eenduidig coherent beeld.



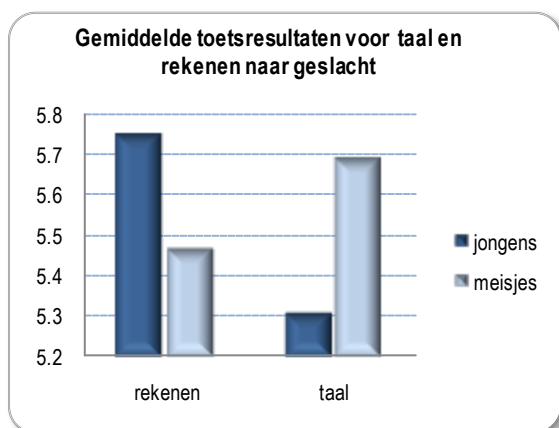
Figuur 14

CONCLUSIE 17.

De toetsresultaten voor rekenen en taal zijn minder naarmate leerlingen de school minder leuk vinden. Het percentage leerlingen die de school meestal niet of nooit leuk vinden bedraagt 12% van alle respondenten en is qua procentuele omvang t.o.v. het jaar 2007 belangrijk geslonken.

Toetsresultaten naar geslacht

De publicatie over de resultaten van de eindtoets basisonderwijs 2007 wees op een interessant en relatief groot seksegerelateerd scoreverschil bij de taal- en rekentoetsen. Een dergelijk duidelijk effect, dat zich niet alleen op nationaal, maar ook schoolniveau manifesteert, laat van nature de vraag opkomen hoe het beeld er voor 2008 uitziet. Het antwoord levert figuur 15. De verschillen wijzen in dezelfde richting en zijn nagenoeg gelijk aan die van 2007. Het ziet er dan ook naar uit dat we hier te maken hebben met een structureel persistent verschijnsel.



Figuur 15

CONCLUSIE 18.

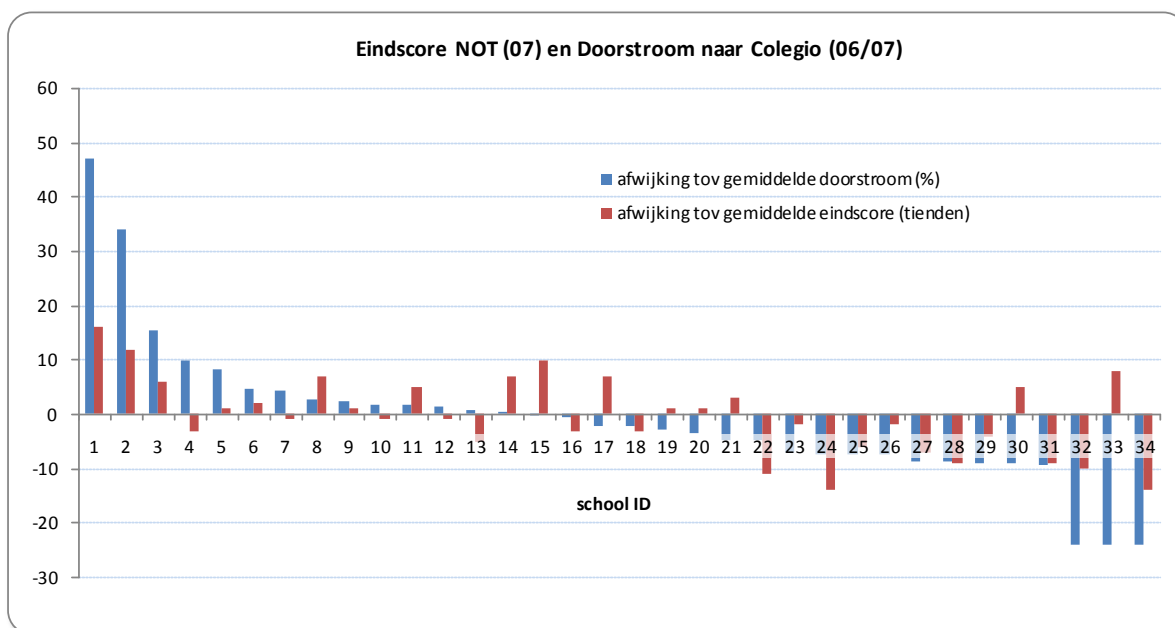
Jongens scoren wederom (t.o.v. 2007) beter in rekenen dan meisjes, terwijl omgekeerd de meisjes beter scoren op taal dan de jongens. Het grootste seksegebonden scoreverschil doet zich voor bij de taaltoetsen, hiervoor scoren de meisjes, net als in 2007, gemiddeld 0.4 punten hoger. Iets minder groot is het gemiddelde scoreverschil tussen jongens en meisjes bij rekenen, dat bedraagt 0.3 punten in 2008 tegenover 0.4 voor 2007.

De bijbehorende histogrammen van zowel de reken- als taalscores zijn als bijlage 3 opgenomen. Deze frequentieverdelingen benaderen allemaal voor beide geslachten redelijk goed de normale verdeling. Ter vergelijking en eventuele verdere analyse zijn ook de vier desbetreffende standaardafwijkingen opgegeven.

Schoolresultaten en doorstroom naar Colegio Arubano (2007)

Één van de directe aanleidingen tot het instellen van een nationale eindtoets basisonderwijs waren de tegenvallende resultaten van groepen leerlingen die het basisonderwijs met relatief goede cijfers verlaten hebben in de ciclo basico van het voortgezet onderwijs (in het bijzonder het Colegio Arubano). De beste manier om dit verschijnsel te onderzoeken is een cohort onderzoek waarbij de resultaten van groepen leerlingen uit het basisonderwijs vergeleken worden met de resultaten van diezelfde groepen in het voortgezet onderwijs. Een dergelijk onderzoek valt echter buiten het kader van een algemene beschrijving van de resultaten eindtoets primair onderwijs. Toch kunnen de schoolresultaten op deze toets een eerste voorzichtige indicatie geven over de omvang van het genoemde probleem. Namelijk als deze schoolresultaten gekoppeld worden aan het percentage leerlingen dat van de desbetreffende school doorgaans naar het voortgezet onderwijs doorstroomt.

In het ideale geval bestaat er een duidelijke correlatie tussen twee schoolranglijsten: één op grond van de schoolresultaten voor de eindtoets basisonderwijs en de ander op grond van het percentage leerlingen dat vanuit klas 6 naar het Colegio Arubano doorstroomt. Er bestaan statistische indicatoren om de correlatiesterkte tussen rangvolgordes in twee verschillende lijsten te meten. Hieronder is echter gekozen voor een visuele benadering omdat deze weergave gemakkelijker te lezen en te interpreteren is. De doorstroomgegevens en de resultaten op de eindtoets primair onderwijs zijn afkomstig uit het jaar 2007.



Figuur 16

In figuur 16 staan de scholen genummerd op de horizontale as en geven de blauwe kolommen aan hoeveel een bepaalde school afwijkt van het gemiddelde doorstroompercentage naar het Colegio Arubano. Dat gemiddelde bedraagt voor alle zesde klas leerlingen in het basisonderwijs 24.1%. Scholen waarbij de blauwe kolom omhoog wijst, scoren hoger dan dit gemiddelde en een omlaaggerichte kolom betekent dat de doorstroom minder dan 24.1% bedraagt. Links staan dus de scholen met de hoogste doorstroom en rechts die met de laagste doorstroom (gemeten voor 06/07).

Ter vergelijking geven de rode kolommen de afwijking aan van de schoolscore t.o.v. het nationale gemiddelde voor rekenen en taal samen van de eindtoets basisonderwijs in 2007 (aangegeven als NOT (07)). Dat

gemiddelde was 5.0 en de afwijking is genomen in tienden van punten. School 1 scoort dus ongeveer 1,6 punten hoger dan het gemiddelde 5.0 en school 34 scoort 1.4 punten lager, oftewel een 3.6.

CONCLUSIE 19.

Momentopname 2007: scholen met een relatief hoog doorstroompercentage van leerlingen naar het Colegio Arubano, scoren in het algemeen ook hoger dan gemiddeld op de eindtoets basisonderwijs en scholen die onder het gemiddelde doorstroompercentage liggen scoren op de eindtoets basisonderwijs in de regel ook onder het gemiddelde.

De meest opvallende uitzonderingen zijn waarschijnlijk: school 33 die relatief goed scoort maar waarvan slechts weinig leerlingen naar het Colegio doorstromen en school 4 waar het omgekeerde het geval is. Buiten die laatste school zijn er niet echt andere scholen aan te wijzen die laag scoren op de NOT en waarvan wel een relatief groot deel van de leerlingen doorstroomt naar het Colegio.

Uiteraard moet bij de interpretatie rekening gehouden worden met de volgende punten:

1. We hebben hier te maken met een momentopname en dan ook nog eens een keer een momentopname van twee verschillende cohorten.
2. De eindtoets basisonderwijs is nog niet geëvalueerd op validiteit en meet lang niet alle competenties van de afzwaaiende basisschoolleerlingen. Bovendien kan ook een valide eindtoets basisonderwijs een vertekend beeld geven, bijvoorbeeld als de eindtermen van het basisonderwijs niet correct aansluiten bij die van de ciclo basico.

4. Aanbevelingen

We evalueren eerst de aanbevelingen zoals omschreven in de rapportage van de eindtoets primair onderwijs uit 2007:

1. Gerichte feedback aan scholen over prestaties ten aanzien van specifieke leerstofonderdelen

Deze feedback is reeds gegeven t.a.v. de resultaten uit 2007, na het uitbrengen van de algemene rapportage. Voor 2008 is deze volgorde omgekeerd: eerst hebben de scholen hun specifieke rapportage ontvangen waarbij de vragen van de reken- en taaltoetsen gebundeld zijn in toetsonderdelen die gekoppeld zijn aan curriculum domeinen/ leerstofonderdelen. Dat heeft als voordeel dat de scholen met deze individuele feedback over de sterke en zwakke kanten van hun onderwijsresultaten direct aan de slag kunnen binnen hun eigen programma's tot kwaliteitsverbetering van het reken- en taalonderwijs op schoolniveau.

2. Naast de schoolexterne factoren ook schoolinterne factoren meenemen in de analyse

De inspectie heeft een goede en zeer uitgebreide vragenlijst samengesteld met betrekking tot de schoolinterne factoren. De gegevens zijn echter dermate individueel dat het niet past deze in een algemene beschouwing te publiceren. Bovendien is, zoals aangegeven in conclusie 8, de variabiliteit in de verschillende antwoordmogelijkheden op de meeste vragen zo groot dat algemene uitspraken of significante (en relevante) verschillen tussen subgroepen nauwelijks/niet te vinden zijn. De gegevens zouden in indirecte zin gebruikt kunnen worden om scholen onderling "eerlijk" te vergelijken. We hebben echter reeds eerder gesteld dat een betrouwbare gedetailleerde absolute schoolrangvolgorde niet mogelijk en ook niet wenselijk is (zie: "Toetsresultaten en schoolcontext"). Vandaar dat deze uitgebreide informatiebron in eerste instantie voorbehouden blijft aan nadere analyses door de inspectie zelf.

3. Koppeling van de onderzoeksbevindingen aan (bekende) gerelateerde problemen

Reeds onderzocht is hoe de in dit onderzoek gemeten schoolprestaties in het primaire onderwijs zich verhouden tot bekende doorstroomgegevens naar de ciclo basico van het Colegio Arubano. Verder heeft zich de relatie tussen het aantal keren doubleren en de toetsresultaten in de 6^{de} klas basisonderwijs wederom gemanifesteerd (conclusie 13). De aanleiding tot nader onderzoek naar de achtergronden van het omvangrijke zittenblijversprobleem wordt hierdoor verder onderbouwd.

4. Toetsevaluatie

De rapporten Eindtoets Primair Onderwijs 2007 en 2008 betreffen een analyse van de toetsresultaten voor rekenen en taal op vooral schoolniveau. Wat ontbreekt en absoluut noodzakelijk is, is een analyse van het toetsinstrument zelf. Aanbevolen werd de op dit gebied aanwezige expertise te bundelen. Initiatieven hiertoe worden op dit moment genomen door het Examenbureau, een eerste rapportage zal binnenkort verschijnen.

5. SEM

De uitbouw van ervaring in het gebruik van "structural equation modeling" middels het programma AMOS (voorbeeld zie bijlage 4) is niet gebeurd omdat de evaluatieperiode van het programma verstreken is en een officiële licentie (nog) ontbreekt.

Naast de niet gerealiseerde aanbevelingen uit 2007 is de belangrijkste aanbeveling na de evaluatie van het proces en de resultaten van de eindtoets basisonderwijs in 2008:

6. Optimalisatie van toetscontextfactoren

Naast het eerder beschreven toetseffect en leereffect worden toetsresultaten ook beïnvloedt door een reeks andere factoren die gemakshalve samengevat kunnen worden onder de term toetscontextfactoren. Een van die factoren zijn frauduleuze handelingen tijdens of na de toets van de kant van de leerlingen of de leerkracht. Bij de analyse van de ijkgegevens, beschreven onder “Toetseffect of leereffect”, is al gebleken dat dit verschijnsel in de praktijk helaas niet uit te sluiten is. Fraude met testresultaten is echter funest voor de validiteit en betrouwbaarheid van de te analyseren gegevens.

Geschikte oplossingen zijn niet gemakkelijk te vinden. Misschien dat op korte termijn een roulatiesysteem van leerkrachten bij de toetsafname dit probleem voornamelijk kan voorkomen. Belangrijker is echter de problematiek bij de wortels aan te pakken. En die liggen op het vlak van onderling vertrouwen denken we. Het gebruik van de resultaten van de eindtoets basisonderwijs door schoolhoofden als persoonlijk repressiemiddel of het door de desbetreffende leerkracht als zodanig ervaren, verkeerd interpreteren, van een poging tot dialoog over deze gegevens zal het beschreven probleem eerder vergoten dan verkleinen. Consensus over de waarde van de eindtoets basisonderwijs voor zowel de betrokken leerlingen als de school en de aard en bedoeling van de feedback over de gescoorde resultaten is essentieel in deze.

Bijlagen

Bijlage 1	Enquête schoolexterne factoren eindtoets primair onderwijs
Bijlage 2	Hulpblad “hoogste behaalde onderwijsniveau ouders”, behorende bij enquêtevraag 9
Bijlage 3	Histogrammen voor de reken- en taaltoetsen, voor jongens en meisjes apart
Bijlage 4	Amos model
Bijlage 5	Model schoolprestaties

+ Vul hier eerst je nummer in (4 cijfers): \longrightarrow

1. Ben je een jongen of een meisje? jongen meisje

+

2. Hoe oud ben je? 10 jaar 11 jaar 12 jaar 13 jaar 14 jaar 15 jaar

3. Waar ben je geboren? Aruba Colombia, Venezuela, Zuid Amerika, Santo Domingo, Jamaica, Caraïbisch gebied
 Nederlandse Antillen, Suriname Nederland, Europa, Verenigde Staten, Canada ander land

4. Welke taal spreek je meestal thuis? Papiamentu Nederlands Spaans Engels anders

5. Hoe lang woon je in totaal al op Aruba? meer dan 10 jaar 8-10 jaar 5-7 jaar 2-4 jaar minder dan 2 jaar

6. Hoeveel broers/zussen die nog naar school gaan, wonen bij je thuis?

Jezelf niet meetellen: 0 1-2 3-4 meer dan 4

7. Woon je met je ouders(s) samen?

Ja, samen met mijn vader en moeder Ja, met alleen mijn moeder Ja, met alleen mijn vader
 Ja, soms bij vader en soms bij moeder Nee, ik woon niet bij familie Nee, ik woon bij andere familieleden

8. Heb je een eigen slaapkamer, voor jou helemaal alleen?

ja nee **+**

9. Wat is de hoogste school (opleiding) die je vader én moeder helemaal hebben afgemaakt? (uitleg door je juf of meneer)

	<div style="display: flex; justify-content: space-between; border-top: 1px solid black; border-bottom: 1px solid black; padding: 5px 0;"> </div>	
vader	<input type="checkbox"/> Basisonderwijs of minder, Educación Primaria o menos, Primary Education or less <input type="checkbox"/> Lager Beroepsonderwijs, Educación Secundaria (Baja y Alta), Secondary Education (Lower and Higher) EPB, Educación (pre)Profesional o (pre)Técnica <input type="checkbox"/> Algemeen Voortgezet Onderwijs MAVO, HAVO, VWO, Bachillerato, (Junior) High School (CSE, GCE, GCSE) <input type="checkbox"/> Middelbaar Beroepsonderwijs EPI, Título Intermedio, Associate Degree <input type="checkbox"/> Hoger Beroepsonderwijs/Universiteit, Graduado y Posgrado, university or postgraduate degree dr., drs., ir., ing., Ma, Ba, PhD, Arquitecto, Licenciado, Diplomado, Médico, ...	moeder

+

ENQUÊTE SCHOOLEXTERNE FACTOREN EINDTOETS PRIMAIR ONDERWIJS 2008 - CBS/DO

+

+



+

10. Praten je ouders met jou over je huiswerk en je proefwerken? ja nee
11. Vinden je ouders het erg als je onvoldoendes haalt? ja nee
12. Hoe lang heb je op de kleuterschool gezeten? 0 jaar (nooit) 1 jaar 2 jaar 3 jaar
13. Ben je blijven zitten in de basisschool? nee, nooit ja, 1 keer ja, 2 keer ja, 3 keer
14. Waar blijf je meestal na school? (alleen één hokje aankruisen) Ik blijf meestal:
- | | | | |
|--|---|--|---|
| <input type="checkbox"/> bij mijn ouder(s) | <input type="checkbox"/> bij andere familieleden | <input type="checkbox"/> bij Tramerdia of bijles | <input type="checkbox"/> bij onze live-in (oppas) |
| <input type="checkbox"/> bij kennissen/buren | <input type="checkbox"/> zonder volwassene(n) thuis | <input type="checkbox"/> op straat (spelen) met vrienden | <input type="checkbox"/> anders |
- 15a. Hoeveel dagen in de week maak je huiswerk?
- | | | | | |
|--|----------------------------------|----------------------------------|---|----------------------------------|
| <input type="checkbox"/> 0 dagen (nooit) | <input type="checkbox"/> 1 dag | <input type="checkbox"/> 2 dagen | + | <input type="checkbox"/> 3 dagen |
| <input type="checkbox"/> 4 dagen | <input type="checkbox"/> 5 dagen | <input type="checkbox"/> 6 dagen | | <input type="checkbox"/> 7 dagen |
- 15b. Hoe lang ben je dan per dag ongeveer bezig met huiswerk?
- | | | | | |
|---|--------------------------------|--------------------------------|---------------------------------|--|
| <input type="checkbox"/> 1 kwartier of minder | <input type="checkbox"/> ½ uur | <input type="checkbox"/> 1 uur | <input type="checkbox"/> 1½ uur | <input type="checkbox"/> 2 uur of meer |
|---|--------------------------------|--------------------------------|---------------------------------|--|
- 16a. Hoeveel dagen in de week besteed je in totaal aan TV /computerspelletjes? (nintendo, playsation, etc.)
- | | | | |
|--|----------------------------------|----------------------------------|----------------------------------|
| <input type="checkbox"/> 0 dagen (nooit) | <input type="checkbox"/> 1 dag | <input type="checkbox"/> 2 dagen | <input type="checkbox"/> 3 dagen |
| <input type="checkbox"/> 4 dagen | <input type="checkbox"/> 5 dagen | <input type="checkbox"/> 6 dagen | <input type="checkbox"/> 7 dagen |
- 16b. Hoe lang ben je dan per dag ongeveer bezig hiermee?
- | | | | | |
|---|--------------------------------|--------------------------------|---------------------------------|--|
| <input type="checkbox"/> 1 kwartier of minder | <input type="checkbox"/> ½ uur | <input type="checkbox"/> 1 uur | <input type="checkbox"/> 1½ uur | <input type="checkbox"/> 2 uur of meer |
|---|--------------------------------|--------------------------------|---------------------------------|--|
17. Vind je leren en huiswerk maken leuk?
- | | | | |
|--------------------------------------|---|--|---|
| <input type="checkbox"/> altijd leuk | <input type="checkbox"/> meestal wel leuk | <input type="checkbox"/> meestal niet leuk | <input type="checkbox"/> helemaal niet (nooit) leuk |
|--------------------------------------|---|--|---|
18. Vind je de school leuk?
- | | | | |
|--------------------------------------|---|--|---|
| <input type="checkbox"/> altijd leuk | <input type="checkbox"/> meestal wel leuk | <input type="checkbox"/> meestal niet leuk | <input type="checkbox"/> helemaal niet (nooit) leuk |
|--------------------------------------|---|--|---|
19. Hoe leuk vinden je beste vrienden leren en huiswerk maken?
- | | | | |
|--------------------------------------|---|--|---|
| <input type="checkbox"/> altijd leuk | <input type="checkbox"/> meestal wel leuk | <input type="checkbox"/> meestal niet leuk | <input type="checkbox"/> helemaal niet (nooit) leuk |
|--------------------------------------|---|--|---|
- +
20. Hoe leuk vinden je beste vrienden de school?
- | | | | |
|--------------------------------------|---|--|---|
| <input type="checkbox"/> altijd leuk | <input type="checkbox"/> meestal wel leuk | <input type="checkbox"/> meestal niet leuk | <input type="checkbox"/> helemaal niet (nooit) leuk |
|--------------------------------------|---|--|---|
- +

Hoogste behaalde onderwijsniveau ouders

Highest attained parental educational level

Nivel máximo de educación alcanzado por los padres

- Dit is vraag 9 uit de enquête behorende bij de eindtoets primair onderwijs.
 - Laat deze vraag thuis door je ouders beantwoorden, (vader én moeder)
 - neem dan dit papier mee naar school en
 - vul vraag 9 op dezelfde manier in.

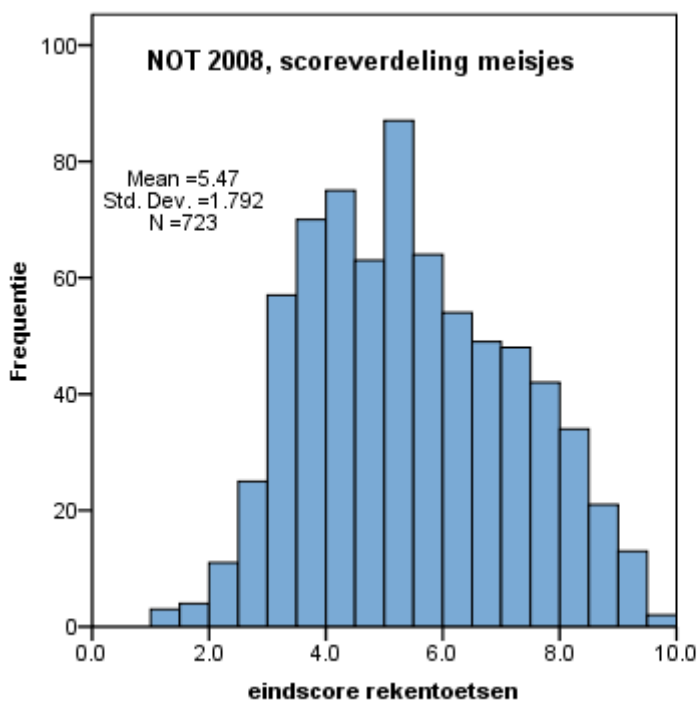
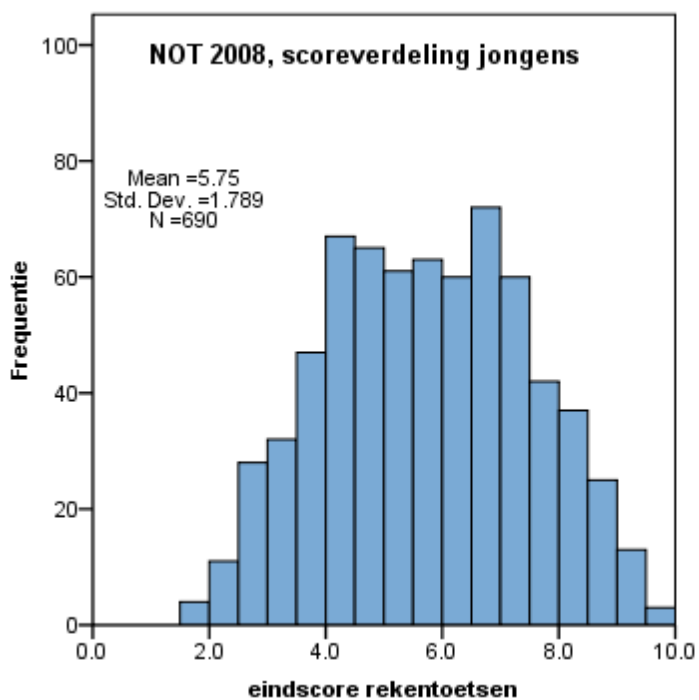


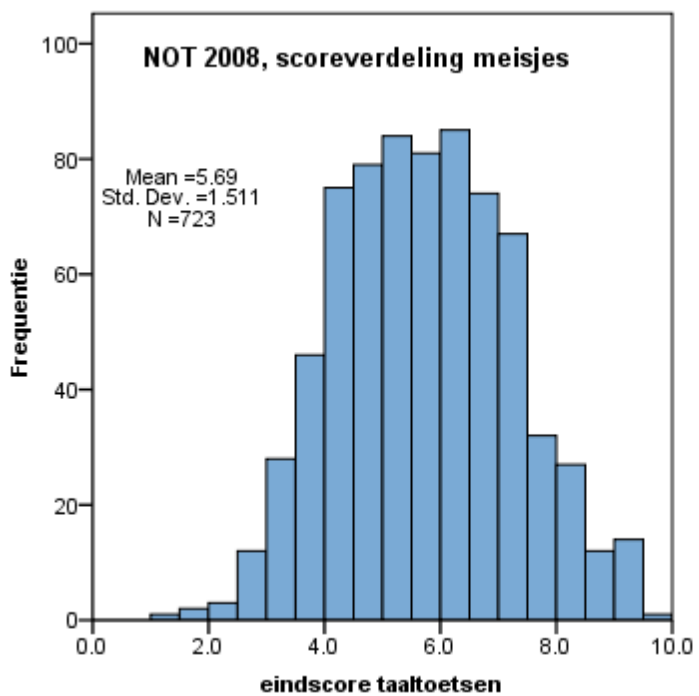
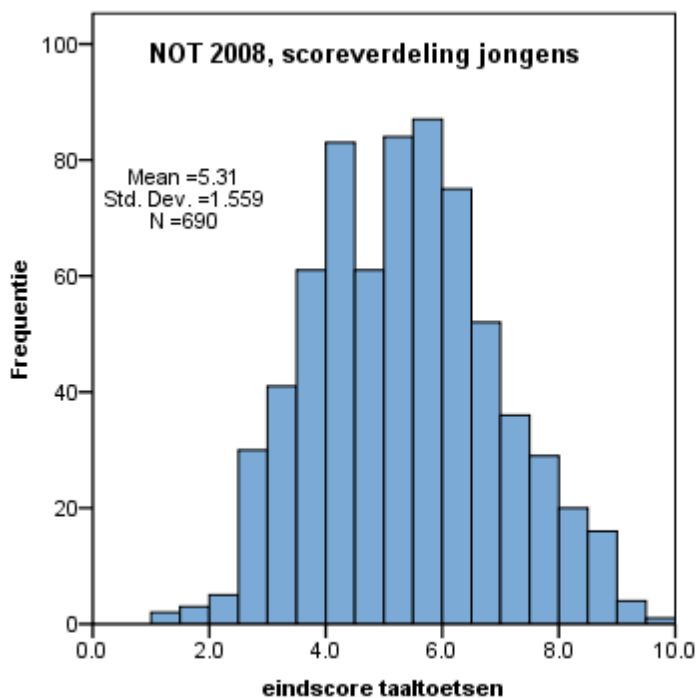
9. Wat is de *hoogste school (opleiding) die je vader én moeder helemaal hebben afgemaakt?* (uitleg door je juf of meneer)

vader		moeder
<input type="checkbox"/>	Basissonderwijs of minder / Educación Primaria o menos / Primary Education or less	<input type="checkbox"/>
<input type="checkbox"/>	Lager Beroepsopleiding / Educación Secundaria (Baja y Alta) / Secondary Education (Lower and Higher) <i>EPB / Educación (pre)Profesional o (pre)Técnica</i>	<input type="checkbox"/>
<input type="checkbox"/>	Algemeen Voortgezet Onderwijs	<input type="checkbox"/>
<input type="checkbox"/>	<i>MAVO, HAVO, VWO / Bachillerato (Junior) High School (CSE, GCE, GCSE)</i>	<input type="checkbox"/>
<input type="checkbox"/>	Middelbaar Beroepsopleiding	<input type="checkbox"/>
<input type="checkbox"/>	<i>EPI / Título Intermedio / Associate Degree</i>	<input type="checkbox"/>
<input type="checkbox"/>	Hoger Beroepsopleiding of Universiteit / Graduado y Posgrado / university or postgraduate degree	<input type="checkbox"/>
	<i>dr., drs., ir., ing., Ma, Ba, Phd, Arquitecto, Licenciado, Médico, ..</i>	

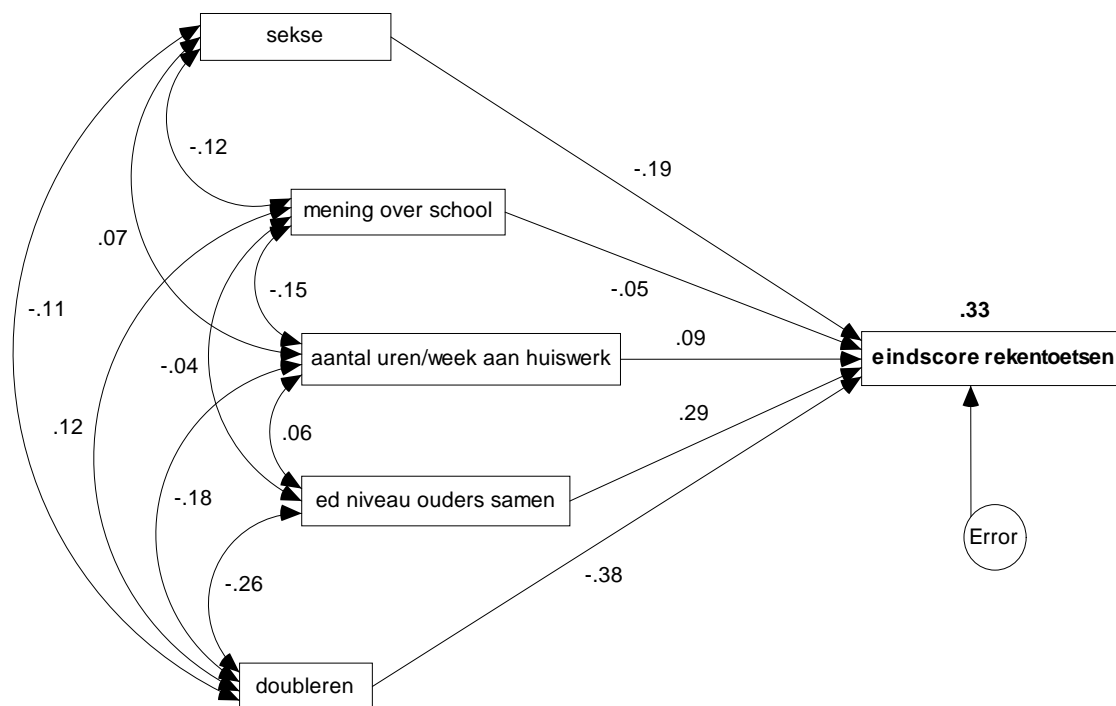
hoogste niveau alleen aankruisen indien men een (of meer) van deze titels bezit
 highest level is only applicable if one (or more) of these degrees is (are) obtained
 el nivel más alto, es solamente aplicable in caso de haber obtenido uno (o más) de estos títulos

Histogrammen voor de reken- en taaltoetsen, voor jongens en meisjes apart.





Een gebruikt AMOS model:



In dit model wordt onderzocht hoe de aan de linkerkant van de figuur aangegeven gemeten variabelen invloed uitoefenen op de eindscore van de rekentoetsen. De enkelvoudige pijlen geven directie invloedsfactoren aan, de dubbele pijlen veronderstelde correlaties tussen de invloedsvariabelen. Alle modellen, en in het bijzonder het bovenstaande relatief eenvoudige model, zijn slechts een benaderende weergave van de complexe werkelijkheid. Zo zal de rekenscore, naast de aangegeven invloedsvariabelen, ook nog door een groot aantal niet aangegeven (niet gemeten) invloedsfactoren bepaald worden. Denk bijvoorbeeld aan motivatie, aanleg en kwaliteit van het onderwijs. Al deze niet meegenomen factoren worden samengevat in de “error” factor. Het boven de eindscore vermelde getal 0.33 geeft aan dat de variatie in de eindscore voor 33% verklaard kan worden uit de aangegeven invloedsfactoren.

De getallen boven de enkelvoudige pijlen zijn de gestandaardiseerde⁸ regressiegewichten, zij geven de relatieve sterkte weer tussen de gebruikte invloedsfactoren. Zo is in het bovenstaande model de sterkste invloedsfactor op de eindscore voor rekenen het doubleren, gevolgd door het gezamenlijk educatieniveau van de ouders. De getallen bij de dubbele pijlen geven de gestandaardiseerde correlatiecoëfficiënten weer. Zo zijn bijvoorbeeld het educatieniveau van de ouders en het aantal keren dat de desbetreffende leerlingen doubleren relatief sterk negatief gecorreleerd.

⁸ dit betekent dat de waarde van de regressiecoëfficiënt onafhankelijk is van de gekozen meeteenheid

Model voor schoolprestaties, met de nadruk op schoolexterne factoren

overgenomen uit: "Het probleem van de overjarige basisschoolleerlingen in nader perspectief", CBS, 2004

