

# Leerpotentieel in de onderbouw

‘Het leerpotentieel van leerlingen uit groep 1 in beeld gebracht en vergeleken met het huidige  
beoordelingssysteem’

Dennis de Wit – 1166441

Educational Studies

Juli 2013

Eerste lezer - Dr. Nadira Saab

Tweede lezer – Dr. Linda Leijenhorst

## Samenvatting

In deze studie wordt het verschil onderzocht tussen scores op leerpotentie-testen en het beoordelingssysteem dat in het basisonderwijs wordt gehanteerd. Leerlingen in het basisonderwijs krijgen lesstof aangeboden aangepast aan hun niveaugroep, waarin zij worden ingedeeld aan de hand van (Cito)toets scores en de niveaubeoordeling van de leerkracht. Bij 124 leerlingen uit groep 1 werden scores op twee onderdelen van de Cito-toets 'Rekenen voor kleuters' en de leerkrachtbeoordeling vergeleken met twee testen waarmee het leerpotentieel met betrekking tot rekenen in beeld wordt gebracht. Een significante relatie tussen de scores op de leerpotentie-testen en het beoordelingssysteem werd niet gevonden. Echter een aantal leerlingen die in het huidige beoordelingssysteem laag scoren, hebben door middel van de leerpotentie-testen laten zien in korte tijd leerstappen te kunnen maken. Om deze reden zouden zij in aanmerking komen voor lesaanbod op een hoger niveau. Gezien het feit dat de leerpotentie-testen niet op validiteit getest zijn, kan de leerpotentieel niet met zekerheid worden vastgesteld.

## Inhoudsopgave

Introductie.....	4
Methode.....	9
<i>Onderzoeksgroep</i> .....	9
<i>Procedure</i> .....	10
<i>Meetinstrumenten</i> .....	10
<i>Leerpotentie-testen</i> .....	11
<i>Cito-toetsen</i> .....	12
<i>Leerkrachtbeoordeling</i> .....	12
<i>Analysemethoden</i> .....	13
<i>De variabelen</i> .....	13
<i>De analyses</i> .....	13
Resultaten.....	15
<i>Leerpotentie-test Camelot</i> .....	15
<i>Leerpotentie-test Vergelijken</i> .....	19
<i>Leerkrachtbeoordeling</i> .....	22
<i>Cito-onderdelen Meetkunde en Getalbegrip</i> .....	26
<i>Opvallende leerlingen</i> .....	27
Discussie.....	28
<i>Beperkingen</i> .....	35
<i>Aanbevelingen</i> .....	36
<i>Conclusie</i> .....	36
Referenties.....	38
Bijlage I. Meetinstrument Camelot.....	42
Bijlage II. Meetinstrument Vergelijken en Tellen.....	48
Bijlage III. Kruistabel Camelot en Leerkrachtbeoordeling.....	55
Bijlage IV. Kruistabel Vergelijken en Leerkrachtbeoordeling.....	56
Bijlage V. Overzicht laag / hoog scorende leerlingen.....	57
Bijlage VI. Gemiddelde scores op alle onderdelen.....	58

## Introductie

De kwaliteit van het onderwijs in Nederland kan worden beschouwd als voldoende. Voor een kenniseconomie zoals die in Nederland bestaat is voldoende echter niet genoeg. Op dit moment is er sprake van een stilstand in de ontwikkeling van het onderwijs en stilstand kan in dit geval worden beschouwd als achteruitgang (Pearson, 2012). Het Nederlandse onderwijs is geschikt voor de opleiding van leerlingen tot basisoniveau, maar excelleert niet (Pearson, 2012). De Inspectie van Onderwijs stelt in het onderwijsverslag over het schooljaar 2011/2012 dat basiskwaliteit geen einddoel mag zijn (Inspectie van het Onderwijs, 2013, p. 8). Voornamelijk bij de groep leerlingen op basisoniveau, is vaak ruimte voor verbetering, dat onder andere tot stand gebracht zou moeten worden door een verbetering in het onderwijs op het gebied van afstemming en maatwerk (Inspectie van het Onderwijs, 2013). Om dit te realiseren moet het aanbod voor leerlingen meer op het individuele leerniveau worden afgestemd, opdat het maximaal haalbare uit het onderwijs en uit de kinderen wordt gehaald. Het feit dat het aanbod voor leerlingen nog niet volledig op het individuele leerniveau is afgestemd, betekent wellicht dat er leerlingen in het Nederlandse onderwijssysteem zijn die onder hun kunnen presteren (Onderwijsraad, 2007). Nederland neemt op de ranglijst met de veertig beste onderwijssystemen in de wereld de zevende plaats in (Economist Intelligence Unit, 2012). Nederland scoort dus bovengemiddeld, maar er is ruimte voor verbetering. Met het aanpassen van het leeraanbod op basis van de individuele potentie van de leerling, zou een hogere plaats op de ranglijst haalbaar zijn (Pfeiffer, 2011). In dit onderzoek zal worden onderzocht of er leerlingen zijn die zich op basis van hun leerpotentieel op een hoger niveau bevinden dan waarop zij in het huidige beoordelingssysteem worden ingedeeld. Leerlingen in het basisonderwijs krijgen lesstof aangeboden aangepast aan de niveaugroep waarin zij zijn ingedeeld (Struiksma, 2012). Deze indeling wordt op geleide van (Cito)toets scores en de beoordeling van de leerkracht gemaakt (Struiksma, 2012). De leerkracht baseert dit op leerresultaten en observaties (Struiksma, 2012).

Om leerlingen op hun eigen niveau te kunnen laten leren wordt in het onderwijs gebruik gemaakt van differentiatie (Inspectie van het Onderwijs, 2010). Differentiatie speelt een belangrijke rol in het passend onderwijs dat steeds meer gebruik wordt (Keesenberg, 2008). Deze vorm van onderwijs houdt in dat zo veel mogelijk leerlingen regulier onderwijs volgen en dat doorverwijzen naar speciaal onderwijs alleen gebeurt als het echt nodig is. Het passend onderwijs heeft er toe geleid dat binnen het reguliere onderwijs veel aandacht wordt besteed aan leerlingen met leerproblemen (Rijksoverheid, z.j.). In het reguliere onderwijs zijn echter ook leerlingen zonder

leerproblemen die baat zouden hebben bij extra onderwijsaanbod met meer uitdaging (van Eijl, 2005). Differentiatie zou in deze gevallen kunnen worden toegepast. Wanneer door de leerkracht wordt gekozen om te differentiëren, kan dit op vier verschillende manieren plaatsvinden (Inspectie van het Onderwijs, 2010). De eerste manier is het aanbieden van extra leer- en verrijkingsstof, waarbij leerlingen die meer uitdaging behoeven dan het standaard lesprogramma extra lesmateriaal aangeboden krijgen op een hoger niveau. Een verrijkingsprogramma is erop gericht om de (begaafde) leerling verder te ontwikkelen (Moon & Rosselli, 2000). De tweede manier is het indikken en compact maken van de lesstof, waarbij de leerling een geselecteerd gedeelte van de lesstof maakt, zodat hij sneller door de stof heen kan gaan. De tijd die een leerling hierdoor overhoudt, kan besteed worden aan verrijkingsmateriaal. De derde manier is het plaatsen van de leerling in een aparte klas met meerdere (hoog)begaafde leerlingen; een zogenaamde plusgroep. Dit gebeurt vaak een aantal dagdelen in de week en wordt gecombineerd met een verrijkt onderwijsaanbod in de eigen klas. De vierde manier om te differentiëren is het 'versnellen', dat inhoudt dat de leerling een klas overslaat en dus sneller het traject van de basisschool aflegt. In groepen 1 en 2 komt met name het 'verrijken' en 'versnellen' voor (Inspectie van het Onderwijs, 2010). Leerlingen krijgen dus eerder moeilijkere stof aangeboden of slaan eerder een klas over, dan dat zij de lesstof compacter krijgen aangeboden. Het aanbod op hoger niveau is afhankelijk van hoge scores op de (Cito)toetsen en het oordeel van de leerkracht. Een leerling die aan de hand van het resultaat op de Cito-toetsen en de leerkrachtbeoordeling geen hoge score behaalt, zal niet worden geselecteerd voor verrijkingsmateriaal. Het selecteren van leerlingen voor verrijkingsprogramma's wordt vaak gebaseerd op statische testresultaten (Doolaard & Oudbier, 2010) zoals de Cito-toets of intelligentietesten. Statische testen zijn momentopnames en weerspiegelen het feitelijke niveau van de ontwikkeling (Vygotsky, 1978). Bij deze manier van testen ligt de nadruk op de kennis die in het verleden is opgedaan en wordt aan het leren van het heden en de toekomst geen waarde gehecht (Resing, 2006). Niet alleen in Nederland wordt op deze manier geselecteerd. Zo wordt in het onderzoek van McClain en Pfeiffer (2012) aangetoond dat de meerderheid van de Verenigde Staten gebruik maakt van statische intelligentietesten om te meten of een leerling meer begaafd is en verrijking nodig heeft. Bij het selecteren van een leerling die begaafd is of verrijking nodig heeft, wordt door leerkrachten soms teveel vastgehouden aan de kwantitatieve aard van IQ-scores (Borland, 2009). Wanneer enkel deze scores zouden worden gebruikt, wordt de grens van de toelating tot verrijkingsprogramma's voor leerlingen op één score, namelijk 130, gesteld (Borland, 2009). Dit leidt ertoe dat enkel aan de harde grens wordt vastgehouden en niet wordt gekeken naar de eventuele doorgroeimogelijkheden van een leerling. Voorgaand onderzoek van Fraser, Walberg, Welch & Hattie (1987) toont aan dat er een hoge correlatie van .71 is tussen het IQ en de

schoolprestaties van leerlingen. Het intelligentiequotum kan worden beschouwd als een indicatie van het intelligentieniveau van een persoon (Dijk & Tellegen, 2004). Intelligentie wordt gedefinieerd als de mogelijkheid van een persoon om zijn vaardigheden op zowel analytisch, creatief als praktisch gebied succesvol te beheersen (Sternberg, 2002). 50% van de variantie van de leerprestaties kan verklaard worden door het IQ van een leerling. In 50% van de variantie kunnen de leerprestaties dus niet door de gemeten intelligentie worden voorspeld. Is het mogelijk dat de leerprestaties voorspeld zouden kunnen worden door andere factoren zoals het leerpotentieel (Calero, Belen & Robles, 2011)? In dit onderzoek wordt gekeken of er een relatie is tussen het leerpotentieel en de leerprestaties. Bij leerpotentieel wordt gekeken naar de mate waarin iemand getraind kan worden en zijn prestatie kan verbeteren (Babad & Budoff, 1974). Er wordt inzicht verkregen in de leerstappen die een kind kan maken. Zijn statische testen wel geschikt om het niveau van presteren in kaart te brengen en kunnen zij een aanleiding geven om te differentiëren op het niveau van de leerling? Het selecteren van leerlingen voor verrijkmateriaal kan aan de hand van leerpotentieel worden gedaan. De door Vygotsky (1978) beschreven zone van de naaste ontwikkeling kan als graadmeter dienen voor intelligentie, gezien het feit dat dit namelijk naast het feitelijke niveau van de ontwikkeling ook de mogelijkheid om te groeien weergeeft (Kingma & Tomic, 1997). Leerpotentieel is een voorspeller voor de ontwikkeling van een leerling en kan ook als voorspeller van hoogbegaafdheid kan worden gezien (Calero, et al., 2011). In dit onderzoek wordt het leerpotentieel van leerlingen vergeleken met de Cito-scores en de indeling op niveau door de leerkracht.

Indien bij het selecteren van leerlingen voor verrijgingsprogramma's enkel naar de statische gegevens wordt gekeken, bestaat de kans dat onderpresterende leerlingen niet worden uitgekozen om mee te doen aan verrijgingsprogramma's, omdat er een discrepantie tussen de schoolprestaties en het cognitief potentieel kan heersen bij onderpresterende leerlingen (Doolaard & Oudbier, 2010; Mulder, Roeleveld & Vierke, 2007). Als deze leerlingen op basis van het niveau wat zij op de statische testen laten zien niet worden geselecteerd, kan dit betekenen dat deze leerlingen leerstofaanbod beneden hun daadwerkelijke niveau krijgen aangeboden. De volgende zeven manieren kunnen gebruikt worden om onderpresteren te meten (Mulder et al., 2007): 1) discrepantie tussen cognitief potentieel en schoolprestaties 2) relatieve positie van een leerling op taal en rekenen ten opzichte van het landelijk gemiddelde 3) discrepantie tussen schoolprestaties en het adviseringsbureau voor een opleiding in het voortgezet onderwijs 4) discrepantie tussen schoolprestatie en de keuze voor een opleiding in het voortgezet onderwijs 5) het oordeel van de leerkracht: vindt hij / zij dat de leerling beter kan presteren dan deze doet? 6) discrepantie tussen de verwachting van de leerkracht en de positie van de leerlingen later in de schoolloopbaan 7) discrepantie tussen de verwachtingen

van de leerkrachten en de feitelijke prestaties van de leerlingen. De manier die als eerste wordt beschreven, onderpresteren als discrepantie tussen cognitief potentieel en schoolprestaties, zal in dit onderzoek worden gehanteerd. In het rapport van de Onderwijsraad (2007) wordt beschreven dat onderpresteren kan worden gezien als een verschil tussen resultaten van een intelligentietest en feitelijke prestaties op schoolse vakken. Een leerling heeft bijvoorbeeld een hoge score op een intelligentietest behaald, maar behaalt lage prestaties op de schoolse vakken. Leerlingen die onderpresteren lopen dus het risico niet geselecteerd te worden voor verrijgingsprogramma's. In dit onderzoek wordt onderzocht of door middel van het meten van het leerpotentieel een grotere groep leerlingen meer uitdaging nodig heeft en voor verrijgingsprogramma's in aanmerking zou kunnen komen dan de groep die nu geselecteerd wordt. Onderzoek van de Onderwijsraad (2007) heeft uitgewezen dat de omvang van het percentage onderpresterende leerlingen varieert van 10 tot 18% van de leerlingengroep, wat betekent dat in iedere klas een groep onderpresterende leerlingen aanwezig is. Oplossingen die de Onderwijsraad aandraagt om dit aantal te verkleinen zijn meer (effectieve) leertijd inzetten voor leerlingen, een gedifferentieerder aanbod in de vorm van maatwerk en meer onderwijstijd voor leerkrachten.

Verwachtingen van leerkrachten over de prestaties van leerlingen die onderpresteren zijn anders dan de verwachtingen van leerlingen die scores boven het gemiddelde behalen en geselecteerd worden voor verrijkmateriaal (de Boer, Bosker & van der Werf, 2010). Leerkrachten zullen meer verwachten van een leerling die goede cijfers behaalt, dan een leerling die cijfers beneden het gemiddelde behaalt (de Boer et al., 2010). Op het moment dat een onderpresterende leerling in beeld wordt gebracht, kan het niveau van deze leerling verhoogd worden door te anticiperen op het niveau en de behoeften van deze leerling (van Wieringen, 2011). Nu vallen deze onderpresterende leerlingen buiten de groep die wordt geselecteerd voor verrijkmateriaal. Binnen de groep onderpresterende leerlingen die niet wordt gesignaleerd, bevinden zich ook leerlingen die worden overgeslagen bij het selecteren voor plusactiviteiten vanwege het risico op ordeverstoringen. Omdat onderpresterende leerlingen in het huidige onderwijs vaak niet op niveau worden aangesproken is er een groot risico op demotivatie, met onderpresteren en eventuele gedragsproblemen als gevolg (Doornekamp, Drent en Bronkhorst, 1999). Uit onderzoek van Kim en VanTassel-Baska (2010) blijkt dat het creatieve vermogen van onderpresterende leerlingen vaak gepaard gaat met gedragsproblemen en dat er onderzoek moet worden gedaan naar dit gedrag om zo een manier te vinden waarop de kwaliteiten van deze leerlingen benut kunnen worden.

In dit onderzoek wordt door middel van dynamisch assessment (Grigorenko & Sternberg, 1998) onderzoek gedaan naar het leerpotentieel van leerlingen uit groep 1. Dynamisch assessment

maakt het mogelijk om de verandering van prestatie van een leerling in beeld te brengen (Tzurriel, 2000). Aangezien de band die een leerling met de leerkracht kan krijgen gedurende de schooltijd de prestaties van de leerling kan beïnvloeden (Spilt, Hughes, Wu & Kwok, 2012), werd gekozen voor leerlingen uit groep 1. Dit heeft als gevolg dat het leerpotentieel dat in dit onderzoek wordt gemeten, het leerpotentieel van het individuele kind zelf optimaal weergeeft (Spilt et al., 2012). Om het leerpotentieel te meten zijn meetinstrumenten samengesteld waarmee de vaardigheden van leerlingen op het gebied van visueel-ruimtelijk inzicht en het vergelijkend, tellend vermogen vastgesteld kunnen worden. Het leerpotentieel van de leerlingen wordt op twee manieren gemeten. Als eerste kan leerpotentieel worden gezien als de mate waarin een leerling in staat is om te profiteren van de instructie die het krijgt (Resing & Drenth, 2007) en wordt de groei die het kind laat zien na het ontvangen van instructie gemeten. Als tweede kan leerpotentieel worden gezien als het aantal hulpstappen dat een kind nodig heeft om tot de oplossing van een opgave te komen. Dit betekent dus dat naarmate het kind minder hulpstappen nodig heeft om tot de goede oplossing te komen, er sprake is van een hoger leerpotentieel (Resing, 2006). Naast het afnemen van leerpotentieel-testen worden resultaten verzameld van de Cito-toetsen van de leerlingen en worden leerkrachten gevraagd de leerlingen in niveaugroepen in te delen. Om te achterhalen of leerlingen een hogere score behalen op de leerpotentieel-testen dan hun scores op de Cito-toets en de leerkrachtenbeoordeling zouden doen vermoeden, worden deze scores met elkaar vergeleken. De vraag die centraal staat in dit onderzoek luidt: 'Wat is de relatie tussen de leerpotentieel-scores bij kleuters uit groep 1 en het beoordelingssysteem wat scholen hanteren om leerlingen op niveau in te delen?'. De hypothese hierbij is dat er leerlingen zijn die op basis van hun leerpotentieel een hoger niveau van zichzelf laten zien dan dat zij doen in het huidige beoordelingssysteem (Pfeiffer, 2011). Om erachter te komen of er relaties bestaan tussen de leerpotentieel-testen, de Cito-onderdelen en de leerkrachtenbeoordeling, worden ze alle drie met elkaar vergeleken. Hiervoor zijn de volgende deelvragen opgesteld.

- Wat is het verschil tussen de leerpotentiescores en de beoordeling van de leerkracht?
- Wat is het verschil tussen de twee leerpotentiescores en de vaardigheidsscores op de twee Cito-onderdelen?
- Wat is het verschil tussen de beoordeling van de leerkracht en de vaardigheidsscores op de twee Cito-onderdelen en wat voor verschil is dit?

Tot op heden is niet bekend of er meer leerpotentieel in leerlingen schuilt dan nu in het onderwijs naar voren komt. In voorgaand onderzoek is de mogelijkheid van het inzetten van leerpotentieel-testen om leerpotentieel in beeld te brengen bij leerlingen onderzocht, maar werd



nooit een vergelijking gemaakt met zowel de leerkrachtbeoordeling als de Cito-scores (Tzurriel, 2000; Calero et al., 2011) zoals in dit onderzoek gebeurt. Zaal (2012) schrijft dat de Cito-toetsen een beperkte representatie van de schoolprestaties van leerlingen zijn, echter zijn deze toetsen momentopnames en een statische vorm van toetsen. Er wordt geen rekening gehouden met de mate waarin een leerling zich ontwikkelt, maar enkel met het feitelijke leerniveau op dat moment (Zaal, 2012). Er is een kans dat leerlingen beschikken over een hoger leerniveau, dan in het huidige systeem wordt gesignaleerd (Zaal, 2012; Pfeiffer, 2011). Dit onderzoek tracht dit in beeld te brengen, zodat in de toekomst kan worden geanticipeerd op deze groep leerlingen en het leerrendement verhoogd kan worden. Wanneer een leerling een hoog leerpotentieel bezit, kan deze leerling meer uitgedaagd worden en lesstof op een hoger niveau worden aangeboden. Op deze manier zou het maximale niveau van leerlingen wellicht bereikt kunnen worden en zou het rendement van het Nederlandse onderwijs kunnen worden verbeterd.

## **Methode**

### *Onderzoeksgroep*

De onderzoeksgroep bestaat uit 124 leerlingen uit groep 1, waarvan 69 meisjes en 55 jongens. De gemiddelde leeftijd is vier jaar. Deze leerlingen zijn afkomstig uit elf verschillende kleutergroepen, waarvan vijf van school 1 (N=59) en zes van school 2 (N=65). Niet alle leerlingen uit de kleutergroepen hebben meegedaan met het onderzoek, enkel de leerlingen die volgend schooljaar naar groep 2 gaan, worden in het onderzoek meegenomen. Er is voor deze leeftijdsgroep gekozen, omdat bij deze groep de kans het grootst is dat bij alle leerlingen het interventiepunt binnen de testopzet valt. Het niveau van alle leerlingen zal het beste in groep 1 binnen één test vallen. In bijvoorbeeld groep 6 is er meer kans op verschil in niveau. Daarnaast is het zo dat hoe jonger de leerling is, hoe minder kans er is op invloed van buitenaf (Spilt et al., 2012). De thuistaal van de leerlingen is geïnventariseerd. Het blijkt dat 80% van de leerlingen thuis Nederlands spreekt, 10% spreekt thuis Nederlands gecombineerd met hun moedertaal en 10% spreekt thuis alleen hun moedertaal (Turks, Marokkaans of Spaans). Er is binnen dit onderzoek enkel getest op het gebied van rekenen, zodat leerlingen met een eventuele taalachterstand hier geen extra problemen van ondervonden. De instructie is wel aan de leerlingen verteld, maar hierbij vond steeds een visuele ondersteuning plaats, door middel van het laten zien van de spelsituaties.

## *Procedure*

Het onderzoek is op twee basisscholen in Rotterdam uitgevoerd. In samenwerking met Stichting Boor (het bestuur voor openbaar onderwijs in Rotterdam) zijn deze twee scholen benaderd. De onderzoeker heeft in samenwerking met een onderwijskundig beleidsmedewerker en een beleidsmedewerker hoogbegaafdheid van Stichting Boor de meetinstrumenten samengesteld. Nadat de scholen hebben toegezegd mee te werken aan het onderzoek, is middels een brief aan de ouders toestemming gevraagd om hun kinderen te testen in dit onderzoek. De onderzoeker heeft na toestemming van de ouders de testen afgenomen met de leerlingen. Het grootste gedeelte is door de onderzoeker zelf afgenomen, maar de onderwijskundig beleidsmedewerker heeft twee ochtenden geholpen met het afnemen van de testen. Dit gebeurde allemaal volgens een vooraf opgestelde standaardprocedure, om verschillen in afname te voorkomen. De testen zijn op school afgenomen in de nabije omgeving van het klaslokaal, zodat de leerlingen zich vertrouwd voelden in de testomgeving. Alle testen zijn op de gang of in een kamertje nabij het klaslokaal afgenomen. Aan de leerlingen is verteld dat ze een spelletje zouden gaan spelen, in plaats van het maken van een test. Door de laagdrempeligheid hebben, op één leerling na, alle leerlingen vrijwillig meegewerkt. De duur van de testafname varieerde per leerpotentietest van 3 tot 15 minuten. Leerlingen maakten nooit de twee testen achter elkaar. Een deel van alle leerlingen maakte de twee testen op dezelfde dag en het grootste deel van de leerlingen maakte de twee testen op twee verschillende dagen.

## *Meetinstrumenten*

Er bestaat geen valide meetinstrument om het leerpotentieel van deze leeftijdsgroep op een juiste manier in kaart te brengen. In deze studie is gebruik gemaakt van twee samengestelde instrumenten, waarvan de resultaten twee leerpotentiescores opleverden. Deze testen zijn niet op validiteit getest, dus er moet rekening worden gehouden dat er niet met zekerheid kan worden gezegd dat deze testen daadwerkelijk het leerpotentieel meten. Daarentegen is dynamisch assessment wel een geschikte manier om leerpotentieel in beeld te brengen (Resing, 2006). Het samenstellen van deze instrumenten is door de onderzoeker in samenwerking met de onderwijskundig beleidsmedewerker en de beleidsmedewerker hoogbegaafdheid gedaan. De instrumenten die beide de leerpotentiescore meten, zijn niet volledig zelf ontworpen, maar zijn gebaseerd op geschikt lesmateriaal voor de leeftijdsgroep. Voor de afname is gebruik gemaakt van dynamisch assessment (Grigorenko & Sternberg, 1998). Er is getoetst volgens een pre-test, interventie en post-test. Eerst is een pre-test afgenomen, waarbij het niveau van het kind in kaart is gebracht. Er is getoetst tot het kind een punt bereikte waarop het interventie nodig had. Op dat punt vond een interventie plaats, totdat het kind

verder kon met de test. Als laatste is er een post-test afgenomen, waarmee kon worden aangetoond of het kind heeft geleerd en een groei heeft doorgemaakt. De meetinstrumenten zijn volgens een standaardprocedure afgenomen, waarbij er een opbouw van moeilijkheidsgraad is gehanteerd. De opbouw van moeilijkheidsgraad moest aanwezig zijn, opdat ieder kind het punt bereikte waar interventie nodig is (Grigorenko & Sternberg, 1998). In dit onderzoek is ieder kind steeds een interventie aangeboden op het moment dat het daadwerkelijk een interventie nodig had. Dit kon dus geen vast moment in de test zijn, omdat dit per kind kan verschillen. Het ene kind had bijvoorbeeld bij de tweede opdracht een interventie nodig en het andere kind bijvoorbeeld pas bij de zevende opdracht. Alleen door dit interventiepunt per kind te bereiken is het mogelijk om vast te stellen of het kind leerpotentieel bezit (Grigorenko & Sternberg, 1998). In de volgende paragrafen zal per meetinstrument een beschrijving worden gegeven.

### *Leerpotentie-testen*

Om het leerpotentieel te meten is gebruik gemaakt van twee meetinstrumenten: 'Camelot' en 'Vergelijken'. In ieder meetinstrument werd een interventie aangeboden. Deze interventies en een verdere beschrijving van de instrumenten worden weergegeven in Bijlage I en II. Bij het meetinstrument 'Camelot' (Bijlage I) werd het ruimtelijk inzicht, oriënterend en lokaliserend vermogen van de leerlingen getest. Zij moesten een pad voor een ridder creëren om bij zijn prinses te komen. Op een houten platform konden kinderen door middel van het neerzetten van torens en trappen een pad maken voor de prins. De prins kon alleen op trappen of op de bovenkant van een muur lopen. Er werd creatief denken en ruimtelijk inzicht van de kinderen gevraagd om tot een oplossing te komen. Het spel heeft een opbouw van moeilijkheidsgraad. Het kind doorliep de levels, welke steeds moeilijker werden. Bij dit meetinstrument is een selectie van levels gemaakt, omdat het niet haalbaar is om in deze test het kind alle levels te laten doorlopen. Bij het selecteren van de levels is gekeken of er iets nieuws in het nieuwe level wordt aangeboden. Was dit niet het geval, dan werd het level eruit gelaten. Op deze manier is de reeks ontwikkeld die voor dit onderzoek is gebruikt.

Bij het meetinstrument 'Vergelijken' (Bijlage II) waren de opdrachten concreet en op de praktijk gericht en werd gebruik gemaakt van materiaal als plastic bekertjes, rietjes, plaatjes en fiches. De test bestond uit acht verschillende metingen. Leerlingen moesten één-op-één relaties kunnen maken aan de hand van voorwerpen die in het dagelijks leven bij elkaar horen, zoals de beker en het rietje. De test bevatte ook één-op-één relaties waarbij de leerlingen plaatjes bij elkaar moesten leggen, zoals een plaatje van een hond en een plaatje van een baasje. Daarnaast was er een

onderdeel waarbij de voorwerpen niets met elkaar te maken hebben, zoals een paperclip en een blokje. Het werd dus minder logisch om de combinaties te maken, maar de leerlingen moesten nog steeds het principe van de één-op-één relatie gebruiken. Zij vergeleken de aantallen door ze één bij één te sorteren. In deze test kwam ook, naast het vergelijken en het herkennen van de één-op-één relaties, het aspect 'meer/minder' aan bod. Daarnaast werd in deze test ook gekeken of kinderen het kardinale aspect binnen het getalbegrip beheersten. Dit betekent dat zij een getal als aanduiding van een aantal kunnen zien en in staat zijn aan te geven van welk aantal er meer of minder zijn. Dit omvat dus het tellend vermogen van de leerlingen.

### *Cito-toetsen*

De scores op de testen zijn vergeleken met vaardigheidsscores van twee onderdelen op de Cito-toets 'Rekenen voor kleuters'. Op school 2 is de Cito-toets 'Rekenen voor kleuters' afgenomen en kon bij deze groep leerlingen een vergelijking gemaakt worden tussen de vaardigheidsscores op de twee onderdelen van deze toets en de leerpotentie-scores van de twee instrumenten. Op school 1 zijn de Cito-toetsen niet afgenomen, waardoor deze vergelijking dus niet met leerlingen van school 1 kon worden gemaakt. Bij de Cito-toetsen zijn de onderdelen 'Meetkunde' en 'Vergelijken' afgenomen. Aan de hand van theorie kan verklaard worden waarom de testen met elkaar vergeleken kunnen worden. De vaardigheidsscore van het onderdeel 'Meetkunde' is vergeleken met de leerpotentie-test 'Camelot'. Beide testen meten het ruimtelijke inzicht en de vaardigheden 'oriënteren en lokaliseren' (Cito, 2010). Daarom is het mogelijk om deze testen met elkaar te vergelijken. De vaardigheidsscore van het onderdeel 'Getalbegrip' is vergeleken met de leerpotentie-test 'Vergelijken'. Beide testen meten het omgaan met getallen en het omgaan met hoeveelheden (Cito, 2010). Deze testen zijn daarom ook met elkaar vergeleken.

### *Leerkrachtbeoordeling*

Naast de Cito-scores zijn de scores op de leerpotentie-testen ook vergeleken met een leerkrachtbeoordeling. Om te meten op welk niveau de leerkracht de leerlingen indeelt, wordt op beide scholen een beoordeling van de leerkracht over de leerlingen afgenomen. De leerkracht is gevraagd om iedere leerling in één van de volgende drie categorieën in te delen: de leerling heeft extra instructie nodig, de leerling heeft het normale aanbod nodig of de leerling heeft verrijking nodig. Als dit van iedere leerling bekend is, kan een vergelijking worden gemaakt tussen de leerpotentie-scores van de testen en de leerkrachtbeoordeling. Deze leerkrachtbeoordeling is op beide scholen afgenomen en vergeleken.

## *Analysemethoden*

### *De variabelen*

De volgende variabelen zijn in dit onderzoek gebruikt: de leerpotentiescores, de vaardigheidsscore die de leerlingen op de Cito-onderdelen 'Meetkunde' en 'Getalbegrip' behalen en de niveaugroep waarin de leerkracht de leerlingen indeelt. De meetinstrumenten zijn gebruikt om het leerpotentieel van de leerlingen in beeld te brengen. Het leerpotentieel is gemeten op het gebied van visueel-ruimtelijk inzicht, vergelijkend en tellend vermogen. De vaardigheidsscores van de (statische) Cito-toetsen op het gebied van rekenen / ruimtelijk inzicht waarmee de kinderen op niveau worden ingedeeld, zijn gebruikt om in beeld te brengen of de score van het leerpotentieel overeenkomt met de statische Cito-toets. Hierdoor kan gekeken worden of er verschil is tussen scores die via een statische of een dynamische test zijn afgenomen. Het niveau waarin de leerling door de leerkracht is ingedeeld wordt vergeleken met de leerpotentie-scores, om te kijken of deze scores overeenkomen. De scores hebben niet dezelfde range, maar door middel van het maken van gelijke groepen is het mogelijk om de variabelen met elkaar te vergelijken (Moore, McCabe & Craig, 2009).

### *De analyses*

Aan de hand van het computerprogramma SPSS 19.0 voor Windows worden correlatietoetsen, gepaarde t-toetsen, Chi-kwadraattoetsen en enkelvoudige variantieanalyses uitgevoerd om te achterhalen of er relaties bestaan tussen de Cito-scores, de leerkrachtbeoordelingen en de leerpotentie-testen. Bij alle analyses wordt uitgegaan van een significantieniveau van  $\alpha = .05$ . Om de hoofdvraag: 'Wat is de relatie tussen de leerpotentie-scores bij kleuters uit groep 1 en het beoordelingssysteem wat scholen hanteren om leerlingen op niveau in te delen?' te kunnen beantwoorden worden correlaties berekend, variantieanalyses uitgevoerd en zal een Chi-kwadratentoets uitgevoerd worden.

Om te achterhalen of er een samenhang is tussen de scores op de verschillende testen wordt een correlatietoets uitgevoerd. De correlatietoetsen worden bij zowel vergelijkingen tussen de testen als vergelijkingen tussen de leerkrachtbeoordeling en de testen uitgevoerd. Bij numerieke variabelen gebeurt dit door middel van de Pearson's correlatietoets en bij categorische variabelen wordt gebruik gemaakt van de Spearman's correlatietoets. Een hoge correlatie betekent dat er veel samenhang is tussen de desbetreffende testen, wat zou betekenen dat een hoge score op de ene

test samengaat met een hoge score op de andere test. Om te achterhalen of de gemiddelden van de scores van de verschillende testen van elkaar verschillen, wordt een t-toets voor gepaarde steekproeven gebruikt. Zo is te zien of de gemiddelden van de scores op de ene test hoger of lager zijn dan op de andere test. Om te achterhalen of er verschil is tussen de groepen (extra instructie nodig, normaal aanbod nodig of verrijking nodig) die door de leerkracht op niveau zijn ingedeeld op de vaardigheidsscore Getalbegrip, de vaardigheidsscore Meetkunde en de leerpotentie-test Camelot wordt een enkelvoudige variantieanalyse uitgevoerd. Hierbij wordt gekeken of de gemiddelden van de groepen op de leerkrachtbeoordeling verschillen van elkaar op de vaardigheidsscore van Getalbegrip, de vaardigheidsscore van Meetkunde en de leerpotentie-score van Camelot. Een verschil zou betekenen dat een groep op de ene test gemiddeld hoger of lager scoort dan op de andere test. Om te kijken of aan de voorwaarden voor de enkelvoudige variantieanalyse wordt voldaan, is de Levene's toets voor gelijkheid van varianties uitgevoerd. Bij deze toets is het belangrijk dat de uitkomsten niet significant zijn, om de variantieanalyse uit te mogen voeren. Indien dit niet zo is, kan er een Kruskal Wallis toets worden uitgevoerd om te achterhalen of er een significant verschil is in de populatiegemiddelden. Als resultaat zal er een p-waarde gerapporteerd moeten worden die kleiner is dan .05. De variantieanalyse is ook uitgevoerd om te achterhalen of er een verschil is in gemiddelden van de leerpotentie-test Vergelijken op de vaardigheidsscore van Getalbegrip. Indien de F-waarde groter is dan 1, is de tussenvariantie groter dan de binnenvariantie. Het merendeel van de totale variantie komt dan voor rekening van verschillen tussen de groepen (tussenvariantie). Hoe groter de F is, des te kleiner de kans dat dit door toeval komt. Als de F-waarde (ongeveer) gelijk is aan 1 zijn binnen- en tussenvariantie (ongeveer) aan elkaar gelijk. De groepen verschillen dan niet echt van elkaar. Bij iedere F wordt een significantieniveau van .05 aangehouden wat bepaalt of de nulhypothese wordt verworpen. Wordt de nulhypothese verworpen dan verschillen de gemiddelden van de subgroepen dus met 95% betrouwbaarheid van elkaar. Om te achterhalen of er een verband bestaat tussen de groepen (extra instructie nodig, normaal aanbod nodig of verrijking nodig) die door de leerkracht op niveau zijn ingedeeld op de leerpotentiescore Vergelijken is een Chi-kwadraattoets uitgevoerd. Er is hierbij gekeken of er een statistisch significant verband bestaat tussen de scores van de groepen op de leerkrachtbeoordeling en de leerpotentiescores Vergelijken. Daarnaast zijn Chi-kwadraattoetsen uitgevoerd om te achterhalen hoe de verhoudingen liggen bij de uitkomsten van verschil in gemiddelden.

## Resultaten

In dit hoofdstuk wordt een beschrijving van de resultaten van de analyses gegeven. In de data-inspectie is gekeken naar de normaliteit van verdelingen, uitbijters en missende waarden. Er is aan alle assumpties voldaan, dus iedere toets kon worden uitgevoerd. Eerst zullen de resultaten van de uitgevoerde toetsen per leerpotentie-test worden beschreven. Daarna worden de resultaten van de uitgevoerde toetsen van de Cito-onderdelen en de Leerkrachtbeoordeling weergegeven.

### *De leerpotentie-testen*

Voor beide leerpotentie-testen is een gepaarde t-toets gebruikt om de gemiddelden van de leerpotentie-testen, de scores op de twee Cito-onderdelen en de Leerkrachtbeoordeling te vergelijken. Alle gepaarde t-toetsen geven een significant verschil aan. Dus bij alle toetsen wordt de nulhypothese verworpen, omdat de gemiddelden significant van elkaar verschillen (Tabel 1). Dit betekent dat de leerpotentie-test Camelot en de vaardigheidsscore Meetkunde, de leerpotentie-test Camelot en de Leerkrachtbeoordeling en de leerpotentie-test Camelot en de leerpotentie-test Vergelijken significant van elkaar verschillen. Dit betekent dat ook de leerpotentie-test Vergelijken en de vaardigheidsscore Getalbegrip en de leerpotentie-test Vergelijken en de Leerkrachtbeoordeling significant van elkaar verschillen.

Tabel 1.

### *Gepaarde t-toets leerpotentie-testen en het beoordelingssysteem*

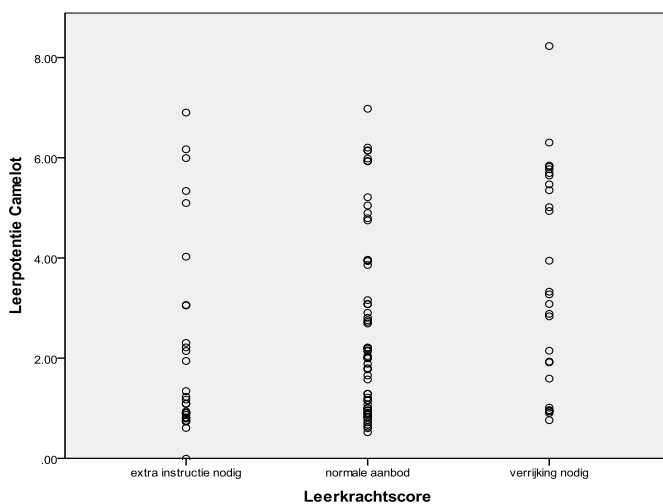
	Gepaarde verschillen				t	df	p
	M	SD	SEM	95% CI			
LP Camelot - Meetkunde	-5.478	3.16	.47	[-6.42, -4.54]	-11.76	45	.00*
LP Camelot - Leerkrachtscore	.737	1.88	.17	[.39, 1.08]	4.26	117	.00*
LP Camelot – LP Vergelijken	.450	2.11	.19	[.07, .83]	2.33	119	.02*
LP Vergelijken – Getalbegrip	-8.59	5.16	.76	[-10.12, 7.05]	-11.28	45	.00*
LP Vergelijken - Leerkrachtscore	.24	.95	.08	[.06, .41]	2.72	117	.01*

\*  $p < .05$ , tweezijdig

### *Leerpotentie-test Camelot*

Om de relatie tussen de leerpotentie-test Camelot en de Leerkrachtbeoordeling te berekenen is gebruik gemaakt van Spearman's correlatietoets. Er is gebruik gemaakt van Spearman's

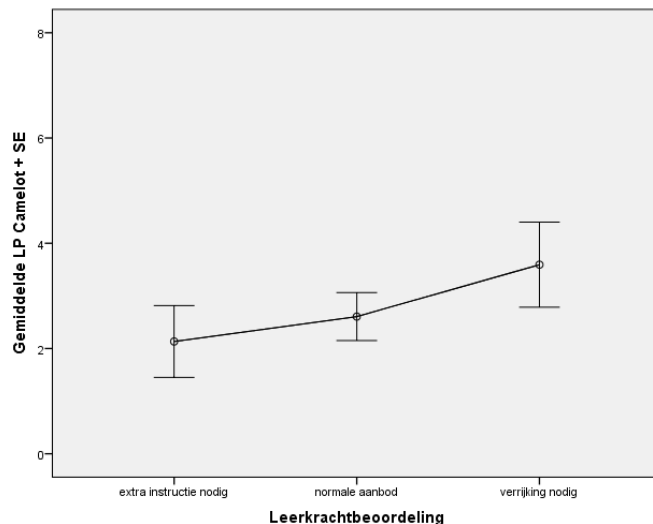
correlatietoets, omdat de variabelen van ordinaal niveau zijn. Er is een zwak positief verband tussen de score op de leerpotentie-test Camelot en de Leerkrachtbeoordeling,  $r(118)=.28$ ,  $p = .002$ . Er wordt 10% of minder van de variantie verklaard. In Figuur 1 is de correlatie van deze combinatie weergegeven. De leerpotentie-score die een leerling op de test Camelot behaalt, komt zeer licht overeen met de score die de leerkracht geeft aan een leerling. Als een leerling een hoge score heeft op de test Camelot, heeft deze ook een hoge score gekregen van de leerkracht. Er zijn wel leerlingen die als 'Extra instructie nodig' worden ingedeeld, maar een hoge score op de leerpotentie-test behalen. Tussen de leerpotentie-test Camelot en de leerpotentie-test Vergelijken is geen sprake van correlatie. Deze twee testen meten binnen het rekenen dus een ander gebied.



*Figuur 1.* Spreidingsdiagram Leerkrachtscore en de leerpotentie-test Camelot.

Omdat via een gepaarde t-toets is achterhaald dat er verschillen zijn tussen de gemiddelden van de leerpotentie-test Camelot en de andere testen, wordt er een variantieanalyse uitgevoerd om te achterhalen of er een verschil is tussen de gemiddelden van de groepen waarin de leerkracht de leerlingen heeft ingedeeld en de score op de leerpotentie-test Camelot. De verschillende niveaugroepen waarin de leerkracht de leerlingen heeft ingedeeld verschillen significant van elkaar,  $F(2,115) = 4.48$ ,  $p = .013$ . Om inzicht te krijgen in wat het verschil tussen de groepen inhoudt is een posthoc Bonferroni-toets uitgevoerd. De groep leerlingen die volgens de leerkracht extra instructie nodig heeft ( $M=2.13$ ,  $SD=1.87$ ), verschilt significant van de groep leerlingen die verrijking nodig heeft ( $M=3.59$ ,  $SD=2.01$ ). In Figuur 2 wordt weergegeven dat de groep die volgens de leerkracht verrijking nodig heeft gemiddeld 1.4 punten hoger scoort op de leerpotentie-test Camelot ( $p = .004$ ).





*Figuur 2.* Lijndiagram gemiddelde scores van de Leerkrachtbeoordeling groepen op de leerpotentie-test Camelot met standaardafwijking.

Na het uitvoeren van de gepaarde t-toets is duidelijk geworden dat er een significant verschil tussen de leerpotentie-test Camelot en de andere testen is. Er is een Chi-kwadraattoets uitgevoerd om te achterhalen hoe de groep leerlingen die laag scoort op het Cito-onderdeel Meetkunde, scoort op de leerpotentie-test Camelot. Om de Chi-kwadraattoets uit te voeren, moeten de percentages van de groepen gelijk zijn. Om te zorgen dat de percentages van de laagst scorende groepen van de twee testen overeenkomen, is gekozen om van de test Camelot de onderste 41% als laagst scorende groep te gebruiken en van Meetkunde de onderste 43.5%. Op de leerpotentie-test Camelot heeft namelijk 41% de score '1' behaald. Een lagere score was niet mogelijk, dus moet deze groep als laagst scorende groep gelden. Het is niet mogelijk om deze groep te splitsen, waardoor deze gehele groep tot laagst scorende groep is benoemd. Leerlingen met dezelfde scores kunnen niet worden gesplitst en in een andere groep worden geplaatst. De groepen moeten bij benadering gelijk zijn, dus voor het Cito-onderdeel Meetkunde is de onderste 43.5% gebruikt. Dit percentage kwam het meest overeen met de 41% van de test Camelot. Deze twee groepen zullen dus gebruikt worden als laagst scorende groep bij deze toets (Moore et al., 2009).

Er behoren 20 leerlingen tot de groep die het laagst scoort op het Cito-onderdeel Meetkunde (43.5%). Van die groep van 20 leerlingen scoren 8 leerlingen (40%) ook op de test Camelot het laagst. De overige 12 leerlingen (60%) behoren niet tot de laagst scorende groep van de test Camelot. Deze 12 leerlingen behoren bij het Cito-onderdeel dus tot de laagst scorende groep, maar laten op de test Camelot een betere score zien. Het andere deel van de groep, 26 leerlingen (66.5%), behoort tot de groep die niet het laagst scoort op het Cito-onderdeel Meetkunde. Van deze groep bevinden 11

leerlingen (42.3%) zich ook in de laagst scorende groep van de test Camelot. En de andere 15 leerlingen (57.7%) behoren bij Camelot ook tot de groep die hoog scoort. De resultaten worden weergegeven in Tabel 2.

Tabel 2.

*Kruistabel met de scores van de leerpotentie-test Camelot en het Cito-onderdeel Meetkunde*

		Camelot 41%		Totaal
		Lage groep	Hoge groep	
	Aantal	8	12	20
Meetkunde 43.5%	Lage groep	% Meetkunde 43.5%	40,0%	60,0%
		% Camelot 41%	42,1%	44,4%
		Aantal	11	15
	Hoge groep	% Meetkunde 43.5%	42,3%	57,7%
		% Camelot 41%	57,9%	55,6%
	Aantal	19	27	46
	Totaal	% Meetkunde 43.5%	41,3%	58,7%
		% Camelot 41%	100,0%	100,0%

Om te achterhalen hoe de leerlingen die op de Leerkrachtbeoordeling laag worden ingedeeld presteren op de leerpotentie-test Camelot, is een Chi-kwadraattoets uitgevoerd. In Tabel 8 (zie bijlage III) zijn de resultaten weergegeven van de niveaugroepen waar de leerkracht de leerling heeft ingedeeld. Uit de resultaten kan worden opgemaakt dat in de groep met de hoogste scores op de test Camelot meer leerlingen zijn vertegenwoordigd die door de leerkracht als ‘normaal aanbod’ of ‘verrijking nodig’ werden beoordeeld. Daarentegen bevinden 6 leerlingen die door de leerkracht als ‘extra instructie nodig’ werden beoordeeld, zich in de bovenste 30% van de scores op de leerpotentie-test Camelot. De bovenste 30% bedraagt 34 leerlingen. Van de zes leerlingen die boven verwachting van de leerkracht scoorden, zijn enkele achtergrondkenmerken bekeken om te kijken of deze mogelijk invloed hebben op deze groep leerlingen. Er is gekeken hoe de achtergrondkenmerken ‘geslacht’ en ‘thuis taal’ verdeeld zijn over deze groep leerlingen.

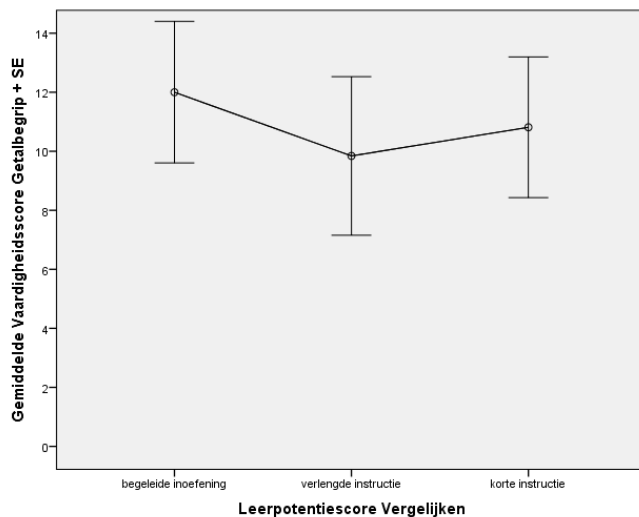
Vijf van de zes leerlingen zijn jongens (83%). In de totale populatie is het percentage jongens 44%. Binnen deze groep van zes leerlingen ligt het percentage ‘jongens’ naar verhouding hoger dan in de totale populatie. De thuis taal is bij drie leerlingen Nederlands (50%), bij twee leerlingen Turks/Nederlands (33%) en bij één leerling is de thuis taal Spaans (17%). In de totale populatie is het percentage thuis taal ‘Nederlands’ 80%, ‘Nederlands gecombineerd met de moedertaal’ 10% en ‘enkel de moedertaal’ 10%. Binnen deze groep van zes leerlingen liggen de percentages ‘Nederlands gecombineerd met de moedertaal’ en ‘alleen de moedertaal’ naar verhouding hoger dan in de totale

populatie. In deze groep zitten meer jongens dan meisjes en is de thuistaal voornamelijk Nederlands of Nederlands gecombineerd met een moedertaal.

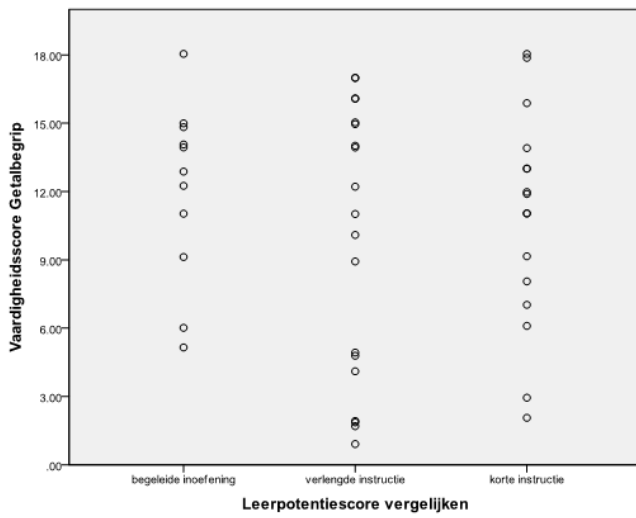
### *Leerpotentie-test Vergelijken*

Om de correlaties te berekenen is gebruik gemaakt van Spearman's correlatietoets. Er is gebruik gemaakt van Spearman's correlatietoets, omdat de variabelen van ordinaal niveau zijn. Er zijn geen significante correlaties gevonden tussen de leerpotentie-test Vergelijken en de leerpotentie-test Camelot ( $r(120) = -.10, p = .28.$ ), Vaardigheidsscore Getalbegrip ( $r(46) = -.07, p = .62$ ) en de Leerkrachtbeoordeling ( $r(118) = .15, p = .10$ ).

Omdat via een gepaarde t-toets is achterhaald dat er verschillen zijn tussen de gemiddelden van de leerpotentie-test Vergelijken en de andere testen, wordt een variantieanalyse uitgevoerd om te achterhalen of er een verschil is tussen de gemiddelden van de niveaugroepen bij de leerpotentie-test Vergelijken op de vaardigheidsscore Getalbegrip. Figuur 3 weergeeft dat de verschillende niveaugroepen van de leerpotentie-test Vergelijken niet significant verschillen van elkaar op het Cito-onderdeel Getalbegrip,  $F(2,43) = 0.63, p = .54$ . Figuur 4 weergeeft dit in een spreidingsdiagram.



*Figuur 3.* Lijndiagram score leerpotentie-test Vergelijken en vaardigheidsscore Getalbegrip.



Figuur 4. Spreidingsdiagram score leerpotentie-test Vergelijken en vaardigheidsscore Getalbegrip.

Om te achterhalen hoe de leerlingen die laag scoren op het Cito-onderdeel Getalbegrip presteren op de leerpotentie-test Vergelijken, is een Chi-kwadraattoets uitgevoerd. Om de Chi-kwadraattoets uit te voeren, moeten de percentages van de groepen gelijk zijn. Om te zorgen dat de percentages van de laagst scorende groepen van de twee testen overeenkomen, is gekozen om de onderste 21,7% als laagst scorende groep op de Cito-toets Getalbegrip te gebruiken. Bij de leerpotentie-test Vergelijken bestaat de onderste groep namelijk uit 19,7%, dit is de groep die begeleide inoefening heeft gekregen. ‘Begeleide inoefening’ was de laagst mogelijke score op de test Vergelijken. Deze groep kan niet worden gesplitst, omdat deze leerlingen allemaal de score ‘begeleide inoefening’ hebben behaald. Leerlingen met dezelfde scoren kunnen niet worden gesplitst en in een andere groep worden geplaatst. De groepen moeten bij benadering gelijk zijn, dus voor het Cito-onderdeel Getalbegrip is de onderste 21.7% gebruikt. Dit percentage kwam het meest overeen met de 19.7% van de test Vergelijken. Deze twee groepen zullen dus gebruikt worden als laagst scorende groep bij deze toets (Moore et al., 2009).

Er behoren 10 leerlingen tot de groep die het laagst scoort op het Cito-onderdeel Getalbegrip (21.7%). Van die groep van 10 leerlingen scoort 1 leerling (10%) ook op de leerpotentie-test Vergelijken het laagst, 7 leerlingen (70%) scoren gemiddeld en 2 leerlingen (20%) scoren het hoogst. De groep ‘hoog’ bestaat dus uit meer leerlingen dan de groep ‘laag’. Het verschil is echter wel 1 leerling. Het andere deel van de groep, 36 leerlingen (78.3%), behoort niet tot de groep die het laagst scoort op het Cito-onderdeel Getalbegrip. Van deze groep scoren 10 leerlingen (27.8%) op de leerpotentie-test Vergelijken het laagst, 12 leerlingen (33.3%) gemiddeld en 14 leerlingen (38.9%) het

hoogst. De groepen zijn dus bijna gelijk, waarbij de lage groep iets kleiner is en de hoogste groep iets groter is. De resultaten worden in Tabel 3 weergegeven.

Tabel 3.

*Kruistabel met de scores van de leerpotentie-test Vergelijken en de vaardigheidsscore Getalbegrip*

		Leerpotentiescore vergelijken			Totaal	
		begeleide inoefening	verlengde instructie	korte instructie		
Getalbegrip 21.7%	Lage groep	Aantal	1	7	2	10
		% Getalbegrip 21.7%	10,0%	70,0%	20,0%	100,0%
	Hoge groep	% LP Vergelijken	9,1%	36,8%	12,5%	21,7%
		Aantal	10	12	14	36
	Totaal	% Getalbegrip 21.7%	27,8%	33,3%	38,9%	100,0%
		% LP Vergelijken	90,9%	63,2%	87,5%	78,3%
	Aantal	11	19	16	46	
	% Getalbegrip 21.7%	23,9%	41,3%	34,8%	100,0%	
	% LP Vergelijken	100,0%	100,0%	100,0%	100,0%	

Om te achterhalen hoe de leerlingen die op de Leerkrachtbeoordeling laag worden ingedeeld presteren op de leerpotentie-test Vergelijken, is een Chi-kwadraattoets uitgevoerd. Via een kruistabel kan gezien worden of er een samenhang is tussen de beoordeling van de leerkracht en de score op de leerpotentie-test Vergelijken. Dit biedt duidelijkheid in hoe die samenhang is verdeeld. De tabel is terug wordt weergegeven in bijlage IV (Tabel 9).

Er behoren 30 leerlingen (25,4%) tot de groep die volgens de leerkracht extra instructie nodig heeft. Van die groep hebben 6 leerlingen (20%) begeleide inoefening, 15 leerlingen (50%) verlengde instructie en 9 leerlingen (30%) korte instructie nodig gehad op de leerpotentie-test Vergelijken. Er hebben meer leerlingen van deze groep korte instructie dan begeleide inoefening gehad. De groep leerlingen die volgens de leerkracht 'extra instructie nodig' hebben en 'korte instructie' heeft gescoord bestaat uit 9 leerlingen. Er zijn dus 9 leerlingen die door de leerkracht laag ingedeeld worden, die aanzienlijk hoger scoren op de leerpotentie-test Vergelijken dan de leerkracht voorspelt. De groep leerlingen die volgens de leerkracht 'extra instructie' nodig heeft en 'begeleide inoefening' heeft gescoord, bestaat uit 6 leerlingen. Er zijn dus 6 leerlingen die door de leerkracht laag ingedeeld worden, die ook daadwerkelijk een lage score behalen op de leerpotentie-test Vergelijken. De groep leerlingen die door de leerkracht laag is ingedeeld en hoger scoort dan de leerkracht heeft voorspeld

is groter dan de groep leerlingen die door de leerkracht laag is ingedeeld en ook laag scoort op de leerpotentie-test. Daarnaast behoren er 61 leerlingen (51.7%) tot de groep die volgens de leerkracht het normale aanbod nodig heeft. Van die groep hebben 13 leerlingen (21.3%) begeleide inoefening, 24 leerlingen (39.3%) verlengde instructie en 24 leerlingen (39.3%) korte instructie nodig gehad op de test Vergelijken. Tot de groep die volgens de leerkracht verrijking nodig heeft behoren 27 leerlingen (22.9%). Van die groep hebben 4 leerlingen (14.8%) begeleide inoefening, 8 leerlingen (29.6%) verlengde instructie en 15 leerlingen (55.6%) korte instructie nodig op de test Vergelijken. Het grootste deel van deze groep heeft de hoogste score gehaald op Vergelijken, dus dit heeft de leerkracht goed ingedeeld.

### *Leerkrachtbeoordeling*

Een gepaarde t-toets is gebruikt om de gemiddelden van Leerkrachtbeoordeling, de leerpotentie-testen en de scores op de twee Cito-onderdelen te vergelijken. Alle gepaarde t-toetsen geven een significant verschil aan. Dus bij alle toetsen wordt de nulhypothese verworpen, omdat de gemiddelden significant van elkaar verschillen (Tabel 4). Dit betekent dat de Leerkrachtbeoordeling en leerpotentie-test Camelot, de Leerkrachtbeoordeling en de vaardigheidsscore Meetkunde en de Leerkrachtbeoordeling en de vaardigheidsscore Getalbegrip significant van elkaar verschillen en dat de scores op de toetsen niet gelijk zijn. De scores van de leerlingen op de ene toets zijn niet gelijk aan de scores van de leerlingen op de andere toets.

Tabel 4.

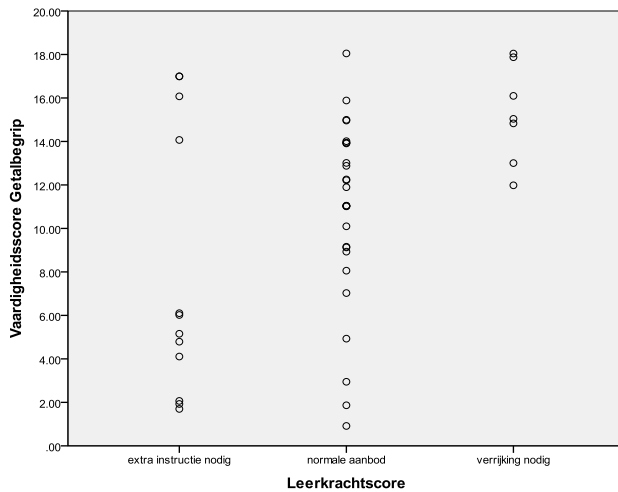
#### *Gepaarde t-toets Leerkrachtbeoordeling*

	Gepaarde verschillen				t	df	p
	M	SD	SEM	95% CI			
Leerkrachtbeoordeling- Camelot	.74	1.88	.17	[.39, 1.08]	4.26	117	.00*
Leerkrachtbeoordeling - Meetkunde	-6.06	2.39	.35	[-6.78, -5.35]	-17.15	45	.00*
Leerkrachtbeoordeling - Getalbegrip	-8.80	4.80	.71	[-10.23, -7.38]	-12.44	45	.00*

\*  $p < .05$ , tweezijdig

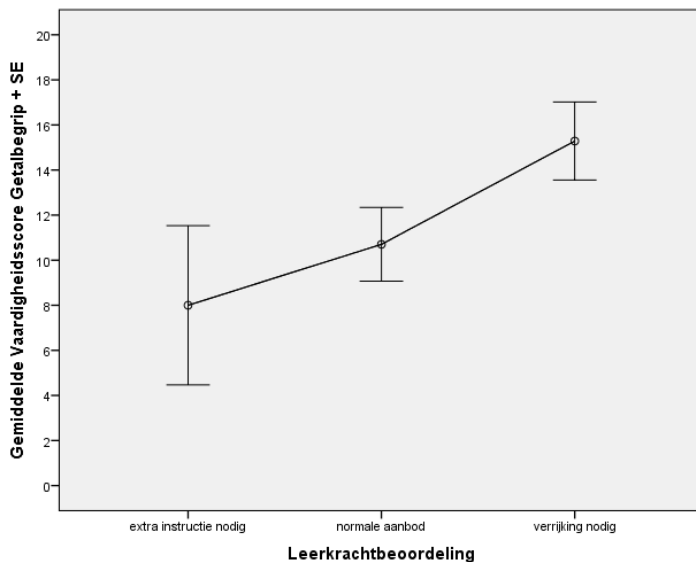
Om te achterhalen of er een verband bestaat tussen de Leerkrachtbeoordeling en de vaardigheidsscore Getalbegrip, is Spearman's correlatietoets uitgevoerd. Er is gebruik gemaakt van Spearman's correlatietoets, omdat de variabelen van ordinaal niveau zijn. Er is een matig sterk positief verband tussen de Leerkrachtbeoordeling en de vaardigheidsscore Getalbegrip,  $r(46) = .39$ ,  $p$

=.006. Bij een correlatiecoëfficiënt van  $r = .30$ , wordt er 10 – 25 % van de variantie verklaard. In Figuur 5 is de correlatie van deze combinatie weergegeven. De vaardigheidsscore die een leerling op het Cito-onderdeel Getalbegrip haalt, komt licht overeen met de score die de leerkracht geeft aan een leerling. Als een leerling een hoge score heeft op het Cito-onderdeel Getalbegrip, heeft deze ook een hoge score gekregen van de leerkracht.



Figuur 5. Spreidingsdiagram Leerkrachtbeoordeling en vaardigheidsscore Getalbegrip.

Omdat via een gepaarde t-toets is achterhaald dat er verschillen zijn tussen de gemiddelden van Leerkrachtbeoordeling en de andere testen, wordt een variantieanalyse uitgevoerd om te achterhalen of er een verschil is tussen de gemiddelden van de groepen. Om te achterhalen hoe de verschillende groepen die door de leerkracht zijn ingedeeld presteren op het Cito-onderdeel Getalbegrip is een variantieanalyse uitgevoerd. De verschillende niveaugroepen waarin de leerkracht de leerlingen heeft ingedeeld verschillen significant van elkaar,  $F(2,43) = 5.53$ ,  $p = .007$ . Om inzicht te krijgen in het verschil tussen de groepen is een Bonferroni-toets uitgevoerd. De groep leerlingen die volgens de leerkracht extra instructie nodig heeft ( $M = 8.00$ ,  $SD = 6.12$ ), verschilt significant van de groep leerlingen die verrijking nodig heeft ( $M = 15.29$ ,  $SD = 2.29$ ). Zoals in Figuur 6 te zien is scoort de groep die volgens de leerkracht verrijking nodig heeft gemiddeld 7.3 punten hoger op het Cito-onderdeel Getalbegrip ( $p = .002$ ).



Figuur 6. Lijndiagram Leerkrachtbeoordeling en vaardigheidsscore Getalbegrip.

Om te achterhalen hoe de leerlingen die laag worden ingedeeld door de leerkracht scoren op het Cito-onderdeel Getalbegrip, is een Chi-kwadraattoets uitgevoerd. Om te zorgen dat de percentages van de laagst scorende groep op Getalbegrip en de groep ‘Extra instructie nodig’ bij de Leerkrachtbeoordeling zo goed mogelijk overeenkomen, is gekozen om de onderste 21.7% als laagst scorende groep op het Cito-onderdeel Getalbegrip te gebruiken. Bij de Leerkrachtbeoordeling is de groep ‘Extra instructie nodig’ 25%. Dit komt niet precies overeen, maar deze groepen liggen het dichtst bij elkaar.

Er behoren 10 leerlingen (21.7%) tot de groep die het laagst scoort op het Cito-onderdeel Getalbegrip. Van die groep van 10 leerlingen worden 6 leerlingen (60%) door de leerkracht beoordeeld als ‘Extra Instructie nodig’, 4 leerlingen (40%) worden door de leerkracht beoordeeld als ‘Normale aanbod nodig’ en 0 leerlingen (0%) van die groep worden door de leerkracht beoordeeld als ‘Verrijking nodig’. Het andere deel van de groep bestaat uit 36 leerlingen (78.3%) en behoort niet tot de groep die het laagst scoort op het Cito-onderdeel Getalbegrip. Van deze groep worden 6 leerlingen (16.7%) door de leerkracht ingedeeld als ‘Extra instructie nodig’, 23 leerlingen (63.9%) als ‘Normale aanbod’ en 7 leerlingen (19.4%) ingedeeld als ‘Verrijking nodig’. Zie Tabel 5 voor de resultaten. De groep leerlingen ‘Extra instructie nodig’ was evenredig verdeeld over de lage en de hoge scores van het Cito-onderdeel. Er waren net zoveel leerlingen uit de groep die hoog scoorde op het Cito-onderdeel Getalbegrip als de groep die laag scoorde op het Cito-onderdeel Getalbegrip. Er is gekeken hoe het achtergrondkenmerk ‘thuisstaal’ verdeeld is in de groep die volgens de leerkracht



extra instructie nodig heeft. Hieruit blijkt dat 55.5% thuis Nederlands spreekt, 27% thuis Nederlands gecombineerd met een andere taal spreekt en 7.5% thuis alleen een andere taal spreekt.

Tabel 5.

*Kruistabel met de scores van het Cito-onderdeel Getalbegrip en de Leerkrachtbeoordeling*

		Leerkrachtbeoordeling			Totaal	
		extra instructie	normale aanbod	verrijking nodig		
Getalbegrip 21.7%	Lage groep	Aantal	6	4	0	10
		% Getalbegrip 21.7%	60,0%	40,0%	0,0%	100,0%
		% Leerkrachtbeoordeling	50,0%	14,8%	0,0%	21,7%
	Hoge groep	Aantal	6	23	7	36
		% Getalbegrip 21.7%	16,7%	63,9%	19,4%	100,0%
		% Leerkrachtbeoordeling	50,0%	85,2%	100,0%	78,3%
Totaal	Aantal	12	27	7	46	
	% Getalbegrip 21.7%	26,1%	58,7%	15,2%	100,0%	
	% Leerkrachtbeoordeling	100,0%	100,0%	100,0%	100,0%	

Om te achterhalen hoe de verschillende groepen die door de leerkracht zijn ingedeeld presteren op het Cito-onderdeel Meetkunde is een variantieanalyse uitgevoerd. De verschillende niveaugroepen waarin de leerkracht de leerlingen heeft ingedeeld verschillen niet significant van elkaar,  $F(2,115) = 1.45$ ,  $p = .25$ . Er is een Chi-kwadraattoets uitgevoerd om inzicht te krijgen in hoe de leerlingen die laag worden ingedeeld door de leerkracht presteren op het Cito-onderdeel Meetkunde. Om te zorgen dat de percentages van de laagst scorende groep op het Cito-onderdeel Meetkunde en de groep 'Extra instructie nodig' bij de Leerkrachtbeoordeling zo goed mogelijk overeenkomen, is gekozen om de onderste 23.9% als laagst scorende groep op het Cito-onderdeel Meetkunde te gebruiken. Bij de Leerkrachtbeoordeling is de groep 'Extra instructie nodig' 25%. Dit komt niet precies overeen, maar deze groepen liggen het dichtst bij elkaar. Er behoren 11 leerlingen (23.9%) tot de groep die het laagst scoort op het Cito-onderdeel Meetkunde. Van die groep van 11 leerlingen worden 6 leerlingen (54.5%) door de leerkracht beoordeeld als 'Extra Instructie nodig', 5 leerlingen (45.5%) worden door de leerkracht beoordeeld als 'Normale aanbod nodig' en 0 leerlingen (0%) van die groep worden door de leerkracht beoordeeld als 'Verrijking nodig'. Het andere deel van de groep bestaat uit 35 leerlingen (76.1%) en behoort niet tot de groep die het laagst scoort op het Cito-onderdeel Meetkunde. Van deze groep worden 6 leerlingen (17.1%) door de leerkracht beoordeeld als 'Extra instructie nodig', 22 leerlingen (62.9%) als 'Normale aanbod' en 7 leerlingen (20.0%) beoordeeld als 'Verrijking nodig'. De categorie 'Extra instructie nodig' bevat bijna net zoveel

leerlingen uit de groep die hoog scoorde op het Cito-onderdeel als de groep die laag scoorde op het Cito-onderdeel Meetkunde. Er zit één leerling verschil in. De resultaten worden in Tabel 6 weergegeven.

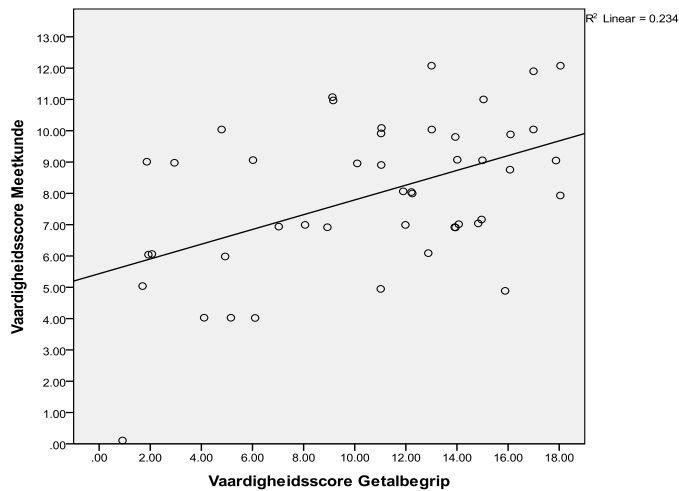
Tabel 6.

*Kruistabel met de scores van het Cito-onderdeel Meetkunde en de Leerkrachtbeoordeling*

		Leerkrachtbeoordeling			Totaal
		extra instructie	normale aanbod	verrijking nodig	
Meetkunde 23.9%	Aantal	6	5	0	11
	Lage groep % Meetkunde 23.9%	54,5%	45,5%	0,0%	100,0%
	% Leerkrachtbeoordeling	50,0%	18,5%	0,0%	23,9%
	Aantal	6	22	7	35
	Hoge groep % Meetkunde 23.9%	17,1%	62,9%	20,0%	100,0%
	% Leerkrachtbeoordeling	50,0%	81,5%	100,0%	76,1%
Totaal	Aantal	12	27	7	46
	% Meetkunde 23.9%	26,1%	58,7%	15,2%	100,0%
	% Leerkrachtbeoordeling	100,0%	100,0%	100,0%	100,0%

#### *De Cito-onderdelen Meetkunde & Getalbegrip*

De Cito-onderdelen 'Meetkunde' en 'Getalbegrip' zijn met elkaar vergeleken. Hiervoor is Pearson's correlatietoets gebruikt. Er is gebruikt gemaakt van Pearson's correlatietoets, omdat er sprake is van numerieke variabelen. Er is een matig sterk positief verband tussen de vaardigheidsscore Meetkunde en de vaardigheidsscore Getalbegrip,  $r(46) = .48$ ,  $p = .001$ . Er wordt 10-25% van de variantie verklaard. In Figuur 7 is de correlatie weergegeven. De vaardigheidsscore die een leerling op het Cito-onderdeel Meetkunde haalt, komt licht overeen met de vaardigheidsscore op het Cito-onderdeel Getalbegrip. Als een leerling een hoge score heeft op het Cito-onderdeel Meetkunde, heeft deze ook een hoge score op het Cito-onderdeel Getalbegrip.



Figuur 7. Spreidingsdiagram vaardigheidsscore Meetkunde en vaardigheidsscore Getalbegrip.

### Opvallende leerlingen

Om te kijken of er leerlingen zijn die tot ‘onderpresteerders’ benoemd zouden kunnen worden, is er gekeken naar welke leerlingen laag scoorden op de Cito-onderdelen en de Leerkrachtbeoordeling. De tabel waarin dit wordt weergegeven is terug te vinden in Bijlage V (Tabel 10). Daarnaast zijn er leerlingen die hogere scores behaalden op de leerpotentie-testen dan dat hun Cito-scores of Leerkrachtbeoordeling deed vermoeden. Tabel 7 geeft deze leerlingen weer. Zij scoren dus laag op het huidige beoordelingsstelsel en hoog op één van de of beide leerpotentie-testen. Er zijn in totaal 22 leerlingen die bij één of meerdere statische metingen bij de laag scorende groep behoren en bij één of meerdere leerpotentie-testen bij de hoog scorende groep behoren.

Tabel 7.

*Opvallende leerlingen*

<b>Combinatie</b>	<b>Leerlingen</b>	<b>Totaal</b>
Lage vaardigheidsscore Meetkunde (onderste 24%), lage vaardigheidsscore Getalbegrip (onderste 26%), extra instructie nodig volgens leerkracht + hoge LP Camelot (bovenste 23%)	J11, K6	2
Lage vaardigheidsscore Meetkunde (onderste 24%), lage vaardigheidsscore Getalbegrip (onderste 26%), extra instructie nodig volgens leerkracht + hoge LP Vergelijken (bovenste 40%)	G3, G4	2
Extra instructie nodig volgens leerkracht + hoge LP Camelot	D8, F6, H13	3
Extra instructie nodig volgens leerkracht + hoge LP Vergelijken	C9, D13, E7, E10, F7, F9 en G7.	7
Lage vaardigheidsscore + hoge LP score	F10 (gb), G10 (gb), H14 (gb), I3 (mk), K10 (mk), K11 (mk/gb)	6
Extra instructie nodig volgens leerkracht + lage vaardigheidsscore Meetkunde (onderste 43,5%)	I4	1
Extra instructie nodig volgens leerkracht + lage vaardigheidsscore Getalbegrip (onderste 26%)	J1	1

**Discussie**

Het doel van deze studie was verschillen in kaart te brengen tussen de scores op de gebruikte leerpotentie-testen en het beoordelingssysteem wat in het basisonderwijs wordt gehanteerd bij leerlingen uit groep 1. Middels vergelijkingen tussen leerpotentie-testen en Cito-scores en leerkrachtenbeoordelingen is achterhaald of er leerlingen zijn die een hoger niveau behalen op de leerpotentie-testen dan het niveau waarop zij nu worden ingedeeld aan de hand van de Cito-testen en de leerkrachtbeoordeling. Het meten van het leerpotentieel is gedaan aan de hand van zelf samengestelde testen die middels dynamisch assessment zijn afgenomen. Er zijn een aantal overeenkomsten gevonden tussen de leerpotentie-testen die in dit onderzoek zijn gebruikt en het beoordelingssysteem dat gehanteerd wordt in de kleutergroep, maar er is geen sprake van een significante relatie tussen deze testen en het beoordelingssysteem. In dit onderzoek is gekeken naar leerlingen die als 'onderpresteerders' gezien kunnen worden en in aanmerking zouden moeten komen voor leerstof op een hoger niveau. Er is een kleine groep leerlingen in beeld gebracht die via

één van de Cito-onderdelen of de leerkrachtbeoordeling op een laag niveau wordt geschat, maar op één van de leerpotentie-testen op een hoger niveau presteren. De leerlingen die zich op basis van het huidige beoordelingssysteem in de laagste groep bevinden, bevinden zich niet altijd in de laagste groep op de leerpotentie-testen. Deze leerlingen hebben door middel van de leerpotentie-testen laten zien dat zij in korte tijd leerstappen kunnen maken. Er moet wel rekening gehouden worden met het feit dat deze leerpotentie-testen geen valide testen zijn, dus het is niet bewezen dat van deze leerlingen ook daadwerkelijk het leerpotentieel is gemeten.

#### *Leerpotentie-testen en de leerkrachtbeoordeling*

In deze paragraaf wordt antwoord gegeven op de vraag wat het verschil is tussen de scores op de leerpotentie-testen en de beoordeling van de leerkracht. Tussen de scores op de leerpotentie-test Camelot en de leerkrachtbeoordeling is een significante relatie waargenomen, al is deze klein. Wanneer een leerling een hoge score behaalde op de leerpotentie-test Camelot, werd deze ook op een hoog niveau ingeschat door de leerkracht. Tussen de leerpotentie-test Vergelijken en de leerkrachtbeoordeling bestaat geen verband. Bij de leerpotentie-test Camelot scoort de groep leerlingen die volgens de leerkracht verrijking nodig heeft, gemiddeld 1.4 punten hoger op de leerpotentie-test Camelot dan de groep leerlingen die volgens de leerkracht extra instructie nodig heeft. Dit betekent dat het niveau waarop de leerkracht de leerlingen indeelt overeenkomt met het niveau dat de leerlingen op de leerpotentie-test Camelot behalen. De groep leerlingen die op de test Camelot hoog scoort, is voornamelijk vertegenwoordigd door leerlingen die door de leerkracht als 'Normaal aanbod' of 'Verrijking nodig' zijn beoordeeld. Er moet wel rekening gehouden worden dat de leerpotentie-test Camelot geen valide test is, dus geen bewijs dat de leerpotentie-test ook daadwerkelijk leerpotentieel meet. De uitkomsten moeten daarom met enige voorzichtigheid worden geïnterpreteerd. Van de totale groep weet de leerkracht het niveau redelijk goed te voorspellen. Ondanks deze correlatie zijn er een aantal leerlingen die hoog scoren op de leerpotentie-test, terwijl de leerkracht hen in de groep 'extra instructie nodig' heeft ingedeeld. Er bevinden zich zes leerlingen in de bovenste 30% van de scores op de leerpotentie-test Camelot, die door de leerkracht als 'Extra instructie nodig' zijn beoordeeld. Deze leerlingen worden dus door de leerkracht lager beoordeeld dan hun prestatie op de leerpotentie-test aangeeft. Deze groep leerlingen is 5% van de totale populatie, maar als het maximale leerniveau uit iedere leerling gehaald moet worden, zouden er geen onderpresteerders zijn en zou dit percentage '0' moeten zijn. Van de totale populatie is 5% een klein gedeelte. Er moet ook realistisch worden gekeken naar de kans dat deze 5% een foutenmarge voor de leerkracht bij het inschatten op niveau kan zijn. Op de leerpotentie-test Vergelijken behaalt 30% van de groep leerlingen die door de leerkracht als 'extra

instructie nodig' wordt ingedeeld (N=30), een hoge score. Er zijn dus negen leerlingen die door de leerkracht laag worden ingeschat en een hoge score behalen op de leerpotentie-test. Er zijn zes leerlingen van wie de leerkracht ze correct heeft ingedeeld in de laagste groep. In verhouding is de groep die het hoogste niveau behaalt op de leerpotentie-test terwijl de leerkracht ze als 'extra instructie nodig' had ingedeeld, groter dan de groep die aan de inschatting van de leerkracht voldeed en het laagste niveau behaalde op de leerpotentie-test. Dit betekent dat leerkrachten de negen leerlingen die boven verwachting hebben gescoord niet snel in verrijgingsprogramma's meenemen, omdat ze denken dat deze leerlingen juist extra instructie nodig hebben. Terwijl de leerlingen bij het dynamisch assessment laten zien dat ze met weinig instructie leerstappen kunnen maken. De verwachting in dit onderzoek was dat leerkrachten hogere verwachtingen hebben van een leerling die goede resultaten behaalt, dan een leerling die resultaten beneden het gemiddelde behaalt (de Boer et al., 2010) en dat de leerkracht niet bij alle leerlingen een juiste verwachting zou hebben. De leerkracht zou dus geen hoge verwachtingen hebben van leerlingen die op de Cito-toetsen resultaten beneden het gemiddelde behalen. Er werd verwacht dat er leerlingen zouden zijn die door de leerkracht laag werden ingeschat, maar hoge scores zouden behalen op de leerpotentie-test. Dit onderzoek laat op de leerpotentie-test Camelot en op de leerpotentie-test Vergelijken een kleine groep leerlingen zien waarbij hier sprake van is.

Het feit dat er toch leerlingen zijn die hoger scoren dan de leerkracht verwacht, komt overeen met voorgaand onderzoek op dit gebied (Keuning & Visser, 2013). De intentie van de leerkracht zal altijd zijn om een leerling op basis van zijn vaardigheid in te schatten. Deze inschatting zal er tevens voor zorgen dat het niveau van de leerling ook niet hoger zal komen dan deze inschatting. Keuning & Visser (2013) spreken hierbij van een *self fulfilling prophecy*. Dit houdt in dat als een leerkracht niet veel van een bepaalde vaardigheid van een leerling verwacht, het aanbod voor deze leerling ook beperkt zal zijn. Op deze manier zal een leerling niet de kans krijgen om hoger te presteren dan de verwachting die de leerkracht van hem heeft. Het is mogelijk dat dit het geval is bij de groep leerlingen uit dit onderzoek naar leerpotentieel en dat deze leerlingen op de leerpotentie-testen hun daadwerkelijke niveau hebben kunnen laten zien. Door de leerkracht worden deze leerlingen ingedeeld in de groep die extra instructie nodig heeft, terwijl zij op basis van hun leerpotentieel wellicht juist meer gebaat zijn bij aanbod op een hoger niveau (Keuning & Visser, 2013). Onderzoek laat zien dat leerkrachten het niveau van leerlingen baseren op de Cito-resultaten (van Batenburg, 2012). Daarnaast baseren zij hun oordeel op eigen observaties, maar dit doen zij in mindere mate. Een leerkracht kijkt dus niet gericht naar de capaciteiten van een leerling, maar voornamelijk naar de feitelijke scores die hij of zij heeft behaald op de Cito-toetsen en vult dit aan

met de eigen observaties in de klas (van Batenburg, 2012). Het onderzoek van Van Batenburg geeft aan dat leerkrachten er vaak naast zitten (van Batenburg, 2012). Dit gebeurt dus wanneer de leerkracht hun beoordeling baseren op basis van Cito resultaten en hun eigen observaties. De leerlingen die in dit onderzoek boven verwachting van de leerkracht scoorden, zouden tot onderpresteerders kunnen worden benoemd (Doolaard & Oudbier, 2010). Dit komt overeen met de vorm van onderpresteren waarbij discrepantie tussen de verwachtingen van de leerkracht en de feitelijke prestaties van de leerlingen bestaat (Mulder et al., 2007). De verwachtingen van de leerkrachten zijn te laag voor deze leerlingen. Onderzoek van Mulder et al. (2007) wijst uit dat leerkrachten van groep 4 bij 13% van de leerlingen te lage verwachtingen had over het vervolg van de schoolcarrière van de leerlingen. In het huidige onderzoek is een lager percentage gemeten. Te lage verwachtingen van leerkrachten, betekent dat de desbetreffende leerlingen eerder naar deze verwachtingen zullen presteren (Jackson & MacKergow, 2003). De leerlingen in dit onderzoek lopen dus ook het risico om in het vervolg van hun schoolcarrière naar de lage verwachtingen van leerkrachten te gaan presteren. Het kan voor leerkrachten lastig zijn om leerlingen met een hoog leerniveau te signaleren, omdat niet alle leerlingen met veel potentie op basis van hoge schoolprestaties opvallen (Mooij, Hoogeveen, Driessen, Van Hell & Verhoeven, 2007). Als bij het indelen op niveau gebruik gemaakt zou worden van leerpotentie-testen kan er gekeken worden naar de groei die een leerling doormaakt (Pfeiffer, 2011). Op deze manier hebben de leerlingen die door de leerkracht op een niveau worden ingeschat dat beneden hun leerpotentieel is, meer kans om hun daadwerkelijke niveau te laten zien.

Van de zes leerlingen die boven verwachting van de leerkracht scoorden, zijn enkele achtergrondkenmerken bekeken. Vijf van de zes leerlingen zijn jongens en de thuistaal is bij drie leerlingen Nederlands, bij twee leerlingen Turks/Nederlands en bij één leerling is de thuistaal Spaans. Deze groep is te klein om te kunnen generaliseren, maar er bevinden zich meer jongens dan meisjes in de groep. Dat er meer jongens bij deze groep horen, kan komen vanwege het feit dat leerkrachten doorgaans meisjes op een hoger niveau inschatten dan jongens (van den Bergh, 1999). De leerkrachten hebben wellicht minder verwachtingen van de jongens en hebben zij daarom boven verwachting gescoord. Daarnaast is het bij kleuters zo dat de jongens het beste presteren op het gebied van rekenen (Penner & Paret, 2008). De jongens laten wellicht eerder een hoge score zien dan de meiden, omdat zij gemakkelijker een hogere score halen bij rekentoetsen.

### *Leerpotentie-testen en Cito-onderdelen*

In deze paragraaf wordt antwoord gegeven op de vraag wat het verschil is tussen de scores op de leerpotentie-testen en de Cito-onderdelen. Bij de vergelijking tussen de leerpotentie-test Camelot en het Cito-onderdeel Meetkunde is gebleken dat van de groep leerlingen die op het Cito-onderdeel Meetkunde een lage score had, 60% niet tot de laag scorende groep op de test Camelot behoort. Er zijn dus leerlingen die op basis van hun score op het Cito-onderdeel tot de laagste groep behoren, die op de leerpotentie-test een hoge score laten zien. Dit betekent dat deze leerlingen aan de hand van hun Cito-scores in het beoordelingssysteem nooit zullen worden ingedeeld in een groep waarin leerlingen verrijgingsstof aangeboden krijgen, omdat op basis van hun scores op de Cito-toetsen verwacht wordt dat ze geen uitdaging aankunnen. Bij de vergelijking tussen de leerpotentie-test Vergelijken en het Cito-onderdeel Getalbegrip is gebleken dat de verschillende niveaugroepen van de leerpotentie-test Vergelijken niet significant van elkaar verschillen op het Cito-onderdeel Getalbegrip. De verschillen in de niveaugroepen op de leerpotentie-test Vergelijken bestaan uit leerlingen die korte instructie, verlengde instructie of begeleide inoefening nodig hadden. Dat er geen significant verschil is betekent dat leerlingen die bijvoorbeeld een hoge score hebben op de test Vergelijken, zowel een lage als een hoge vaardigheidsscore kunnen behalen op het Cito-onderdeel Getalbegrip. Leerlingen die een lage score op het Cito-onderdeel halen, kunnen net zo hoog scoren op de leerpotentie-test Vergelijken als leerlingen die een hoge score op het Cito-onderdeel hebben. Dit betekent dat ofwel de leerpotentie-test Vergelijken geen goede weergave is van het niveau van de leerlingen of dat de leerlingen op de leerpotentie-test op een ander niveau presteren dan op het Cito-onderdeel. De leerpotentie-test is niet getest op validiteit, waardoor niet met zekerheid kan worden gezegd dat de test ook daadwerkelijk het leerpotentieel van de leerlingen meet. Hierdoor is het aannemelijk om te stellen dat de leerpotentie-test niet het juiste niveau van de leerlingen meet en dat de scores daarom ver uit elkaar liggen. De Cito-toets 'Rekenen voor kleuters' is overigens wel getest op validiteit en is hierop positief beoordeeld (Koerhuis & Keuning, 2011). In dit onderzoek is vooral gekeken naar leerlingen die op basis van het huidige beoordelingssysteem op een laag niveau worden ingedeeld. Wanneer de resultaten van de groep leerlingen die op het Cito-onderdeel Getalbegrip het laagst scoort worden bekeken, behaalt 90% van die groep een gemiddelde of hoge score op de leerpotentie-test. Deze groep is 19% van de totale groep leerlingen van de school waarvan de Cito-resultaten bekend zijn. Het is opvallend dat maar één leerling op beide testen bij de laagste groep zit. De andere leerlingen uit de laagst scorende groep op het Cito-onderdeel Getalbegrip bewijzen meer leerpotentieel te bezitten, dan hun vaardigheidsscore op het Cito-onderdeel Getalbegrip doet vermoeden. Zij behoren bij het Cito-onderdeel tot de laagst scorende



groep en bij de leerpotentie-test behoren zij niet tot de laagst scorende groep. Op beide leerpotentie-testen zijn er leerlingen die tot onderpresteerders benoemd kunnen worden. De manier van onderpresteren waarbij er een discrepantie is tussen het cognitief potentieel en schoolprestaties is op deze leerlingen van toepassing (Mulder et al., 2007). De schoolprestaties geven een lager niveau aan dan het de leerpotentie-testen. Onderzoek van de Onderwijsraad (2007) heeft uitgewezen dat de omvang van het percentage onderpresterende leerlingen varieert van 10 tot 18% van de leerling groep. Het percentage leerlingen wat onderpresteert, gebaseerd op de discrepantie tussen de leerpotentie-test Vergelijken en het Cito-onderdeel Getalbegrip, is 19%. Op dit onderdeel komt het percentage dus overeen met het percentage wat de Onderwijsraad heeft ontdekt.

De Cito-toetsen zijn statische toetsen en zoals in voorgaand onderzoek is aangetoond, zitten er bepaalde nadelen aan de statische manier van toetsen (Zaal, 2012; Borland, 2009; Doolaard & Oudbier, 2010; Resing, 2006). Bij statische toetsen wordt geen rekening gehouden met de mate waarin de leerling zich heeft ontwikkeld, het is een momentopname en het levert geen beeld op van onderpresterende leerlingen (Zaal, 2012; Borland, 2009; Doolaard & Oudbier, 2010; Resing, 2006). Intelligentietesten zijn vaak ook statische testen en bij dit soort toetsen is er een beperkte relatie tussen de scores en de educatieve praktijk (Reschly, 1997) en de aandacht wordt meer gericht op het product dan op de psychologische processen (Wagner & Sternberg, 1984). De uitkomsten op de leerpotentie-testen laten zien dat de dynamische manier van testen bij sommige leerlingen een andere score oplevert. De ontwikkelingspsycholoog Piaget (1972) vindt het bij testen van belang dat kennismeting als uitgangspunt wordt gebruikt. Het draait volgens Piaget (1972) om het meten van 'wat een kind nu kan'. Dit hoort te gebeuren volgens gestandaardiseerde testafnames en op statische wijze (Piaget, 1972). In dit onderzoek is er middels dynamisch assessment getoetst en is vooral gekeken naar 'wat een kind kan leren' in plaats van 'wat een kind nu kan'. Er is geen significante relatie gemeten tussen de twee leerpotentie-testen, dus de twee testen laten geen verband zien. De twee leerpotentie-testen toetsten ook een ander deelgebied binnen rekenen. Dit zou een verklaring kunnen zijn voor het feit dat de twee testen niet dezelfde scores opleverden. Er zijn wel leerlingen die op de leerpotentie-testen een hoger niveau bereiken dan dat zij op de Cito-onderdelen en/of de leerkrachtbeoordeling doen. Voorgaand onderzoek geeft aan dat het onverstandig is om het onderwijsaanbod al vroeg te baseren op leerprestaties (Keuning & Visser, 2013). Is het relevant om te kijken naar leerprestaties (zoals resultaten op de Cito-toets) bij het selecteren van een ander onderwijsaanbod voor een kind? Het leerrendement kan niet met zekerheid worden vastgesteld op basis van een bepaald startniveau (Keuning & Visser, 2013). Er moet dus worden nagedacht of de statische manier van testen, waarvan gebruikt gemaakt wordt bij Cito-toetsen, een geschikte manier

is om het niveau van leerlingen te meten (Doolaard & Oudbier, 2010; Mulder et al., 2007). Bij de Cito-toetsen wordt getest wat een kind op een bepaald moment kan en niet in hoeverre het kind groeit (van Batenburg, 2012). Bij leerpotentie-testen kan er gekeken worden naar de groei die een leerling doormaakt (Pfeiffer, 2011). Op deze manier hebben de leerlingen die op basis van hun Cito-scores op een niveau worden ingedeeld dat beneden hun leerpotentieel is, meer kans om hun daadwerkelijke niveau te laten zien. Er kan daarom door schoolbesturen en leerkrachten een afweging worden gemaakt of de resultaten van leerpotentie-testen een basis kunnen vormen om leerlingen op niveau in te delen (van Batenburg, 2012). Mogelijk zouden een omvangrijk instituut als Cito een afweging kunnen maken om leerpotentie-testen te gaan ontwikkelen, zodat er meer mogelijkheden voor scholen zijn om leerpotentie-testen in te zetten.

#### *Leerkrachtbeoordeling en Cito-onderdelen*

Om te achterhalen wat het verschil is tussen de beoordeling van de leerkracht en de vaardigheidsscores op de twee Cito-onderdelen, Getalbegrip en Meetkunde, is een correlatietoets uitgevoerd en daar is uitgekomen dat er een significante positieve relatie is tussen de leerkrachtbeoordeling en de scores op het Cito-onderdeel Getalbegrip. Dit betekent dat als een leerling een hoge score op het Cito-onderdeel Getalbegrip heeft, deze leerling ook een hoge score van de leerkracht heeft gekregen. Dit betekent dat de manier waarop de leerkracht de leerlingen indeelt, overeen komt met de score die zij op het Cito-onderdeel Getalbegrip halen. De groep die volgens de leerkracht verrijking nodig heeft, scoort gemiddeld 7.3 punten hoger op de test Getalbegrip dan de groep die volgens de leerkracht extra instructie nodig heeft. Maar er is ook gekeken hoe de leerlingen uit de groep 'Extra instructie nodig' zich verhouden tot de test Getalbegrip. De groep die door de leerkracht laag wordt ingeschat, is evenredig verdeeld over de hoog scorende en laag scorende groep bij de test Getalbegrip. Dit duidt dus aan dat de leerkrachtbeoordelingen van de lage groep voor de helft overeenkomen met de scores op het Cito-onderdeel Getalbegrip. Zoals eerder beschreven laat voorgaand onderzoek zien dat leerkrachten het niveau van leerlingen baseren op de Cito-resultaten, aangevuld met hun eigen observaties (Van Batenburg, 2012). Het is niet bekend waar de leerkrachten uit dit onderzoek hun beoordelingen op baseren. Hen is gevraagd om de leerlingen op één van de drie (instructie)niveaus in te delen. In dit onderzoek is er sprake van een significante positieve relatie tussen de leerkrachtbeoordeling en de scores op het Cito-onderdeel Getalbegrip. Bij de groep 'Extra instructie nodig' komt echter maar de helft van de scores overeen met de scores op het Cito-onderdeel Getalbegrip. Dit kan betekenen dat de leerkrachtbeoordeling niet precies overeenkomt met dit specifieke Cito-onderdeel of dat de

leerkracht de beoordeling voor deze leerlingen niet laat afhangen van de Cito-resultaten. Onderzoek van Brophy (1983) laat zien dat de verwachtingen van leerkrachten beïnvloed kunnen worden door factoren als sociaaleconomische status, etniciteit, leeftijd en motivatie. Dit kan leiden tot een afwijking van de daadwerkelijke capaciteiten van een leerling (de Boer et al., 2010). Het achtergrondkenmerk 'thuisstaal' is in de groep die volgens de leerkracht extra instructie nodig heeft bekeken en de verdeling hiervan leverde geen bijzonder resultaat op.

### *Beperkingen*

Een beperking van dit onderzoek is dat de meetinstrumenten 'Camelot' en 'Vergelijken' geen valide meetinstrumenten zijn. Er kan niet met zekerheid worden gezegd dat deze twee testen daadwerkelijk leerpotentieel meten. Dit maakt de resultaten moeilijker interpreteerbaar, ondanks de ruime steekproefgrootte (N=124). Daarentegen is dynamisch assessment wel een geschikte manier om leerpotentieel in beeld te brengen. Resing (2006) stelt dat op basis van dynamisch testen het vermogen om te leren gemeten kan worden. Aan de hand van het dynamisch testen kan een uitspraak worden gedaan over de mogelijke toekomstige leerprestaties van een kind. Daarnaast is bij de afname ook rekening gehouden met de theorie over het in beeld brengen van leerpotentieel. Er zijn twee manieren waarop er naar leerpotentieel gekeken kan worden en beide testen zijn op één van deze twee manieren afgenomen (Resing & Drenth, 2007; Resing, 2006). Er is beschreven dat er geen correlatie bestaat tussen de twee leerpotentieel-testen. Dit kan zijn, omdat de testen twee verschillende deelgebieden binnen het rekenen meten. Dit kan ook zijn, omdat de twee manieren van het meten van leerpotentieel een andere score opleveren. Om dit vast te stellen, zal dit in vervolgonderzoek verder getest moeten worden. Voorgaand onderzoek laat zien dat het functioneren van leerlingen beïnvloed kan worden door andere factoren, namelijk: aangeboren factoren, zoals persoonlijkheidskenmerken, en omgevingsfactoren spelen een rol bij het functioneren van leerlingen (Hoogeveen, Van Hell, Mooij & Verhoeven, 2005). De resultaten van dit onderzoek gaan over de prestaties bij rekenen. Deze beslissing is bewust genomen, maar er kan niet gegeneraliseerd worden naar het totale leerpotentieel, want taal en rekenen meten niet hetzelfde en dit levert verschillende meetgebieden (Ministerie van Onderwijs, 2009). Dit onderzoek brengt maar een kleine groep leerlingen in beeld die op de leerpotentieel-testen hoog presteren, terwijl hun Cito-resultaten en/of leerkrachtbeoordeling laag zijn. Het percentage van de groep leerlingen ten opzichte van het totaal kan beschouwd worden als een foutenmarge, maar kan ook een aanleiding vormen voor verder onderzoek.

## *Aanbevelingen*

Vervolgonderzoek kan meer inzicht leveren in de efficiëntie van leerpotentie-testen. Kunnen leerpotentie-testen als betere voorspeller van het niveau van leerlingen dienen? Om dit te achterhalen zou gedurende langere termijn op een dynamische manier getoetst moeten worden. Dit zou ook in meerdere leeftijdscategorieën gedaan moeten worden, zodat niet alleen een uitspraak wordt gedaan over de kleutergroep. Er zijn toetsen ontwikkeld om de cognitieve capaciteiten van leerlingen te meten en te vergelijken met de inschatting van de leerkracht, zoals de NSCCT (Rijksuniversiteit Groningen, z.j.). Dit is echter geen dynamische test en dat is wel van belang bij het onderzoeken van leerpotentieel. Als laatste aanbeveling geldt dat leerkrachten beter begeleid moeten worden bij het signaleren van leerlingen die onderpresteren. Wanneer zij begeleiding krijgen bij het interpreteren van resultaten en het bedenken van opvolgende acties, zullen zij meer in staat zijn om op de capaciteiten van de leerlingen te anticiperen in plaats van de gemeten resultaten.

## *Conclusie*

De vraag die in dit onderzoek centraal heeft gestaan luidt: ‘Wat is de relatie tussen de leerpotentie-scores bij kleuters uit groep 1 en het beoordelingssysteem wat scholen hanteren om leerlingen op niveau in te delen?’. De hypothese die hierbij is aangehouden is dat er leerlingen zijn die meer leerpotentieel bezitten dan hun Cito-scores of leerkrachtbeoordeling aangeven (Pfeiffer, 2011). In dit onderzoek is aangetoond dat er een kleine groep leerlingen is die op basis van het huidige beoordelingssysteem op een laag niveau wordt geschat, maar op één van de leerpotentie-testen een hoger niveau van zichzelf laat zien. Deze leerlingen hebben door middel van de leerpotentie-testen laten zien dat zij in korte tijd leerstappen kunnen maken. Zij presteren op de leerpotentie-testen namelijk niet op het niveau van de laagst presterende groep. Bij deze conclusie moet rekening worden gehouden met het feit dat deze leerpotentie-testen geen valide testen zijn, dus het is niet bewezen dat van deze leerlingen ook daadwerkelijk het leerpotentieel is gemeten. Op basis van de scores van deze leerlingen op de leerpotentie-testen zou verrijkmateriaal of leerstofaanbod op een hoger niveau overwogen moeten worden. Wanneer er van hen hogere verwachtingen zijn, zullen zij hoger presteren (Rosenthal & Jacobson, 1968). Het selecteren van leerlingen voor verrijkmateriaal is niet gebaat bij leerprestaties en de leerkrachtbeoordeling alleen (Keuning & Visser, 2013; van Batenburg, 2012). Dit onderzoek heeft laten zien dat dit voor een kleine groep leerlingen van toepassing is. Het is belangrijk om bij het leerpotentieel van leerlingen aan te sluiten, zodat de talenten van de leerlingen tot ontwikkeling kunnen komen (Gagné, 2010). Wanneer het

leerpotentieel in beeld is gebracht, kan worden aangesloten op het niveau van de leerling in plaats van de verwachting van de leerkracht. Lage verwachtingen van leerkrachten en een inadequaar leerstofaanbod kunnen bijdragen aan het ontstaan van onderpresteren (Drent & van Gerven, 2007). Potentie wordt niet altijd omgezet in prestaties, omdat een goede werkhouding, omgevingsfactoren en persoonlijkheidsfactoren ook een belangrijke rol spelen bij het excelleren van leerlingen (Platform Bètatechniek, 2012). Er kan, vanwege de invloed van deze factoren, niet vanuit worden gegaan dat leerlingen met potentie vanzelf zullen excelleren in het onderwijs (Platform Bètatechniek, 2012). Het is belangrijk dat deze leerlingen gesignaleerd worden en goede begeleiding krijgen. Wanneer onderpresterende leerlingen in beeld zijn gebracht, is het belangrijk dat het onderwijs hierop inspeelt. Riemersma & Maslowksi (2007) beschrijven dat het hierbij van belang is om inzicht in leerpotentieel te hebben, effectieve leertijd te hanteren, duidelijke doelen voor de leerling te stellen, de motivatie op te wekken en te zorgen voor individuele aandacht. Deze factoren dragen bij aan verbetering van het onderwijs aan onderpresteerders (Riemersma & Maslowksi, 2007). Aan het begin van dit artikel was te lezen dat de Inspectie van het Onderwijs stelt dat basiskwaliteit geen einddoel mag zijn in het onderwijs (Inspectie van het Onderwijs, 2013). Om te zorgen dat het Nederlandse onderwijs meer kans heeft om te excelleren, zal er meer aandacht besteed moeten worden aan het feit dat er leerlingen zijn die meer leerpotentieel bezitten dan van hen wordt verwacht. Het niveau van deze leerlingen kan worden verhoogd en als deze leerlingen door middel van dynamisch assessment al leerstappen laten zien, hoeft basiskwaliteit geen einddoel te zijn voor hen. Houd leerlingen geen plafond boven het hoofd wanneer het draait om groeien, maar benut de potentie die in een leerling schuilt. Wanneer hier meer op ingespeeld wordt, zal het Nederlandse onderwijs hier baat bij hebben. Het Nederlandse onderwijs kan worden verbeterd en wellicht hoger eindigen dan de zevende plaats op de ranglijst met beste onderwijssystemen in de wereld (Pearson, 2012). En wat nog van belangrijkere orde is: alle leerlingen krijgen de kans om zich op hun maximale niveau te ontwikkelen.

## Referenties

- Babad, E. Y. & Budoff, M. (1974). Sensitivity and Validity of Learning-Potential Measurement in Three Levels of Ability. *Journal of Educational Psychology*, 66(3), 439–447.
- Van Batenburg, A. (2012). Het signaleren en aanpakken van onderpresteren op de basisschool. GION; Gronings Instituut voor onderzoek van Onderwijs.
- Van den Bergh, D. (1999). Jongens versus meisjes: zelf- en leerkrachtbeoordeling op de CBSK en CBSL. *Journal of Educational Psychology*, 75, 631–661.
- Brophy, J. E. (1983). Research on the Self-fulfilling Prophecy and Teacher Expectations. *Journal of Educational Psychology*, 75, 631–661.
- De Boer, H., Bosker, R. J., van der Werf, M. P. C. (2010). Sustainability of Teacher Expectation Bias Effects on Long-Term Student Performance. *Journal of Educational Psychology*, 102(1), 168–179.
- Borland, J. H. (2009). Myth 2: The Gifted Constitute 3% to 5% of the Population, Moreover, Giftedness Equals High IQ, Which is a Stable Measure of Aptitude. *Gifted Child Quarterly*, 53(4), 237.
- Calero, M. D., Belen, G. M. M., & Robles, M. A., (2011). Learning Potential in High IQ children: The contribution of dynamic assessment to the identification of gifted children. *Learning and Individual Differences*, 21, 176–181.
- Cito (2010). Toetspakket Rekenen voor Kleuters.
- Dijk, H. V. & Tellegen, P. J. (2004). Nederlandse intelligentietest voor onderwijsniveau. *Handleiding en verantwoording*. Amsterdam: Boom test uitgevers.
- Doolaard, S. & Oudbier, M. (2010). Onderwijsaanbod aan (hoog)begaafde leerlingen in het basisonderwijs.
- Doornekamp, G., Drent, S. & Bronkhorst, E. (1999). Een slimme aanpak voor slimme leerlingen. Enschede: SLO.
- Drent, S. & van Gerven, E. (2007). Professioneel omgaan met hoogbegaafde leerlingen in het basisonderwijs. Assen: Uitgeverij van Gorcum, 3<sup>e</sup> druk.

- Economist Intelligence Unit (2012). *The Learning Curve 2012*. Pearson.
- Eijl, P. van, Wientjes, H., Wolfensberger, M. V. C., & Pilot, A. (2005). Het uitdagen van talent in onderwijs. In *Onderwijs in thema's* (pp. 117-156). Den Haag: Onderwijsraad.
- Feuerstein, R., Rand, Y. & Snijders, J. E. (1993). *Laat me niet zoals ik ben. Een baanbrekende methode om de cognitieve en sociale ontwikkeling te stimuleren*. Lemniscat Rotterdam.
- Fraser, B. J., Walberg, H. J., Welch, W. W. & Hattie, J. A. (1987). Synthesis of Educational Productivity research. *International Journal of Educational Research*, 11, 145–252.
- Grigorenko, E. L. & Sternberg, R. J. (1998). Dynamic Testing, *Psychological Bulletin*, 124(1), 75–111.
- Guldemon, H., Bosker, R., Kuyper, H., & Werf, Van der G. (2003). *Hoogbegaafden in het voortgezet onderwijs*. Groningen: GION.
- Hoogeveen, L. Van Hell, J. Mooij, T. & Verhoeven, L. (2005). *Onderwijsaanpassingen voor hoogbegaafde leerlingen. Meta-analyses en overzicht van internationaal onderzoek*. Nijmegen: Radboud Universiteit, CBO / ITS.
- Inspectie van het Onderwijs (2010). *Het Onderwijsaanbod aan Hoogbegaafde Leerling in het Basisonderwijs*.
- Inspectie van het Onderwijs (2013). *De staat van het onderwijs. Hoofdlijnen uit het onderwijsverslag 2011/2012*.
- Jackson, P. Z. & Mackergow, M (2003). *Oplossingsgericht denken*. Uitgeverij Thema, 98–100.
- Keesenberg, H. (2008). *Passend Onderwijs*. Kluwer.
- Keuning, J. & Visser, J. (2013). De (on)nauwkeurigheid van een leerrendementsverwachting: Ervaringen met de LVS-toetsen Rekenen-Wiskunde. *Tijdschrift voor Orthopedagogiek*, 52, 164-176.
- Kim, K. H. & VanTassel-Baska, J. (2010). The Relationship Between Creativity and Behavior Problems Among Underachieving Elementary and High School Students. *The Creativity Research Journal*, 22(2), 185–193.
- Kingma, J. & Tomic, W. (1997). *Leerpotentieel als Graadmeter van de Intelligentie*. Uitgeverij Swets & Zeitlinger, 99–114.

- Koerhuis, I. & Keuning, J. (2011). Wetenschappelijke verantwoording van de toetsen Rekenen voor kleuters.
- McClain, M. C., & Pfeiffer, S. I. (2012). Education for the gifted in the United States today: A look at state definitions, policies, and practices. *Journal of Applied School Psychology, 28*(1).
- Ministerie van Onderwijs. (2009). Referentiekader taal en rekenen: de referentieniveaus.
- Mooij, T., Hoogeveen, L., Driessen, G., Van Hell, J., & Verhoeven, L. (2007). Eindverslag van het onderzoek naar succescondities voor onderwijs aan hoogbegaafde leerlingen. Nijmegen: Radboud Universiteit, Instituut voor Toegepaste Sociale wetenschappen / Centrum voor begaafdheidsonderzoek.
- Moon, S. M., & Rosseli, H. C. (2000). Developing gifted programs. In: K.A. Heller, F.J. Mönks, R.J. Sternberg, and R.F. Subotnik (Eds.), *International Handbook of Giftedness and Talent*, 499-521. Amsterdam: Elsevier.
- Moore, D. S., McCabe, G. P. & Craig, B. A. (2009). *Introduction to the Practice of Statistics*. W.H. Freeman and Company New York.
- Mulder, L., Roeleveld, J. & Vierke, H. (2007). Onderbenutting van capaciteiten in basis- en voortgezet onderwijs. Studie ITS en SCO-Kohnstamm Instituut in opdracht van Onderwijsraad. Den Haag: Onderwijsraad.
- Onderwijsraad (2007). Presteren naar Vermogen. Uitgave van de Onderwijsraad, februari 2007.
- Pfeiffer, S. (2011). Current Perspectives on the Identification and Assessment of Gifted Students. *Journal of Psychoeducational Assessment, 30*(1), 3–9.
- Piaget, J. (1972). Intellectual Evolution from Adolescence to Adulthood. *Human Development 15*, 1-12, in *Human Development 2008, 51*, 40–47.
- Platform Bèatechniek, (2012). Drieluik Excellentie: De doorstroom van excellente leerlingen door het Primair Onderwijs.
- Reschly, D. J. (1997). Diagnostic and Treatment Utility of Intelligence Tests, in: Flanagan, D.P., Genshaft, J.L. & Harrison, P.L. (Eds), *Contemporary Intellectual Assessment: Theories, Tests and Issues*. New York, Guilford Press.



- Resing, W. C. M. (2006). Zicht op potentieel. Over dynamisch testen, variabiliteit in oplossingsgedrag en leerpotentieel van kinderen. Oratie. Universiteit Leiden: maart 2006.
- Resing, W. C. M. & Drenth, J. P. D. (2007). *Intelligentie: Meten en Weten*. Amsterdam: uitgeverij Nieuwezijds.
- Riemersma, F. & Maslowski, R. (2007). *Onderpresteren in het Primair en Voortgezet Onderwijs. Advies om onderbenutting in het onderwijs tegen te gaan*.
- Rijksoverheid, (z.j.) *Passend Onderwijs, extra aandacht op de basisschool*. Op 14 maart 2013 ontleend aan: <http://www.rijksoverheid.nl/onderwerpen/passend-onderwijs/extra-aandacht-op-de-basisschool>
- Rijksuniversiteit Groningen, (z.j.). *Niet Schoolse Cognitieve Capaciteiten Test*. Op 16 juni 2013 ontleend aan: <http://www.nscct.nl>
- Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the Classroom*. New York: Holt, Reinhart & Winston.
- Spilt, J. L., Hughes, J. N., Wu, J. Y. & Kwok, O. M. (2012). Dynamics of Teacher-Student Relationships: Stability and Change Across Elementary School and the Influence on Children's Academic Success. *Child Development*, 83(4), 1180–1195.
- Sternberg, R. J. (2002). *Succesvolle Intelligentie. Hoe praktische en creatieve intelligentie succes bepalen*. Uitgeverij Swets & Zeitlinger.
- Struiksma, C. (2012) *Duiden en doen*. Versie voor het basisonderwijs. Rotterdam: CED-groep.
- Tzuriel, D. (2000). Dynamic Assessment of Young Children: Educational and Intervention Perspectives. *Educational Psychology Review*, 12(4).
- De Vocht, A. (2011). *Basishandboek SPSS 19*. Bijleveld Press.
- Vygotsky, L. S., (1978). Interaction Between Learning and Development. *Mind and Society*, 79–91.
- Wagner, R. & Sternberg, R. J. (1984). Alternative Conceptions of Intelligence and their Implications for Education. *Review of Educational Research*, 54, 179–223.
- van Wieringen, F. (2011). *Verwachtingsvol Onderwijs*. Uitgeverij van de Onderwijsraad.
- Zaal, J. N. (2012). Schoolgedrag, schoolprestaties en de CITO toets. *Onderzoek en Analyse*.

## Bijlage I. Meetinstrument Camelot

### *Beschrijving van het meetinstrument.*

Camelot is een spel wat ingezet kan worden als verrijkingsmateriaal dat geschikt is voor kinderen tussen de 4 en 8 jaar oud. Dit spel is ontworpen door Smartgames (<http://www.smartgames.eu/nl>). De bedoeling van het spel is om een pad voor de ridder te creëren om bij zijn prinses te komen. Op een houten platform kunnen kinderen door middel van het neerzetten van torens en trappen een pad te maken voor de prins. De prins kan alleen trappen of op de bovenkant van een muur lopen. Hij kan niet springen of klimmen. Er wordt creatief denken van de kinderen gevraagd om een oplossing te bedenken. Het spel heeft een opbouw van moeilijkheidsgraad. Het kind doorloopt de levels, welke steeds moeilijker zullen worden. Er zijn verschillende niveaus, namelijk: starter, junior, expert en master.



Bij dit meetinstrument is er een selectie van levels gemaakt, omdat het niet haalbaar is om in deze test het kind alle levels te laten doorlopen. Bij het selecteren van de levels is er gekeken of er iets nieuws in het nieuwe level wordt aangeboden. Was dit niet het geval, dan werd het level eruit gelaten. Op deze manier is de reeks ontstaan die gebruikt wordt voor dit onderzoek.

### *Wat wordt er gemeten met deze afname?*

Het ruimtelijk inzicht van kinderen tussen 4 en 8 jaar wordt hiermee onderzocht. Door te kijken welke mogelijkheden er zijn om een toren of trap te plaatsen, laten de kinderen zien hoeveel ruimtelijk inzicht zij hebben om een goed pad voor de ridder te creëren.

### *Vertaling naar dynamisch assessment.*

In deze test wordt er middels dynamisch assessment gewerkt (Grigorenko & Sternberg, 1998). Dit houdt in dat er een pre-test, een interventie en een post-test moet plaatsvinden. Ieder kind zal vooraf een algemene uitleg krijgen, zodat het weet hoe de opdrachten gemaakt kunnen worden. Deze uitleg zal vooraf gestandaardiseerd worden, zodat ieder kind dezelfde uitleg krijgt, afhankelijk van het punt waarop zij stranden. De pre-test omvat het spelen van het spel tot het kind niet meer zelfstandig verder kan. De testleider heeft de goede antwoorden, dus weet wanneer een kind een fout maakt. Op dit moment volgt de interventie en zal de proefleider instructie geven over het punt waarop het kind vastliep. Na de interventie volgt een post-test. Het kind zal verder gaan met de level waar het is gestrand. Het aantal levels dat het kind verder komt, toont de leerpotentie op dat moment aan. Wanneer de leerling opnieuw vastloopt, wordt het onderzoek beëindigd. Dynamisch

testen omvat onderzoek naar leerpotentie, dynamisch assessment en meetmethoden waarbij gestructureerde testinterventie in uiteenlopende vormen plaatsvindt (Resing, 2006). Grigorenko en Sternberg (1998) spreken bij dynamisch testen over een grote verbale paraplu waaronder een heterogene diversiteit van gestructureerde testinterventie schuilgaat. Wanneer dynamisch assessment toegepast wordt, maken instructie en feedback beide deel uit van het adaptief inspelen op hetgeen wat de individu uitvoert tijdens de afname. De situatie die Resing (2006) beschrijft is ook van toepassing op deze test: 'Bij de meting van leerpotentieel van een kind gaat het er altijd om na te gaan wat het verschil is tussen probleemoplossen met en zonder hulp waarbij het merendeel van de onderzoekers de valkuil van de verschilcores wenst te vermijden.' Dit is hetgeen wat er met deze testen in beeld gebracht moet worden. Het verschil tussen het statisch en dynamisch toetsen kan andere resultaten opleveren. Resing (2006) stelt dat er op basis van dynamisch testen het vermogen om te leren in beeld gebracht kan worden. Aan de hand van het dynamisch testen kan er dus een uitspraak worden gedaan over de mogelijke toekomstige leerprestaties van een kind.

### *Leerpotentie*

Leerpotentie wordt in deze test gezien als 'de mate waarin of de efficiëntie waarmee iemand in staat is te profiteren van instructie' (Resing & Drenth, 2007).

### Handleiding afname.

Proefleider: *'Ik heb hier een leuk spel dat je mag gaan spelen. De prins en prinses zijn verliefd op elkaar, maar het probleem is dat ze niet bij elkaar kunnen komen. Kijk maar.'*

De proefleider zet het startplateau klaar en laat de beginposities van de prins en prinses zien.

Proefleider: *'Gelukkig hebben we hier een aantal torens en trappetjes staan, zodat de prins toch nog naar de prinses toe kan gaan. Jij moet hem daarbij helpen. Eerst zal ik eens laten zien hoe hij naar haar toe kan gaan.'*

De proefleider doet level 1 voor, zodat het kind ziet dat de prins een trap op kan lopen en over de bovenkant van een toren heen mag. Om aan te tonen dat de torens aan moeten sluiten, laat de proefleider ook zien wat niet kan. Er mogen geen lege plaatsen op het plateau blijven, trappen mogen niet op het smalste punt leunen en de prins kan niet springen of klimmen.

Proefleider: *'Kijk, als er een gat tussen twee torens zit, kan de prins er niet langs, want dan valt hij naar beneden. De torens of trappen moeten dus altijd naar elkaar staan, zodat de prins niet valt. En jammer genoeg kan de prins ook niet springen of klimmen. Het moet dus allemaal precies naar elkaar staan, zodat de prins over de muur heen kan wandelen.'*

De proefleider laat ook even zien dat de blokken ook gedraaid mogen worden, zolang zij maar in balans staan en niet met de trap op de grond staan. Dit is een punt dat kinderen moeten weten, maar er is ook een kans dat kinderen hier later tegenaan lopen en dit als interventie nodig zullen hebben.

Proefleider: *'Er is één klein probleem, we kunnen niet steeds alle blokken gebruiken. We moeten dus goed kijken hoe we de blokjes goed kunnen gebruiken. Kun jij nu misschien kijken of jij met de volgende blokjes de prins bij de prinses kan laten komen?'*

Op het moment dat het kind vastloopt, vraagt de proefleider eerst of het kind nog verder kan gaan. Als het kind op een juiste manier doorgaat, wordt dit niet als interventie gezien en mag het kind verder gaan. Als het kind niet door kan gaan, start de interventie:

Proefleider: *'Kun je nog verder gaan of niet?'* JA: *'Oke, goed'*. NEE: *'Kijk eens wat je ook kunt doen...'*  
En de proefleider geeft de bijbehorende instructie.

De proefleider laat het kind de volgende levels spelen: 1, 2, 3, 4, 5, 9, 10 en 11. Dit is de fase 'starter'. Als het kind nog geen fouten heeft gemaakt (en dus geen interventie nodig heeft gehad) gaan we

door met level 13, 14, 16, 19, 20 en 24. Dit is de fase 'junior'. Als het kind nog geen fouten heeft gemaakt gaan we door met level 25, 26, 27, 30, 31 36. Dit is de fase 'expert'. Als het kind nog geen fouten heeft gemaakt (waarschijnlijk lukt het eerstejaarskleuters niet om tot zover te komen) gaan we door met level 37, 39, 42, 43 en 48. Als het kind nog geen interventie nodig heeft gehad, is er voor dit kind geen interventie mogelijk bij de spel, omdat het alles beheerst. De kans dat dit bij een eerstejaarskleuter gebeurt, is zeer klein. De levels nemen in moeilijkheidsgraad toe en de interventie is gericht op het punt waarop een leerling de fout ingaat. Dus wanneer het een bepaalde blok niet juist draait, zal de proefleider hierop aansturen.

### *Interventie*

Onder interventie wordt in het onderzoek het moment verstaan wanneer de proefleider uitleg geeft over bepaalde aspecten van het spel, wanneer het kind is uitgevallen. De proefleider mag tussendoor geen uitleg geven, dus ook niet beamen wanneer een leerling bijvoorbeeld een blokje de goede kant op draait. De proefleider stelt de kinderen wel op hun gemak en complimenteert hen als ze een level hebben doorlopen. Tussendoor mogen er echter geen aanwijzingen gegeven worden. In deze interventie wordt er niet specifiek gelet op hoeveel uitleg een kind nodig heeft, omdat de leerpotentie in deze test aangegeven wordt door het aantal stappen wat de leerling na de interventie vooruit gaat. De interventie draait er dus om dat het kind de level snapt en verder kan.

### Toegestaan:

- Uitleg bij de start van het spel → volgens handleiding.
- Het kind gerust stellen wanneer het zenuwachtig is. Bijvoorbeeld: 'Doe maar rustig aan' of 'Je doet heel goed je best'. Dit moet echter op de oppervlakte blijven en mag niet inhoudelijk over de opdracht gaan.
- Het kind belonen wanneer een level is doorlopen.

### Niet toegestaan:

- Bevestigen wanneer een kind iets goed doet. Wanneer het een blokje neerzet zeggen: 'Goed zo'. Mocht dit toch gedaan worden, dan moet dit genoteerd worden op het scoreformulier!
- Antwoorden voorzeggen.
- Buiten de startinstructie en de interventie om dingen voordoen.

## Scoreformulier – Camelot

Kleurcode leerling:

Omcirkelen bij welk level het kind uitvalt en het interventiemoment dus plaatsvindt. Onderstrepen tot welk level het kind de tweede keer (met interventie) komt:

1 2 3 4 5 9 10 11 **(S)** | 13 14 16 19 20 24 **(J)**

25 26 27 30 31 36 **(E)** | 37 39 42 43 48 **(M)**

Aanvinken wat er in de interventie is behandeld:

- globale uitleg over de mogelijkheid het van het draaien van blokken.
- specifiekere uitleg: laten zien dat een lange trap ook op een paarse muur kan leunen (overlapping aantonen).
- het level voordoen.

Datum en tijdstip van de proef:

\_\_\_\_\_

Overige opmerkingen:

\_\_\_\_\_  
\_\_\_\_\_

Literatuur:

Resing, W. C. M. & Drenth, J. P. D. (2007). Intelligentie: Meten en Weten. Amsterdam: uitgeverij Nieuwezijds.

Resing, W. C. M. (2006). *Zicht op potentieel. Over dynamisch testen, variabiliteit in oplossingsgedrag en leerpotentieel van kinderen*. Oratie. Universiteit Leiden: maart 2006.

Grigorenko, E. L. & Sternberg, R. J. (1998). Dynamic Testing, *Psychological Bulletin* 124 (1), 75–111.

## **Bijlage II. Meetinstrument Vergelijken en tellen**

### *Beschrijving van het meetinstrument*

De test is samengesteld uit opdrachten afkomstig uit het boek 'Hulpboek 3: Vergelijken, Tellen, Getallen' van Citogroep primair onderwijs. Dit is een observatie- en hulpprogramma voor kleuters in de basisschool. Dit instrument is ontwikkeld door het Instituut voor Toetsontwikkeling (Cito) te Arnhem. De opdrachten zijn concreet en op de praktijk gericht, er worden dus voornamelijk materialen gebruikt. De test bestaat uit een aantal verschillende metingen. Kinderen moeten één-op-één relaties kunnen maken aan de hand van voorwerpen die in het dagelijks leven bij elkaar horen, zoals een beker en een rietje. De test betreft ook één-op-één relaties waarbij de kinderen plaatjes bij elkaar moeten leggen, zoals een plaatje van een hond en een plaatje van een baasje. Daarnaast zal er een onderdeel zijn waarbij de voorwerpen niets met elkaar te maken hebben, zoals een paperclip en een blokje. Het wordt dus minder logisch om de combinaties te maken, maar de kinderen moeten nog steeds het principe van de één-op-één relatie gebruiken. Zij vergelijken de aantallen door ze één bij één te sorteren. Bij het volgende onderdeel zal er ook met fiches gewerkt worden, zodat er geen concrete betekenis meer aan de voorwerpen zit en het meer om het tellen zelf draait. En in deze test komt ook, naast het vergelijken en het herkennen van de één-op-één relaties, het aspect 'meer/minder' aan bod. De aantallen van de voorwerpen en plaatjes die de kinderen krijgen, zullen namelijk niet altijd gelijk zijn. Bij het ene onderdeel wordt een kind gevraagd om te kijken of er genoeg rietjes voor het aantal bekers is en bij de andere vraag wordt er bijvoorbeeld gevraagd of er minder van de blokjes of van de paperclips zijn. Deze vragen verschillen per opdracht, omdat de opdrachten in moeilijkheidsgraad oplopen.

### *Wat wordt er gemeten met deze afname?*

Bij deze opdrachten wordt er vooral gekeken naar het vergelijkend vermogen van leerlingen, omdat zij de één-op-één relaties moeten kunnen maken. Daarnaast wordt er gekeken of kinderen het kardinale aspect binnen het getalbegrip beheersen. Dit betekent dat zij een getal als aanduiding van een aantal kunnen zien en in staat zijn aan te geven van welk aantal er meer of minder zijn. Dit omvat dus het tellend vermogen van de kinderen.

### *Vertaling naar dynamisch assessment*

In deze test wordt er middels dynamisch assessment gewerkt (Grigorenko & Sternberg, 1998). Dit houdt in dat er een pre-test, een interventie en een post-test moet plaatsvinden. De pre-test omvat



de test, tot het punt waarop de kinderen een fout maken. De testleider heeft de goede antwoorden, dus weet wanneer een kind een fout maakt. De interventie houdt in dat de leerlingen instructie krijgen om verder te komen in de test. In dit geval wordt er met drie niveaus van instructie gewerkt: korte instructie, verlengde instructie en begeleide inoefening (Leenders et al., 1999). Er is voor korte instructie gekozen, omdat er bij het meten van leerpotentie gekeken wordt of de leerling met zo min mogelijk instructie verder kan komen (Resing, 2006). Bij de verlengde instructie leren de kinderen vooral aan de hand van modellen (Bandura, 1965). Begeleide inoefening houdt in dat de proefleider stap voor stap het proces doorloopt met het kind (Leenders et al., 1999). De post-test omvat het deel na de interventie en dit geeft aan hoe ver het kind na de interventie kan komen.

Dynamisch testen omvat onderzoek naar leerpotentie, dynamisch assessment en meetmethoden waarbij gestructureerde testinterventie in uiteenlopende vormen plaatsvindt (Resing, 2006). Grigorenko en Sternberg (1998) spreken bij dynamisch testen over een grote verbale paraplu waaronder een heterogene diversiteit van gestructureerde testinterventie schuilgaat. Wanneer dynamisch assessment toegepast wordt, maken instructie en feedback beide deel uit van het adaptief inspelen op hetgeen wat de individu uitvoert tijdens de afname. Resing (2006) beschrijft een situatie van twee imaginaire kinderen die over eenzelfde niveau van actuele mentale ontwikkeling kunnen beschikken. Zij halen bijvoorbeeld beiden een score van 80 op een intelligentietest, maar het ene kind kan echter, met hulp, meer en complexere testopgaven oplossen dan het andere kind. Het eerste kind beschikt over een groter leerpotentieel dan het tweede. Dit is hetgeen wat er met deze testen in beeld gebracht moet worden. Het verschil tussen het statisch en dynamisch toetsen kan andere resultaten opleveren. Resing (2006) stelt dat er op basis van dynamisch testen het vermogen om te leren in beeld gebracht kan worden. Aan de hand van het dynamisch testen kan er dus een uitspraak worden gedaan over de mogelijke toekomstige leerprestaties van een kind.

### *Scoringssysteem*

In deze test wordt er gekeken naar het leerpotentieel van kinderen. Dit wordt in beeld gebracht door te kijken hoeveel interventie leerlingen nodig hebben om verder te kunnen met de test. Er wordt in deze test onderscheid gemaakt tussen drie niveaus van interventie: korte instructie, verlengde instructie en begeleide inoefening. Er wordt dus genoteerd of het kind korte instructie, verlengde instructie of begeleide inoefening heeft ontvangen. Als een leerling na de korte instructie nog niet aan de slag kan, dan gaat de proefleider verder met verlengde instructie. Als het kind hierna nog niet zelfstandig verder kan, dan wordt gaat de proefleider verder met begeleide inoefening. Na de interventie gaan de kinderen weer zelfstandig aan de slag en wordt er doorgegaan tot zij wederom

op een interventiepunt komen. Hierna wordt er niet meer verder gegaan met de proef. Korte instructie geeft aan dat het kind een hoge leerpotentie heeft, verlengde instructie geeft een gemiddelde leerpotentie aan en begeleide inoefening geeft een lage leerpotentie aan.

### *Leerpotentie*

De leerpotentie wordt in deze test gezien als het aantal hulpstappen dat een kind nodig heeft om tot de oplossing van een opgave te komen (Resing, 2006). Dit betekent dus dat naarmate een kind minder hulpstappen nodig hebben om tot de goede oplossing te komen, zij een hoger leerpotentieel hebben.

## Interventie: Vergelijken

Korte instructie

Als de proefleider hulp aanbiedt in de vorm van:

- Zit in ieder rietje nu in een beker?
- Stop in elke beker een rietje.
- *Niks voordoen, enkel instructie.*

Leerpotentie kan gemeten worden aan de hand van het aantal hints dat een kind nodig heeft om verder te kunnen. Bij de korte instructie draait het in deze proef dus alleen om het stellen van vragen en het geven van hints, waardoor het kind dus met weinig begeleiding verder zou kunnen met de oefening.

Verlengde instructie:

Als de proefleider hulp biedt in de vorm van:

- Modellen: 'Kijk, ik pak een rietje en stop het in de beker.'
- Modellen: 'Ik geef bij elke beker een rietje.'
- Modellen: 'Kijk, ik leg bij elk rood fiche een geel fiche.'
- Modellen: 'Kijk, ik zie nu dat deze twee honden geen baasje hebben. Er zijn dus meer honden dan baasjes.'
- *Modellen / voordoen om het kind verder op weg te helpen.*

(Bandura, 1965).

Begeleide inoefening:

Als de proefleider hulp biedt in de vorm van:

- Als we nu eens samen de poppetjes tellen, dan zie je dat er nog teveel poppetjes alleen zijn.
- Kijk, als ik deze twee bij elkaar leg, kun jij dan de volgende twee bij elkaar leggen?
- *Samen met het kind het desbetreffende level afmaken en daar waar nodig dingen voordoen.*

(Leenders et al., 1999).

## **De test**

Opdracht 1: De vraag wordt gesteld: 'Zijn er genoeg rietjes voor alle kinderen?'

Nodig: 10 plastic bekers, 8 rietjes.

Opdracht 2: De vraag wordt gesteld: 'Heeft elke hond een baasje om hem uit te laten?'

Nodig: 9 plaatjes van een baasje, 11 plaatjes van een hond.

Opdracht 3: Er liggen 7 blokjes en 10 paperclips. De vraag wordt gesteld: 'Waarvan zijn er minder?'

Nodig: 7 blokjes, 10 paperclips.

Opdracht 4: Er liggen 9 rode fiches en 11 gele fiches op tafel. De vraag wordt gesteld: 'Waarvan zijn er meer?'

Nodig: 9 rode fiches, 11 gele fiches.

Opdracht 5a: De leerling heeft 9 rode en 11 gele fiches liggen. De vraag wordt gesteld: 'Hoeveel gele fiches zijn er meer?'

Opdracht 5b: De vraag wordt gesteld: Als er een rood fiche weggehaald worden, hoeveel gele fiches zijn er dan meer?'

Opdracht 5c: De vraag wordt gesteld: 'Hoeveel fiches gele fiches moet je weghalen om te zorgen dat er één geel fiche minder is?'

## Scoreformulier – Vergelijken en tellen

Code leerling:

Pre-test (voltooide opdrachten): \_\_\_\_\_

Interventie: korte instructie    verlengde instructie    begeleide inoefening

Post-test (voltooide opdrachten): \_\_\_\_\_

Tijd afname: \_\_\_\_\_

Duur afname: \_\_\_\_\_

Datum afname: \_\_\_\_/\_\_\_\_/\_\_\_\_

## Literatuur

- Resing, W. C. M. (2006). *Zicht op potentieel. Over dynamisch testen, variabiliteit in oplossingsgedrag en leerpotentieel van kinderen*. Oratie. Universiteit Leiden: maart 2006.
- Grigorenko, E. L. & Sternberg, R.J. (1998). Dynamic Testing, *Psychological Bulletin* 124 (1), 75–111.
- Leenders, Y., Meyer, P. & Sanders, M. (1999). Effectieve instructie: leren onderwijzen met behulp van het Directe-Instructiemodel. Hoevelaken: CPS.
- Bandura, A. (1965). Influence of Models' Reinforcement Contingencies on the Acquisition of Imitative Responses. *Journal Of Personality and Social Psychology* 1 (6), 589–595.

### Bijlage III. Kruistabel Camelot en Leerkrachtscore

Tabel 8.

*Kruistabel met de scores van de leerpotentie-test Camelot en de Leerkrachtbeoordeling*

LP		Leerkrachtscore			Totaal
		extra instructie nodig	normale aanbod	verrijking nodig	
Camelot	0	1	0	0	1
	% LP Camelot	100.0%	0.0%	0.0%	100.0%
	% LK score	3.3%	0.0%	0.0%	0.8%
	% Totaal	0.8%	0.0%	0.0%	0.8%
	1	17	23	6	46
	% LP Camelot	37.0%	50.0%	13.0%	100.0%
	% LK score	56.7%	37.7%	22.2%	39.0%
	% Totaal	14.4%	19.5%	5.1%	39.0%
	2	4	14	4	22
	% LP Camelot	18.2%	63.6%	18.2%	100.0%
	% LK score	13.3%	23.0%	14.8%	18.6%
	% Totaal	3.4%	11.9%	3.4%	18.6%
	3	2	8	5	15
	% LP Camelot	13.3%	53.3%	33.3%	100.0%
	% LK score	6.7%	13.1%	18.5%	12.7%
	% Totaal	1.7%	6.8%	4.2%	12.7%
	4	1	4	1	6
	% LP Camelot	16.7%	66.7%	16.7%	100.0%
	% LK score	3.3%	6.6%	3.7%	5.1%
	% Totaal	0.8%	3.4%	0.8%	5.1%
	5	2	5	4	11
	% LP Camelot	18.2%	45.5%	36.4%	100.0%
	% LK score	6.7%	8.2%	14.8%	9.3%
	% Totaal	1.7%	4.2%	3.4%	9.3%
	6	2	6	6	14
	% LP Camelot	14.3%	42.9%	42.9%	100.0%
	% LK score	6.7%	9.8%	22.2%	11.9%
	% Totaal	1.7%	5.1%	5.1%	11.9%
	7	1	1	0	2
	% LP Camelot	50.0%	50.0%	0.0%	100.0%
	% LK score	3.3%	1.6%	0.0%	1.7%
	% Totaal	0.8%	0.8%	0.0%	1.7%
	8	0	0	1	1
	% LP Camelot	0.0%	0.0%	100.0%	100.0%
	% LK score	0.0%	0.0%	3.7%	0.8%
	% Totaal	0.0%	0.0%	0.8%	0.8%
Total		30	61	27	118
	% LP Camelot	25.4%	51.7%	22.9%	100.0%
	% LK score	100.0%	100.0%	100.0%	100.0%
	% Totaal	25.4%	51.7%	22.9%	100.0%

## Bijlage IV. Kruistabel Vergelijken en Leerkrachtbeoordeling

Tabel 9.

### *Kruistabel Leerpotentie-test Vergelijken en Leerkrachtbeoordeling*

		Leerkrachtbeoordeling			Totaal	
		extra instructie nodig	normale aanbod	verrijking nodig		
LP Vergelijken	begeleide inoefening	Aantal	6	13	4	23
		% Leerkrachtbeoordeling	20,0%	21,3%	14,8%	19,5%
		% Totaal	5,1%	11,0%	3,4%	19,5%
	verlengde instructie	% LP Vergelijken	26,1%	56,5%	17,4%	100,0%
		Aantal	15	24	8	47
		% Leerkrachtbeoordeling	50,0%	39,3%	29,6%	39,8%
	korte instructie	% Totaal	12,7%	20,3%	6,8%	39,8%
		% LP Vergelijken	31,9%	51,1%	17,0%	100,0%
		Aantal	9	24	15	48
Totaal	% Leerkrachtbeoordeling	30,0%	39,3%	55,6%	40,7%	
	% Totaal	7,6%	20,3%	12,7%	40,7%	
	% LP Vergelijken	18,8%	50,0%	31,2%	100,0%	
Totaal		Aantal	30	61	27	118
		% Leerkrachtbeoordeling	100,0%	100,0%	100,0%	100,0%
		% LP Vergelijken	25,4%	51,7%	22,9%	100,0%



## Bijlage V. Overzicht laag / hoog scorende leerlingen

Tabel 10.

*Laag scorende leerlingen op de Cito- en leerkrachtscore en hoog scorende leerlingen op de leerpotentie-testen*

Test	Leerlingen	Totaal
Laagscorende leerlingen Cito meetkunde $\leq 6$ (24%)	G3, G4, G11, I3, J2, J11, J12, K1, K2, K6, K9.	11
Indien laagste groep $\leq 7$ (42,5%)	F10, G3, G4, G11, G12, H3, H9, I3, I4, J2, J11, J12, K1, K2, K3, K5, K6, K9, K10, K11.	17
Laagscorende leerlingen Cito Getalbegrip – $\leq 6$ (26%)	G3, G4, G8, G10, J1, J2, J8, J11, J12, K1, K2, K6.	12
Indien laagste groep $\leq 19$ (39,1%):	F10, G3, G4, G8, G10, H12, H14, J1, J2, J4, J8, J11, J12, K1, K2, K6, K10, K11.	18
Leerkrachtscore - Extra instructie nodig	A1, A2, C9, C13, D8, D12, D13, E2, E7, E10, F6, F7, F9, G2, G3, G4, G7, G8, H7, H10, H13, I2, I4, I6, I7, J1, J2, J11, J12, K6.	30
Hoogscorende leerlingen LP Camelot - $\geq 5$ (23%)	A4, A6, A11, B4, B5, B6, C2, C7, C8, C11, D1, D3, D4, D8, D15, E11, F5, F6, F8, G5, G9, H13, I3, J3, J11, K6, K10, K11	28
Hoogscorende leerlingen LP Vergelijken – korte instructie (40%)	A3, A4, A5, A6, A7, A8, A9, A11, A14, B1, B4, B8, C3, C4, C9, C11, C12, D4, D10, D13, E3, E4, E7, E8, E9, E10, E12, F3, F5, F7, F9, F10, G1, G3, G4, G7, G9, G10, H3, H11, H14, I5, J3, J6, J7, K3, K8, K9, K11.	49

## Bijlage VI. Gemiddelde scores op alle onderdelen

Tabel 11. *Vergelijking van gemiddelde scores 'Extra instructie nodig' en de andere groep*

	N	Gemiddelde 'Extra instructie nodig'	Sd		N	Gemiddelde zonder 'Extra instructie nodig'	Sd
Citoscore	12	7.17	2.7	Citoscore Meetkunde	34	8.24	2.3
Meetkunde	12	8.00	6.1	Citoscore Getalbegrip	34	11.65	5
Citoscore	30	2.13	1.8	LP Camelot	88	2.91	4.3
Getalbegrip	30	2.10	.71	LP Vergelijken	88	2.25	3
LP Camelot	12			Valid N (listwise)		34	1.9
LP Vergelijken							3
Valid N (listwise)							.76