

Onderzoek Bachelorthesis

-Rekendrempels op de basisschool-

Studenten:	Romy Gabriels	3360180
	Annika Verkooijen	3339955
Begeleidster:	Bernadette van de Rijt	
Onderzoeksgroep:	9	
Subgroep:	1	
Datum:	24-06-2011	

Rekendrempels op de basisschool.*R. J. M. Gabriels (3360180) & A. M. Verkooijen (3339955)**Universiteit Utrecht, 2011*

Samenvatting

In dit onderzoek is gekeken of het drempelmodel terugkomt binnen het huidige realistisch rekenonderwijs. Dit is getoetst aan de hand van rekentoetsen, afgeleid van het drempelmodel, op vier basisscholen bij groep 3 tot en met 7. Bij de leerlingen uit deze groepen zijn niveautoetsen en automatiseringstoetsen afgenomen. De analyses zijn uitgevoerd op de gemiddelde drempelscores van leerlingen. Uit de resultaten kan de conclusie worden getrokken dat de vijf opgestelde drempels terug te vinden zijn binnen het huidige rekenonderwijs, dat zichtbaar is zowel binnen als tussen groepen. Dit levert empirisch bewijs voor het drempelmodel. Verder blijkt dat er geen eenduidig beeld is in de verschillen tussen jongens en meisjes in rekenprestaties. Daarnaast is er een groot verschil tussen zwakke en sterke rekenaars binnen het drempelmodel. Omdat dit onderzoek verschillende onderzoeksbeperkingen kent en niet longitudinaal is, wordt aanbevolen meer onderzoek te doen naar het drempelmodel.

Het realistisch rekenonderwijs

In de media is de laatste jaren veel aandacht gegeven aan rekenvaardigheid. Dit komt doordat men zich grote zorgen maakt over het niveau van de rekenvaardigheid in Nederland, vooral sinds de vernieuwing naar het realistisch rekenonderwijs (Gelderblom, 2007). Dit nieuwe rekenen is sinds de jaren '70 een nieuw fenomeen binnen het rekenonderwijs. De meeste basisscholen in Nederland maken al jaren gebruik van deze realistische methodes (Goffree, Munk, Markusse & Olofsen, 2005). De huidige realistische rekenmethoden zijn nog steeds onderhevig aan kritieken (Meelissen & Drent, 2008). Zo blijkt onder andere dat niet alle leerkrachten in Nederland deze nieuwe manier van rekenen optimaal toepassen in de klas (Houtveen et al., 2007).

Het nieuwe rekenen heeft het traditioneel mechanistisch rekenen naar de achtergrond gebracht (Nelissen & Van Oers, 2000). Dit was voor de jaren '70 de manier van rekenen. De kenmerken van dit traditioneel rekenonderwijs zijn onder andere het aanleren van één bepaalde strategie om rekensommen aan te pakken, het oefenen van sommen zonder context en de nadruk op automatisering van rekenen (Nelissen & Van Oers, 2000). Pas als de leerling het beoogde rekenniveau beheerst, worden sommen in een context aangeboden (Braams & Milikowski, 2008). Leren rekenen wordt gezien als een opeenstapeling van regels en geheugenkennis; de kennis en de vaardigheden worden steeds tot een hoger niveau getild. In dit onderwijs wordt leren beschouwd als het overbrengen van kennis volgens een vaststaande structuur (Milo & Ruijssenaars,

2003). Iedere leerling volgt dezelfde leerweg en er is weinig ruimte voor het verschil in inzicht van leerlingen (Nelissen & Van Oers, 2000). Kritiek op deze methode is dat de leerlingen vaak mechanisch sommen uitrekenen, waardoor ze deze operaties niet kunnen generaliseren naar soortgelijke sommen die anders gepresenteerd worden (Nelissen & Van Oers, 2000).

Het realistisch rekenonderwijs heeft andere kenmerken waarbij kinderen onder andere verschillende strategieën mogen toepassen, sommen worden aangeboden in een context die kinderen aanspreekt (Nelissen & Van Oers, 2000) en die bestaan in de echte wereld (Doorman et al., 2007). De leerkracht geeft een banende instructie, zodat er ruimte is voor eigen inbreng van het kind (Ruijsenaars, Van Luit & Van Lieshout, 2004). Realistisch rekenen is niet alleen gericht op het resultaat maar ook op de processen die dat resultaat moeten opleveren. Dit wordt verwezenlijkt in het interactief onderwijs, waarbij leerlingen worden geconfronteerd met oplossingen van anderen en dit bespreken en evalueren (Milo & Ruijsenaars, 2003a). Verschillen tussen leerlingen worden positief gewaardeerd, wat er toe leidt dat zij niet allemaal dezelfde strategieën of werkwijzen hanteren (Nelissen & Van Oers, 2000). Er wordt dan ook gepleit voor een unieke ontwikkeling van het individu. Leerlingen hebben zelf de mogelijkheid om te ontdekken, vondsten te gebruiken en problemen op te lossen. Doordat leerlingen zich individueel kunnen ontwikkelen, leren ze op hun eigen snelheid bepaalde rekenstappen te begrijpen. Als zij een bepaalde stap bereikt hebben, kunnen ze een stap verder gaan en meer ingewikkelde sommen aan. Op deze manier verloopt de rekende ontwikkeling sprongsgewijs (Nelissen & Van Oers, 2000).

Ondanks dat Nederland behoort tot één van de tien best presterende landen ter wereld op het gebied van rekenen-wiskunde (Meelissen & Drent, 2008), zijn er een aantal signalen die aangeven dat het realistisch rekenonderwijs niet zonder risico's is. Allereerst is er in Nederland een afname in rekenprestaties geconstateerd tussen 1995 en 2007 in groep 6 (Meelissen & Drent, 2008). De daling blijkt deels toe te schrijven aan de lagere prestaties van meisjes en de daling vindt plaats in de hoge vaardigheidsniveaus (Meijerink, 2008). In andere landen, zoals Engeland en de Verenigde Staten, is er een stijging van toetsprestaties waardoor de positie van Nederland verslechterd (Meelissen & Drent, 2008).

Rekenen bestaat uit drie gebieden; getallen en bewerkingen, breuken en procenten en meten en meetkunde. Uit onderzoek van Harskamp (2007, zoals geciteerd in Meijering, 2008) blijkt dat tussen 1997 en 2004 de prestaties van leerlingen uit groep 8 op het gebied van getallen en bewerkingen gelijk zijn gebleven, met uitzondering van het deelgebied vermenigvuldigen waar steeds meer fouten worden gemaakt. Wanneer gekeken wordt naar rekenen met breuken en procenten is een kleine vooruitgang te zien. Bij meten en meetkunde is er een kleine achteruitgang tussen 1997 en 2004 te zien.

Deze achteruitgang in prestaties op het gebied van meetkunde worden ook gevonden in het onderzoek van Janssen, Van der Schoot, Hemker en Verhelst (1999). Uit dit onderzoek blijkt dat ongeveer de helft van de leerlingen moeite heeft met oppervlaktematen en het uitvoeren van herleidingen met inhoudsmaten. De achteruitgang van het rekenonderwijs kan mogelijk verklaard worden door het realistisch onderwijs, waarin minder geoefend wordt met het automatiseren van basiskennis in vergelijking met het voorgaande mechanistisch onderwijs (Gelderblom, 2007).

Een ander signaal wat aangeeft dat het realistisch rekenen risico's met zich meebrengt, is kritiek op de vormgeving van realistisch rekenen. Het huidige rekenonderwijs legt het accent op het hoofdrekenen tot en met groep 5 of 6 en geeft vanaf groep 6 aandacht voor schriftelijk rekenen (Danhof et al., 2008; Inspectie van het Onderwijs, 2008). Dit heeft er onder andere toe geleid dat het automatiseren en de basiskennis van het rekenen naar de achtergrond zijn gedrongen. Op deze manier hebben steeds meer leerlingen moeite om opdrachten procedureel op te lossen met behulp van kladpapier. Uit onderzoek van Doorman en collega's (2007) blijkt dat ruim de helft van de kinderen aan het begin van groep zes geen gebruik maakt van kladpapier als zij het antwoord op een rekensom niet weten.

Andere kritiek gaat over het feit dat het realistisch rekenonderwijs weinig empirisch ondersteund is (Meelissen & Drent, 2008) en dat contextopgaven een te groot beroep doen op vaardigheden die niet direct met rekenen samenhangen (Inspectie van onderwijs, 2008). Kortom, het realistisch onderwijs is nog steeds in ontwikkeling en onderhevig aan kritieken omdat het risico's met zich meebrengt. Deze risico's zijn onder andere de dalingen van scores op bepaalde rekengebieden, het accent op hoofdrekenen en het ontbreken van empirische fundering.

Het drempelmodel

De sprongsgewijze ontwikkeling van het leren rijgen is door Danhof en collega's (2008) met een drempelmodel in kaart gebracht. Rijgen is een manier van hoofdrekenen, waarbij een getal als basis wordt gebruikt en op dit getal getalbewerkingen worden uitgevoerd ($38+46 \Rightarrow 38+40=78 \rightarrow 78+6=84$). Het drempelmodel van Danhof en collega's (2008) bestaat uit vijf mijlpalen voor rekenen die een kind behoort te doorlopen tijdens de basisschool. Drempel 1: Het automatiseren van de sommen tot tien; het betreft de plussommen, de minssommen en de splitsingen tot tien. Drempel 2: Het getalbegrip tot 100; voor het leren rijgen is met name het kunnen springen op de getallenlijn van belang. Drempel 3: Sommen tot 20; naast de automatisering van de plus- en minssommen, gaat het om de beheersing van het vlot kunnen springen over tien. Drempel 4: Bouwstenen voor het rijgen; voor het leren rijgen is het vlot kunnen uitvoeren van tientallen eraf en erbij en de sprong over het tiental tot 100 van belang.

Drempel 5: Rijen tot 100; voor het kunnen uitvoeren van de rijstrategie moeten de bouwsteensommen van drempel 4 vlot gecombineerd worden. De beheersing van de drempels bepaalt de kans op het kunnen leren rijen (Danhof et al., 2008). Dit drempelmodel wordt door onderzoek van Danhof en collega's (2009) empirisch ondersteund.

Het doel van dit onderzoek naar drempels in het rekenen op de basisschool is het toetsen in welke mate de verschillende drempels in het huidige realistisch rekenonderwijs terug te vinden zijn. De volgende vraagstelling staat centraal: 'Is het drempelmodel terug te vinden binnen het huidige realistische basisonderwijs?'. Dit sluit aan bij de vraag vanuit de schoolpraktijk om de leerontwikkeling beter in kaart te brengen (Danhof et al., 2009). De resultaten van dit onderzoek kunnen een bijdrage leveren aan het wel of niet ondersteunen van het drempelmodel. Bij een aanname van het drempelmodel kunnen methodeontwikkelaars en leerkrachten van deze informatie gebruik maken bij de vormgeving van het rekenonderwijs. Daarnaast geeft dit onderzoek een overzichtelijk beeld van de rekenscores bij het volgen van de realistische rekenmethode en waar zich mogelijk problemen voordoen. Dit kan een implicatie zijn voor een verbetering van het rekenonderwijs. Ook levert het wetenschappelijke informatie op over de toetsing van het drempelmodel binnen het huidige rekenonderwijs.

Binnen deze vraagstelling kan gekeken worden naar verschillen tussen leerlingen op drempelscores. Ieder kind kan een unieke ontwikkeling doormaken binnen het realistisch rekenonderwijs en dit kan leiden tot verschillen in rekenprestaties tussen leerlingen (Nelissen & Van Oers, 2000). Op verschillende manieren kan een verschil ontstaan in de rekenprestaties van kinderen. Allereerst verschillen leerlingen in het gebruik van rekenstrategieën, wat kan leiden tot verschillen in rekenprestaties (Nelissen & Van Oers, 2000). Een andere samenhangende factor is ervaring. Meer ervaring hangt samen met betere rekenprestaties. Dit is terug te zien in de kleuterklas, waarin kinderen gestimuleerd worden met rekenkundige aspecten bezig te zijn (Nelissen & Van Oers, 2000). Andere factoren zijn motivatie, interpersoonlijke vaardigheden en schoolbetrokkenheid (DiPerna, Volpe & Elliott, 2005). Kortom, er zijn veel manieren waarop kinderen kunnen verschillen op rekenscores. De twee leerling-kenmerken die nader bestudeerd worden in dit onderzoek zijn sekse en het niveau van rekenen.

Het verschil tussen jongens en meisjes op rekenprestaties is door verschillende studies onderzocht, met verschillende resultaten. Op bijna alle gebieden van het rekenen, zoals meetkunde en delen en vermenigvuldigen, zijn de prestaties van meisjes lager dan die van jongens. Alleen op het gebied van bewerkingen zijn de prestaties gelijk (Janssen et al., 1999). Ook halen jongens ieder jaar hogere reken-wiskundescores op de CITO-eindtoets (Van den Heuvel-Panhuizen & Vermeer, 1999). De afgelopen jaren is dit verschil kleiner geworden en sommige onderzoeken laten zelfs een gelijke verdeling zien

(Meelissen & Luyten, 2008). Het onderzoek van Meelissen en Luyten (2008) laat zien dat slechts 0.1 procent van de verschillen in rekenontwikkeling verklaard kan worden door sekse. Ander onderzoek vindt een klein verschil bij het hoofdrekenen, in het voordeel van jongens (Lynn & Irwing, 2008).

Een verklaring voor deze mogelijke verschillen tussen jongens en meisjes is dat het realistische rekenonderwijs beter aansluit bij de natuurlijke vaardigheden van jongens; het handige strategieën toepassen, zich schatstrategieën eigen maken en maatkennis ontwikkelen. Om deze vaardigheden te ontwikkelen zijn meisjes meer aangewezen op sturend onderwijs. Het realistisch onderwijs biedt, vergeleken met het mechanistisch onderwijs, niet veel sturing. De prestaties van meisjes kunnen hierdoor minder zijn (Nelissen & Van Oers, 2000). Uit onderzoek van Lynn en Irwing (2008) blijkt dat jongens meer aanleg hebben voor hoofdrekenen dan meisjes, op welk aspect van rekenen veel nadruk ligt tot en met groep 5. Dit komt mogelijk doordat het werkgeheugen, wat gebruikt wordt bij hoofdrekenen, bij jongens beter ontwikkeld is (Lynn & Irwing, 2008).

De variatie in rekencores tussen jongens en meisjes kan mogelijk voor een deel verklaard worden door de relatie tussen sekse en zelfvertrouwen op het gebied van rekenen. Jongens hebben meer zelfvertrouwen op het gebied van rekenvaardigheid, wat ertoe leidt dat zij ook beter scoren (Meelissen & Luyten, 2008). Uit onderzoek van Jacobs, Lanza, Osgood, Eccles en Wigfield (2002) blijkt dat dit zelfvertrouwen afneemt naarmate de leerlingen ouder worden. Dit neemt sneller af voor jongens dan voor meisjes, wat ertoe leidt dat het zelfvertrouwen op het gebied van rekenen gelijk is aan het begin van de middelbare school. Deze ontwikkeling kan leiden tot kleinere sekseverschillen in rekenprestaties naar mate de klas hoger wordt. In het onderzoek van Skaalvik en Skaalvik (2004) is tevens naar voren gekomen dat jongens een hoger zelfbeeld hebben dan meisjes op het gebied van rekenen, wat ondersteund wordt door het stereotype beeld dat in de samenleving bestaat over prestaties van jongens in het rekenonderwijs. Op deze manier spelen culturele en opvoedingsfactoren een rol bij rekenprestaties (Nelissen & Van Oers, 2000). Volgens Halpern en collega's (2007) is er geen eenvoudige verklaring voor een verschil tussen jongens en meisjes, maar is dit afhankelijk van vele factoren, zoals biologische, evolutionaire, sociale en omgevingsfactoren.

Behalve naar sekse is er ook onderzoek gedaan naar het verschil in rekencores tussen zwakke en sterke rekenaars. Vlotte rekenaars kijken eerst naar de opgave en bepalen daaruit welke strategie zij gebruiken (Nelissen & Van Oers, 2000). Echter, zwakke rekenaars maken vaak gebruik van één of een gering aantal strategieën (Nelissen & Van Oers, 2000). Het risico wat binnen het realistische rekenonderwijs bestaat is dat vooral zwakke rekenaars lang verkeerde of onvolledig ideeën over de oplossingstrategie behouden en stagneren (Ruijsenaars et al., 2004). Internationaal

onderzoek toont aan dat zwakke leerlingen minder profijt hebben van moderne rekenmethodes zonder schriftelijke uitwerkingen (Woodward, Monroe & Baxter, 2001). Een belangrijk kenmerk van realistisch onderwijs is banende instructie, wat minder gunstig is voor zwakke rekenaars (Ruijssenaars et al., 2004). Zo leidt een duidelijk gestuurd en herkenbare instructie van de leerkracht tot betere rekenprestaties bij kinderen met een leerstoornis (Butler, Miller, Lee & Pierce, 2001) en moeilijk lerende kinderen (Woodward et al., 2001).

Aan de hand van deze informatie over mogelijke oorzaken van verschillen tussen leerlingen en informatie over rekenscores bij de twee leerling-kenmerken zal dit onderzoek zich richten op de verschillen binnen een groep en tussen groepen van de basisschool. De eerste onderzoeksvraag betreft: 'Is er een verschil in drempelscores tussen leerlingen van dezelfde groep?'. Hierbij gaat de aandacht uit naar de gehele verdeling van drempelscores per groep, het verschil in drempelscores tussen jongens en meisjes en het verschil in drempelscores tussen zwakke en sterke rekenaars. De volgende onderzoeksvraag is opgesteld voor het verschil tussen groepen: 'Is er een verschil in drempelscores tussen verschillende groepen?'. Binnen deze onderzoeksvraag wordt gekeken naar de verschillen tussen groepen en waar die verschillen zitten. Daarnaast wordt er gekeken naar de verschillen tussen jongens en meisjes en of deze toe- of afnemen.

De resultaten van deze onderzoeksvragen dragen bij aan beter inzicht van verschillen tussen leerlingen bij het gebruik van realistische rekenmethoden. Leerkrachten en methodeontwikkelaars kunnen zich op deze inzichten afstemmen, zodat ieder kind gelijke kansen krijgt en optimaal kan profiteren van het huidig realistisch rekenonderwijs.

Methode

In dit kwantitatieve onderzoek is gebruik gemaakt van een literatuurstudie om een vraagstelling op te stellen naar aanleiding van het drempelmodel. De volgende vraagstelling staat centraal: 'Is het drempelmodel terug te vinden binnen het huidige realistische rekenonderwijs?'. Door middel van de scores van leerlingen op de opgestelde drempels wordt beoogd een antwoord te geven op de vraagstelling.

Als eerste richt dit onderzoek zich op de variatie van drempelscores binnen groepen. Hierbij gaat de aandacht als eerste uit naar de gehele verdeling van rekenscores binnen een groep, om te kijken of dat de drempels oplopen in moeilijkheid. De verwachting is dat op de oplopende drempels door leerlingen steeds slechter gescoord wordt. Daarnaast wordt er gekeken naar het verschil in rekenscores tussen jongens en meisjes van dezelfde groep. De verwachting is dat er geen verschil is in drempelscores tussen jongens en meisjes binnen een groep. Uit de literatuurstudie blijkt dat dit verschil

steeds kleiner is geworden de afgelopen jaren en dat het op dit moment de vraag is of er nog wel een verschil is. Daarnaast wordt het verschil in drempelscores tussen sterke en zwakke rekenaars per groep onderzocht, waarbij de verwachting is dat dit verschil groot is. Groot wordt in dit onderzoek gezien als scores die meer dan 25 procent van elkaar verschillen ten opzichte van de maximale score. Dit grote verschil wordt verwacht omdat het huidig realistisch onderwijs gericht is op individuele aanpakken van leerlingen bij het rekenen waardoor verschillen duidelijk zichtbaar kunnen worden. De zwakke en sterke rekenaars worden in dit onderzoek respectievelijk gezien als de leerlingen die scores hebben die behoren tot de onderste en bovenste zeven en een half procent.

Daarnaast wordt de variatie in drempelscores bekeken tussen de groepen. Hierbij gaat de aandacht als eerste uit naar verschillen in drempelscores tussen de groepen. De verwachting is dat er een significant verschil is tussen de groepen, zodat de oplopende groepen steeds beter scoren op de drempels. Ook blijkt uit de literatuurstudie dat leerlingen in de loop van hun basisschool steeds meer rekenkennis opdoen. Door het toenemen van de rekenkennis zullen de scores van de hogere groepen tevens hoger zijn. Tevens wordt binnen deze onderzoeksvraag bekeken of er een toe- of afname is in het verschil tussen jongens en meisjes tussen de verschillende groepen. Er wordt verwacht dat er een afname is in de verschillen tussen jongens en meisjes, naarmate zij in een hogere groep zitten. Uit de literatuurstudie blijkt dat dit mogelijk komt door het zelfvertrouwen in rekenen.

Deze onderzoeksvragen vragen zijn opgesteld om het drempelmodel van Danhof en collega's (2008) empirisch te toetsen. Omdat het tijdrovend is een longitudinaal onderzoek uit te voeren naar de ontwikkeling van het rekenen wordt gebruik gemaakt van een cross-sectioneel design. Dit houdt in dat kinderen van verschillende leeftijden de toets maken om op deze manier de rekenontwikkeling te onderzoeken. Daarnaast heeft dit onderzoek de opzet van een niet experimenteel toetsend onderzoek, omdat alleen getracht wordt het huidig rekenniveau in kaart te brengen.

Instrumenten

De onderzoeksvraag en deelvragen zijn in de praktijk getoetst door middel van een empirisch onderzoek waarbij leerlingen rekentoetsen maken die opgesteld zijn door Danhof en collega's (2008). De leerlingen uit groep 3 tot en met 6 krijgen twee rekentoetsen, een niveautoets en een automatiseringstoets. Groep 7 maakt alleen een gedeelte van de automatiseringstoets. Beide toetsen bevatten alle onderdelen van het drempelmodel van Danhof en collega's (2008), zodat kan worden onderzocht hoe de leerlingen scoren op de verschillende drempels. De niveautoets heeft het karakter van een vaardigheidstoets die wordt gemaakt zonder vaste tijd per onderdeel, zodat de leerling kan laten zien wat hij of zij kan. De automatiseringstoets wordt afgenomen op

tijd om te kijken wat kinderen presteren binnen een bepaald tijdsbestek. Uit de rekentoetsen kunnen drempelscores worden vastgesteld, die per niveautoets en automatiseringstoets apart geanalyseerd worden. In groep drie worden drempelscores vastgesteld voor drempel 1, 2 en 3 van de niveautoets en drempel 1 en 2 van de automatiseringstoets. Voor groep 4, 5 en 6 worden de drempelscores van alle drempels van beide toetsen berekend. Aan de hand van de resultaten van de toetsen van groep 7 kunnen alleen de drempelscores van drempel 3 en 4 van de automatiseringstoets vastgesteld worden.

De toetsen worden afgenomen door de onderzoekers of de school zelf. Door Danhof en collega's (2008) wordt aanbevolen om de niveautoetsen en automatiseringstoetsen op een verschillend tijdstip af te nemen. De toetsen zijn tussen eind januari en begin februari afgenomen om het rekenniveau halverwege het schooljaar in kaart te brengen. De leerlingen ontvangen de rekentoetsen op papier en de bladzijden worden een voor een ingevuld, waarbij de onderzoeker of de leraar kinderen begeleid en de tijd in de gaten houdt. Deze data wordt vervolgens anoniem verwerkt in het dataverwerkingsprogramma SPSS.

Participanten.

De participanten van dit onderzoek zijn leerlingen uit groep 3 tot en met 7 van vier reguliere basisscholen in Nederland binnen de provincie Noord-Brabant die voor het aanleren van rekenen gebruik maken van een realistische rekenmethode. Deze basisscholen zijn geselecteerd op basis van beschikbaarheid en woonomgeving van twee onderzoekers. Er is sprake van een niet random cluster steekproef, omdat de scholen geselecteerd zijn door de onderzoekers. De participanten zijn in principe leerlingen met een 'normale ontwikkeling', zodat aan de hand van dit onderzoek het rekenniveau en de rekenontwikkeling van de gemiddelde leerling bekeken kan worden van de reguliere basisschool. In totaal hebben er 417 participanten meegedaan aan dit onderzoek; 226 jongens (54,2%) en 191 meisjes (45,8%).

Voorafgaand aan het onderzoek is toestemming gevraagd aan de scholen om de toetsen af te nemen, waarna de scholen dit zelf gerapporteerd hebben naar de leraren, ouders en leerlingen. Alle kinderen uit groep 3 tot en met 7 zijn waar mogelijk betrokken in dit onderzoek. Als leerlingen door bijzonderheden de toets niet naar eigen niveau kunnen maken of als leerlingen niet aanwezig zijn in de klas, wordt de toets door deze leerling niet gemaakt. Ook wordt vermeld aan de scholen, dat er vertrouwelijk wordt omgegaan met de gegevens, zodat namen van kinderen en scholen niet terug zijn te vinden bij de dataverwerking.

Betrouwbaarheid en validiteit.

Het drempelmodel van Danhof en collega's wordt enkel empirisch ondersteund door een eerder onderzoek uitgevoerd door de Rijksuniversiteit Groningen (Danhof et al., 2009). Meer studies zijn nodig om aan te tonen of dit drempelmodel daadwerkelijk aanwezig is binnen dit huidig realistisch rekenonderwijs. Ook de betrouwbaarheid en validiteit van de instrumenten die gebruikt worden om het drempelmodel te toetsen zijn niet bekend. Om deze reden is het onduidelijk of de scores op de toetsen informatie geven over het drempelmodel. Met de resultaten uit de analyses van dit onderzoek moet dus met voorzichtigheid worden omgegaan.

Analyses

De scores van de leerlingen op de rekentoetsen zijn ingevoerd en geanalyseerd, waarbij de verwachtingen van het opgestelde drempelmodel zijn getoetst. Bij het analyseren van de rekenscores van de leerlingen is onderscheid gemaakt tussen de niveautoets (Niv) en de automatiseringstoets (Aut). De totale groep participanten betreft 417 leerlingen die aan één of beide toetsen geheel heeft deelgenomen. Van deze groep hebben 316 leerlingen de niveautoets en 399 leerlingen de automatiseringstoets gemaakt. Verder bestaat de totale groep uit 226 jongens (54,2%) en 191 meisjes (45,8%). De drempelscores zijn samengesteld door een totaalscore te berekenen van het aantal goede antwoorden op de sommen die behoren tot een drempel. De beschrijvende statistieken voor beide toetsen staan vermeld in Tabel 1 en 2, waarin de minimale waarde, maximale waarde en het percentage goede antwoorden is weergegeven.

De eerste onderzoeksvraag is onderzocht door het analyseren van de scores binnen de groepen. Hierbij gaat de aandacht uit naar de algemene scores binnen groepen, het verschil tussen jongens en meisjes en het verschil tussen zwakke en sterke rekenaars. De tweede onderzoeksvraag gaat over het verschil tussen de groepen, waarbij het verschil in drempelscores tussen groepen en de ontwikkeling van verschillen in sekse is onderzocht. Alle analyses zijn uitgevoerd met een significantieniveau van vijf procent. Van alle significant resultaten binnen de groepen is de effectgrootte berekend aan de hand van de Cohen's *d*. Bij alle significante resultaten tussen de groepen is dit gedaan aan de hand van de partiële *eta*. Bij de keuze van het soort analyse wordt verondersteld dat leerlingen de rekentoetsen onafhankelijk van elkaar hebben ingevuld.

Algemene scores binnen groepen

Het overzicht van alle scores op de drempels uit Tabel 1 en 2 geeft een beeld van de verdeling van drempelscores binnen een groep. Het percentage goed beantwoorde sommen per drempel geeft een indicatie van de moeilijkheidsgraad. Deze percentages nemen bij de meerderheid van de groepen per opeenvolgende drempel af, behalve

tussen drempel 2 en 3 van de niveautoets in groep 3 en 6 (zie Tabel 1). De beschrijvende statistieken ondersteunen de verwachting dat het drempelmodel bestaat uit verschillende niveaus die oplopen in moeilijkheidsgraad.

Tabel 1. *Beschrijvende statistieken van de drempelscores op de niveautoetsen*

<i>Groep</i>	<i>Drempel</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>% Goed</i>
3	Niv 1	90	15.61	3.12	4.00	18.00	86.72
	Niv 2	90	3.46	2.61	0.00	8.00	43.33
	Niv 3	90	7.82	3.78	0.00	12.00	65.20
4	Niv 1	86	17.12	1.60	10.00	18.00	95.11
	Niv 2	86	7.29	1.32	0.00	8.00	91.11
	Niv 3	86	10.40	2.29	0.00	12.00	86.67
	Niv 4	86	14.15	6.74	0.00	24.00	58.96
	Niv 5	86	2.27	2.29	0.00	8.00	28.38
5	Niv 1	76	17.62	0.71	15.00	18.00	97.89
	Niv 2	76	7.66	0.66	5.00	8.00	95.75
	Niv 3	76	11.38	0.85	9.00	12.00	94.83
	Niv 4	76	20.41	2.98	12.00	24.00	85.04
	Niv 5	76	5.76	2.30	0.00	8.00	72.00
6	Niv 1	64	17.78	0.83	12.00	18.00	98.78
	Niv 2	64	7.66	0.76	4.00	8.00	95.75
	Niv 3	64	11.55	1.59	0.00	12.00	96.25
	Niv 4	64	22.48	3.22	0.00	24.00	93.67
	Niv 5	64	7.06	1.55	0.00	8.00	88.25

Noot. Theoretisch minimum Niv1 = 0.00 en theoretisch maximum = 18.00
 Theoretisch minimum Niv2 = 0.00 en theoretisch maximum = 8.00
 Theoretisch minimum Niv3 = 0.00 en theoretisch maximum = 12.00
 Theoretisch minimum Niv4 = 0.00 en theoretisch maximum = 24.00
 Theoretisch minimum Niv5 = 0.00 en theoretisch maximum = 8.00

Tabel 2. *Beschrijvende statistieken van de drempelscores op de automatiseringstoetsen*

<i>Groep</i>	<i>Drempel</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>% Goed</i>
3	Aut 1	84	44.46	15.62	7.00	81.00	54.89
	Aut 2	84	13.83	13.27	0.00	57.00	23.05
4	Aut 1	80	69.19	12.90	27.00	81.00	85.42
	Aut 2	80	43.48	14.05	6.00	60.00	72.30
	Aut 3	80	23.33	10.95	4.00	48.00	38.88
	Aut 4	80	22.24	14.45	0.00	63.00	18.53
	Aut 5	80	2.64	2.92	0.00	14.00	6.60
5	Aut 1	76	76.25	8.13	41.00	81.00	94.14
	Aut 2	76	55.95	6.13	25.00	60.00	93.25
	Aut 3	76	38.03	12.31	12.00	60.00	63.38
	Aut 4	76	55.76	22.74	17.00	115.00	46.46
	Aut 5	76	8.96	6.64	0.00	33.00	22.40
6	Aut 1	68	79.35	3.42	60.00	81.00	97.96
	Aut 2	68	58.49	4.62	30.00	60.00	97.48
	Aut 3	68	50.15	11.46	18.00	60.00	83.58
	Aut 4	68	80.56	22.30	18.00	114.00	67.13
	Aut 5	68	16.57	9.31	0.00	36.00	41.43
7	Aut 3	91	55.38	7.17	26.00	60.00	92.30
	Aut 4	91	96.55	16.84	52.00	120.00	80.46

Noot. Theoretisch minimum Aut1 = 0.00 en theoretisch maximum = 81.00
 Theoretisch minimum Aut2 = 0.00 en theoretisch maximum = 60.00
 Theoretisch minimum Aut3 = 0.00 en theoretisch maximum = 60.00
 Theoretisch minimum Aut4 = 0.00 en theoretisch maximum = 120.00
 Theoretisch minimum Aut5 = 0.00 en theoretisch maximum = 40.00

Verskil tussen jongens en meisjes binnen een groep

Het verschil tussen jongens en meisjes op de drempelscores is getoetst door de gemiddelden met elkaar te vergelijken aan de hand van onafhankelijke *t*-toetsen. De verwachting is dat de gemiddelde drempelscores tussen jongens en meisjes niet verschillen op beide toetsen. Alleen significante resultaten zijn gerapporteerd en concreter uitgewerkt, zodat duidelijk is waar dit verschil zich bevindt en hoe groot dit verschil is. De gemiddelde drempelscores per sekse zijn beschreven in Tabel 3 en 4.

Na uitvoering van de analyses voor de niveautoetsen blijkt dat er een significant verschil is gevonden tussen jongens en meisjes op drempel 5 in groep 4, $t(84) = 2.29$, $p=.02$. Jongens ($M=2.73$) scoren gemiddeld hoger op drempel 5 dan meisjes ($M=1.60$). De effectgrootte is 0.49, wat een groot verschil aanduidt. Kijkend op somniveau, dan verschillen jongens en meisjes op het onderdeel 'minsommen over het tiental' significant. Jongens scoren gemiddeld hoger op de som '46-37', $t(84)=2.52$, $p=.01$ en op de som '58-49', $t(84)=2.80$, $p=.01$.

Op de drempelscores van de automatiseringstoetsen zijn drie significante verschillen gevonden tussen jongens en meisjes. Jongens ($M=17.26$) scoren gemiddeld hoger dan meisjes ($M=9.68$) op drempel 2 in groep 3, $t(82)=2.70$, $p=.01$. Het verschil op de drempelscores is groot, de effectgrootte is 0.57. Uit *t*-toetsen per onderdeel van deze drempel blijkt dat dit verschil significant is voor 'het getalbegrip tot 100 verder', $t(82)=2.81$, $p=.01$, 'het getalbegrip tot 100 terug', $t(82)=2.06$, $p=.04$ en 'de sprong over 10 verder', $t(82)=2.06$, $p=.04$. Daarnaast scoren jongens ($M=25.54$) gemiddeld hoger dan meisjes ($M=17.28$) op drempel 4 in groep 4, $t(78)=2.70$, $p=.01$. De effectgrootte van 0.57 duidt op een groot verschil. Uit de analyses per onderdeel blijkt dat dit verschil zich bevindt bij de 'minsommen over 10', $t(72)=2.32$, $p=.02$. Als laatste scoren jongens ($M=40.90$) ook hoger dan meisjes ($M=35.03$) op drempel 3 in groep 5, $t(74)=2.12$, $p=.04$. De effectgrootte van dit verschil is ook groot, namelijk 0.48. Dit verschil is significant voor het onderdeel 'tientallen erbij', $t(78)=1.93$, $p=.02$.

De uitkomsten van de analyses laten zien dat jongens op vier drempelscores (10.8%) van de 37 drempelscores beduidend hoger scoren dan meisjes gezien de effectgroottes. Op het merendeel van de drempelscores (89.2%) wordt niet verschillend gescoord tussen jongens en meisjes. Aan de hand van deze resultaten is de verwachting grotendeels bevestigd dat er geen verschil is tussen jongens en meisjes op de drempelscores.

Verskil tussen zwakke en sterke rekenaars

Het verschil tussen zwakke en sterke rekenaars is onderzocht door twee niveaugroepen met elkaar te vergelijken; de onderste 7.5 procent scorende leerlingen en de bovenste 7.5 scorende leerlingen op de niveautoets en de automatiseringstoets. Aan de hand van

Mann-Whitney toetsen is het verschil tussen beide niveaugroepen geanalyseerd per groep en toets. De verwachting is dat er een groot verschil is tussen zwakke en sterke rekenaars op de gemiddelde drempelscores. Naast de toetsende analyse is het verschil in

Tabel 3. *Gemiddelde drempelscores op de niveautoetsen per sekse*

Groep	Drempel	Sekse					
		Jongen			Meisje		
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
3	Niv 1	50	15.30	3.55	40	16.00	2.47
	Niv 2	50	3.64	2.59	40	3.23	2.64
	Niv 3	50	7.68	3.80	40	8.00	3.79
4	Niv 1	51	16.92	1.68	35	17.40	1.44
	Niv 2	51	7.27	1.47	35	7.31	1.08
	Niv 3	51	10.02	2.72	35	10.94	1.33
	Niv 4	51	14.98	7.47	35	12.94	5.40
	Niv 5*	51	2.73	2.53	35	1.60	1.72
5	Niv 1	41	17.61	0.70	35	17.63	0.73
	Niv 2	41	7.63	0.62	35	7.69	0.72
	Niv 3	41	11.54	0.67	35	11.20	0.99
	Niv 4	41	20.46	2.75	35	20.34	3.27
	Niv 5	41	6.05	2.27	35	5.43	2.32
6	Niv 1	34	17.71	1.09	30	17.87	0.35
	Niv 2	34	7.65	0.69	30	7.67	0.84
	Niv 3	34	11.47	2.11	30	11.63	0.67
	Niv 4	34	21.97	4.19	30	23.07	1.36
	Niv 5	34	7.06	1.63	30	7.07	1.48

* Significant verschil tussen jongens en meisjes

Tabel 4. *Gemiddelde drempelscores op de automatiseringstoetsen per sekse*

Groep	Drempel	Sekse					
		Jongen			Meisje		
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
3	Aut 1	46	47.04	17.90	38	41.34	11.83
	Aut 2*	46	17.26	15.47	38	9.68	8.49
4	Aut 1	48	69.33	14.55	32	68.97	10.15
	Aut 2	48	45.02	13.43	32	41.16	14.83
	Aut 3	48	24.60	11.76	32	21.41	9.47
	Aut 4*	48	25.54	14.97	32	17.28	12.24
	Aut 5	48	3.02	3.19	32	2.06	2.40
5	Aut 1	39	76.67	9.24	37	75.81	6.86
	Aut 2	39	54.90	7.48	37	57.05	4.12
	Aut 3*	39	40.90	11.99	37	35.03	12.09
	Aut 4	39	59.85	24.43	37	51.46	20.26
	Aut 5	39	10.10	7.38	37	7.76	5.61
6	Aut 1	38	79.05	4.15	30	79.73	2.16
	Aut 2	38	58.82	3.82	30	58.07	5.50
	Aut 3	38	52.24	11.41	30	47.50	11.13
	Aut 4	38	82.92	24.43	30	77.57	19.27
	Aut 5	38	17.34	9.43	30	15.60	9.23
7	Aut 3	44	54.84	7.78	47	55.89	6.59
	Aut 4	44	98.55	17.70	47	94.68	15.96

* Significant verschil tussen jongens en meisjes

goede antwoorden uitgedrukt in percentage van de theoretische maximale score per drempel. Een verschil groter dan 25 procent tussen zwakke en sterke rekenaars wordt in dit onderzoek geïnterpreteerd als een groot verschil. De gemiddelden en het procentueel verschil zijn opgenomen in Tabel 5 en 6. Opvallend is dat veel sterke rekenaars op de niveautoetsen de theoretische maximale score behalen (zie Tabel 5). Alleen niet significante verschillen tussen de twee groepen zijn gerapporteerd, omdat een verdere analyse naar de grootte van dit verschil dan overbodig is.

Uit de analyse van de niveautoetsen blijkt dat het verschil in scores, naast drempel 3 in groep zes, op alle drempels significant is. In groep 6 verschillen de gemiddelde drempelscores van zwakke ($M=9.00$) en sterke rekenaars ($M=12.00$) niet significant van elkaar, $U=22.50$, $z=-1.94$, $p=.05$. Kijkend naar de vuistregel van een verschil groter dan 25 procent, dan is er geen groot verschil tussen zwakke en sterke rekenaars op drempel 1 in groep vier en drempel 1, 2 en 3 in groep 5 en 6 (zie Tabel 5).

Op de automatiseringstoetsen is alleen het verschil tussen zwakke ($M=6.25$) en sterke ($M=8.00$) rekenaars op drempel 2 in groep 6 niet significant, $U=5.00$, $z=-1.62$, $p=.11$. Volgens de vuistregels is het verschil niet groot op drempel 5 in groep 4, drempel 2 in groep 5 en drempel 1 en 2 in groep 6 (zie Tabel 6).

Deze resultaten laten zien dat er op 11 drempelscores (29,7%) van de 37 vergeleken drempelscores geen groot verschil is tussen leerlingen binnen een groep. Op het merendeel van de drempelscores is het verschil tussen leerlingen groot (70,3%), wat ondersteund wordt door de significante resultaten van de Mann-Whitney toetsen. De verwachting dat er een groot verschil tussen zwakke en sterke rekenaars bestaat wordt niet volledig ondersteund door de variatie in procentuele verschillen en het niet significante verschil van twee drempelscores.

Verskil tussen de groepen

Eerst wordt er gekeken of er een verschil is tussen de verschillende groepen. Bij deze analyses is de verwachting dat er een verschil is in drempelscores tussen de groepen. Hierbij wordt verwacht dat de hoge groepen gemiddeld hogere drempelscores hebben dan de lage groepen. Als er verschil wordt gevonden in drempelscores tussen de verschillende groepen kan dit een ondersteuning zijn van het drempelmodel. Om deze verwachting te toetsen wordt er gebruik gemaakt van een meervoudige ANOVA. Bij deze toets worden alleen de significante resultaten gerapporteerd en verder uitgewerkt. De beschrijvende statistieken van deze toets zijn weergegeven in Tabel 1 en Tabel 2.

Uit de analyses blijkt dat de groep een significant multivariaat hoofdeffect heeft op de drempelscores van de niveautoetsen en de automatiseringstoetsen samengenomen, $Wilks' \Lambda=0.003$, $F(40, 1443)=138.92$, $p<.01$, $\eta^2=.78$. Dit houdt in dat 78 procent van de variatie in drempelscores verklaard wordt door de groep waarin kinderen zitten.

De vuistregel is dat deze minimaal 0.25 moet zijn, er is hier dus sprake van een hoge verklaarde variantie. Om dit hoofdeffect verder te analyseren wordt er gekeken naar de univariate uitkomsten van de verschillende drempeltoetsen. Hieruit blijkt dat op alle

Tabel 5. Gemiddelde drempelscores op de niveautoetsen van zwakke en sterke rekenaars

Groep	Drempel	Niveau van rekenen						% verschil
		Zwak			Sterk			
		N	M	SD	N	M	SD	
3	Niv 1	9	10.11	3.62	7	17.71	0.49	42.22
	Niv 2	9	1.11	0.49	7	7.86	0.38	84.38
	Niv 3	9	0.89	4.72	7	11.43	0.79	87.83
4	Niv 1	6	16.00	1.41	7	18.00	0.00	11.11
	Niv 2	6	5.50	1.64	7	8.00	0.00	31.25
	Niv 3	6	5.12	3.43	7	11.86	0.38	56.17
	Niv 4	6	3.50	1.64	7	22.71	1.38	80.04
	Niv 5	6	0.33	0.52	7	7.00	0.58	83.38
5	Niv 1	6	16.83	0.98	13	18.00	0.00	6.50
	Niv 2	6	6.67	1.21	13	8.00	0.00	16.63
	Niv 3	6	10.17	1.17	13	11.85	0.38	14.00
	Niv 4	6	15.67	2.82	13	23.46	0.52	32.46
	Niv 5	6	1.50	1.76	13	7.85	0.38	79.38
6	Niv 1	4	15.75	2.63	15	18.00	0.00	12.50
	Niv 2	4	6.25	2.06	15	8.00	0.00	21.88
	Niv 3	4	9.00	6.00	15	12.00	0.00	25.00*
	Niv 4	4	14.50	9.71	15	24.00	0.00	39.58
	Niv 5	4	2.75	2.22	15	8.00	0.00	65.63

* Geen significant verschil in drempelscores

Tabel 6. Gemiddelde drempelscores op de automatiseringstoetsen van zwakke en sterke rekenaars

Groep	Drempel	Niveau van rekenen						% verschil
		Zwak			Sterk			
		N	M	SD	N	M	SD	
3	Aut 1	6	20.00	6.63	6	70.83	8.28	62.75
	Aut 2	6	1.83	2.14	6	46.83	9.97	75.00
4	Aut 1	6	42.00	12.03	6	80.50	0.84	47.53
	Aut 2	6	18.33	11.74	6	57.83	1.17	65.83
	Aut 3	6	9.00	5.40	6	44.00	3.74	58.33
	Aut 4	6	5.33	4.37	6	50.17	6.70	37.37
	Aut 5	6	1.33	2.42	6	5.83	4.62	11.38
5	Aut 1	5	58.00	14.82	5	80.00	1.22	27.16
	Aut 2	5	46.40	13.63	5	59.60	0.54	22.00
	Aut 3	5	19.00	7.42	5	58.80	1.30	66.33
	Aut 4	5	23.60	5.98	5	105.20	6.76	68.00
	Aut 5	5	3.40	2.79	5	20.20	10.13	42.25
6	Aut 1	5	75.00	5.20	5	80.40	0.89	6.67
	Aut 2	5	53.60	9.48	5	59.40	0.89	9.67*
	Aut 3	5	24.60	5.55	5	59.80	0.44	58.67
	Aut 4	5	33.00	10.61	5	109.20	3.42	63.50
	Aut 5	5	3.80	2.86	5	31.60	3.21	69.50
7	Aut 3	6	38.00	7.18	7	59.86	0.38	36.43
	Aut 4	6	60.17	5.78	7	119.43	0.53	49.38

* Geen significant verschil in drempelscores

drempels, voor zowel de niveautoets als de automatiseringstoets, er scores van groepen zijn die hierop significant van elkaar verschillen.

De groep heeft een significant univariaat effect op drempel 1 van de niveautoetsen, $F(1,4)=568.46$, $p<.01$, $\eta^2=.85$. Dit significante effect is terug te vinden in het verschil in drempelscores tussen groep 3 en groep 5. Groep 5 ($M=17.13$) scoort gemiddeld hoger dan groep 3 ($M=15.66$). De groep waarin kinderen zitten heeft tevens een significant effect op drempel 2 van de niveautoetsen, $F(1,4)=313.93$, $p<.01$, $\eta^2=.76$. Bij deze drempel scoren groep 4 ($M=7.17$), 5 ($M=7.58$) en 6 ($M=7.25$) significant hoger dan groep 3 ($M=3.48$). Ook bij drempel 3 van de niveautoetsen is er een significant effect van de groep te vinden, $F(1,4)=296.21$, $p<.01$, $\eta^2=.75$. Hier scoren groep 4 ($M=10.18$), 5 ($M=7.58$), en 6 ($M=7.25$) significant hoger dan groep 3 ($M=8.05$). Groep 4 scoort significant hoger dan groep 5. De groep waarin de leerlingen zitten heeft tevens een significant effect op drempel 4 van de niveautoetsen, $F(1,4)=471.34$, $p<.01$, $\eta^2=.83$. Bij deze drempel scoren groep 5 ($M=20.11$) en 6 ($M=21.46$) significant hoger dan groep 4 ($M=13.43$). Bij de laatste drempel van de niveautoetsen, drempel 5 is er ook een significant effect te vinden van de groep waarin kinderen zitten, $F(1,4)=262.67$, $p<.01$, $\eta^2=.73$. Hierbij scoren groep 5 ($M=5.66$) en 6 ($M=6.69$) significant hoger dan groep 4 ($M=2.07$). Ook scoort groep 6 significant hoger dan groep 5.

Bij de automatiseringstoetsen heeft de groep waarin leerlingen zitten een significant effect op de drempelscores bij drempel 1, $F(1,4)=930.94$, $p<.01$, $\eta^2=.91$. Hierbij scoren groep 4 ($M=69.15$), 5 ($M=76.24$) en 6 ($M=79.39$) significant hoger dan groep 3 ($M=44.19$). Ook scoort groep 6 significant hoger dan groep 4 en groep 5. Bij drempel 2 heeft groep tevens een significant effect op de drempelscores. Bij drempel 2 scoren groep 4 ($M=13.47$), 5 ($M=55.98$) en 6 ($M=58.44$) tevens significant hoger dan groep 3 ($M=13.47$). Tevens scoren groep 5 en 6 significant hoger dan groep 4. $F(1,4)=655.56$, $p<.01$, $\eta^2=.87$. Ook bij drempel 3 wordt een significant effect gevonden van de groep waarin leerlingen zitten, $F(1,4)=485.10$, $p<.01$, $\eta^2=.83$. Op deze drempel scoren groep 5 ($M=37.95$), 6 ($M=49.87$) en 7 ($M=55.37$) hoger dan groep 4 ($M=23.01$). Ook scoort groep 6 significant hoger dan groep 5 en scoort groep 7 significant hoger dan groep 5 en 6. Bij drempel 4 worden tevens significante resultaten gevonden, $F(1,4)=467.31$, $p<.01$, $\eta^2=.83$. Op deze drempel scoren groep 5 ($M=55.65$), 6 ($M=80.24$) en 7 ($M=96.61$) significant hoger dan groep 4 ($M=21.41$). Groep 6 scoort significant hoger dan groep 4 en 5. Tot slot wordt bij drempel 5 een significant effect gevonden van de groep waarin leerlingen zitten, $F(1,4)=149.77$, $p<.01$, $\eta^2=.61$. Op deze drempel scoren groep 5 ($M=8.93$) en 6 ($M=16.47$) significant hoger dan groep 4 ($M=2.54$) en groep 6 significant hoger dan groep 5.

Uit deze resultaten blijkt dat de scores op de drempeltoetsen significant hoger worden naarmate de groep hoger wordt. Dit ondersteunt de verwachting dat leerlingen

steeds beter worden op de drempelscores en levert een bijdrage aan de ondersteuning van het drempelmodel.

Verskil tussen jongens en meisjes tussen de groepen

Vervolgens wordt gekeken of er een toe- of afname te vinden is in de verschillen tussen jongens en meisjes tussen de verschillende groepen. Hierbij wordt verwacht dat de verschillen tussen jongens of meisjes af zullen nemen naarmate de groep hoger wordt. Om deze vraag te beantwoorden is eerst met een MANOVA onderzocht of er een significant effect is van sekse op de drempelscores. Hieruit blijkt dat sekse een significant effect heeft op de drempelscores, $Wilks' \Lambda = 0.946$, $F(10,380) = 2.16$, $p = .02$, $\eta^2 = .54$. Dit houdt in dat 54 procent van de variantie in drempelscores verklaard kan worden door sekse. Aangezien er geen test is om de toe- of afname te bekijken is er gebruikt gemaakt van de verschillen. Deze zijn berekend aan de hand van de gemiddelde drempelscores van jongens en meisjes (zie Tabel 3 en Tabel 4). Deze verschillen zijn omgezet in procenten om zo de toe- of afname van de verschillen tussen jongens en meisjes te bekijken. De resultaten hiervan zijn te vinden in Tabel 7, waarbij de verschillen zijn uitgerekend aan de hand van verschil ten opzicht van de maximale theoretische score.

Tabel 7. *Verskil in drempelscores meisjes - jongens per groep in procenten*

Drempel	Groepen				
	Groep 3	Groep 4	Groep 5	Groep 6	Groep 7
Niv 1	3.88	2.67	0.11	0.88	-
Niv 2	-5.10	0.50	0.75	2.00	-
Niv 3	2.67	7.67	-2.83	1.33	-
Niv 4	-	-8.50	-0.50	4.58	-
Niv 5	-	-14.10	-7.75	0.13	-
Aut 1	-7.04	-0.44	-1.06	0.84	-
Aut 2	-12.63	-6.43	3.58	-1.25	-
Aut 3	-	-5.32	-9.78	-7.90	1.75
Aut 4	-	-6.88	-6.99	-4.46	-3.23
Aut 5	-	-2.40	-5.85	-4.35	-

Wat opvalt na een interpretatie van Tabel 7 is dat meisjes vaker gemiddeld hoger scoren op de niveautoetsen en jongens vaker gemiddeld hoger scoren op de automatiseringstoetsen. Ook is in deze tabel te zien dat de niveautoetsen geen duidelijk beeld van een toe- of afname in de verschillen tussen jongens en meisjes geeft. Bij de niveautoetsen wordt de verwachting dat de verschillen tussen jongens en meisjes afnemen dus niet bevestigd. Bij de automatiseringstoetsen is er een duidelijker beeld van

een afname van het verschil tussen jongens en meisjes, waarbij de verschillen op de meeste drempels afnemen. De verwachting wordt op deze toets grotendeels bevestigd. In het algemeen wordt de verwachting dat verschillen tussen jongens en meisjes afnemen naarmate de groep hoger deels bevestigd, met name voor de automatiseringstoetsen.

Conclusie en discussie

Conclusie

In dit onderzoek is gekeken naar de rekenprestaties van leerlingen op de afgenomen rekentoetsen. Deze rekenprestaties zijn omgezet naar drempelscores, zodat het drempelmodel van Danhof en collega's (2008) getoetst kan worden. Of dit drempelmodel terug te vinden is binnen het huidige realistisch rekenonderwijs is getoetst aan de hand van twee onderzoeksvragen. Deze onderzoeksvragen hebben betrekking op het verschil binnen een groep en het verschil tussen groepen van de basisschool.

Uit de resultaten van de eerste onderzoeksvraag blijkt dat de drempels in moeilijkheidsgraad oplopen. Over het algemeen neemt het percentage goed beantwoorde vragen per opeenvolgende drempel af, wat empirische ondersteuning biedt voor het drempelmodel. Daarnaast scoren jongens en meisjes binnen een groep op 89.2 procent van de drempelscores niet verschillend. Dit komt overeen met de meest recente wetenschappelijke artikelen naar het verschil tussen jongens en meisjes, waar een klein verschil of zelfs geen verschil gevonden wordt (Meelissen & Luyten, 2008; Lynn & Irwing, 2008). Concluderend kan gezegd worden dat het drempelmodel binnen de groepen bevestigd wordt voor beide seksen en dit dus ook toepasbaar is voor beide. Uit de analyses naar drempelscores van zwakke en sterke rekenaars blijkt dat over het algemeen deze twee groepen een groot verschil in drempelscores laten zien. Dit komt overeen met de literatuur die aangeeft dat vooral de verschillen tussen leerlingen in het huidige realistisch rekenonderwijs groot zijn (Ruijsenaars et al., 2004; Woodward, Monroe & Baxter, 2001). Een verschil in scores tussen zwakke en sterke rekenaars kan onder andere verklaard worden doordat zwakke rekenaars nadelen ondervinden aan het realistisch rekenonderwijs, wat in de literatuurstudie aan bod is geweest. Deze resultaten geven aan dat het drempelmodel rekening moet houden met mogelijk grote verschillen tussen leerlingen.

De resultaten van de tweede onderzoeksvraag wijzen uit dat de groep waarin leerlingen zitten een significant effect heeft op de drempelscores. Leerlingen uit hogere groepen scoren significant hoger dan lagere groepen. Dit is een bevestiging van het drempelmodel, waarbij iedere groep nieuwe vaardigheden aanleert en beter wordt in de vaardigheden die aangeleerd zijn in voorgaande groepen. De drempels, die oplopend zijn

in moeilijkheidsgraad, worden in de loop van de basisschool door leerlingen beter beheerst. Samen met de conclusie dat ook de drempels oplopen in moeilijkheidsgraad binnen een groep geeft dit een dubbel empirisch bewijs voor drempelmodel binnen het huidige rekenonderwijs, dat ook in het onderzoek van Danhof en collega's (2009) naar voren is gekomen. Daarnaast is gekeken naar de verschillen in ontwikkeling tussen jongens en meisjes. Uit de niveautoetsen blijkt geen duidelijk beeld in het verschil tussen jongens en meisjes tussen de groepen. Bij de automatiseringstoetsen is er een algemene afname in het verschil in drempelscores tussen seksen, waarbij jongens hoger scoren dan meisjes. Op het gebied van automatisering lijkt er een bevestiging te zijn van de literatuur over dit onderwerp, die aangeeft dat verschillen tussen jongens en meisjes afnemen naarmate ze ouder worden (Lanza et al., 2002; Skaalvik & Skaalvik, 2004). Dit geldt echter niet voor de niveautoetsen. Deze resultaten zijn niet eenduidig, waardoor er geen conclusie kan worden getrokken wat dit voor invloed heeft op het drempelmodel.

Uit de resultaten blijkt dat de vijf opgestelde drempels terug te vinden zijn binnen het rekenonderwijs, zowel binnen als tussen groepen. Dit levert empirisch bewijs voor het drempelmodel. Verder blijkt dat er geen eenduidig beeld is in de verschillen tussen jongens en meisjes in rekenprestaties en is er vaak een groot verschil tussen zwakke en sterke rekenaars binnen een groep.

Beperkingen van het onderzoek

De bovenstaande resultaten dienen met voorzichtigheid te worden geïnterpreteerd, vanwege een aantal beperkingen van dit onderzoek. Ten eerste is de steekproef geselecteerd en betreft het alleen basisscholen uit provincie Noord-Brabant, waardoor dit moeilijk generaliseerbaar is naar de rest van Nederland. Daarnaast is de steekproef van het onderzoek is niet groot en zijn de gemaakte groepen zwakke en sterke rekenaars klein. Hierdoor kan de invloed van omstandigheden en fouten een grote invloed hebben op de uitkomst van dit onderzoek. Ook zijn de participanten niet random gekozen, maar geselecteerd. Om deze redenen zijn de resultaten moeilijk te generaliseren naar andere scholen, leerlingen en omstandigheden. Dit betekent dat de externe validiteit van het onderzoek niet hoog is.

Vervolgens is de afname ook gedaan door zowel de school zelf als de onderzoekers. Wanneer de toets is afgenomen door de school, is het onzichtbaar voor de onderzoekers hoe de toets is afgenomen. Daarnaast zijn de toetsen door twee verschillende onderzoekers afgenomen. Hierdoor kan er een verschil zijn in de afnameprocedure, wat invloed kan hebben op de resultaten. Wel is er een richtlijn opgesteld door Danhof en collega's (2008) om de toetsen af te nemen, zoals de maximale tijd per onderdeel en de instructie die gegeven wordt. Dit kan de intra beoordelaars betrouwbaarheid vergroten, omdat de instructors meer gelijkenis vertonen.

Door Danhof en collega's (2008) is aangegeven dat de twee toetsen in twee verschillende meetmomenten afgenomen moeten worden om de concentratie van de leerlingen te behouden. Op twee scholen zijn de automatiseringstoetsen en niveautoetsen in één meetmoment afgenomen, waardoor de concentratie en inzet van deze leerlingen lager kan zijn. Dit kan ervoor zorgen dat de resultaten van de vier scholen niet vergelijkbaar zijn.

Uit de observaties bij de afnamen is gebleken dat bepaalde groepen kinderen gefrustreerd worden van het feit dat ze bladzijden te moeilijk vinden of niet af mogen maken. Een aantal leerlingen zijn afgehaakt tijdens het meetmoment of hebben geen poging gedaan om bepaalde bladzijden te maken. De leerlingen in groep vier raakten bijvoorbeeld gefrustreerd door de te moeilijke sommen. Hierdoor is de kans aanwezig dat zij niet hun huidige rekenpotenties hebben laten zien. De leerlingen die afgehaakt zijn, zijn uit de steekproef gehaald, maar de leerlingen die frustraties lieten zien en de volledige toets hebben gemaakt zijn in de steekproef gebleven. Deze leerlingen kunnen een negatieve invloed hebben op de gemiddelde drempelscores.

Ook is de betrouwbaarheid en validiteit van de rekentoetsen die behoren tot het opgestelde drempelmodel niet bekend. Het is niet bekend of de drempeltoets werkelijk de vaardigheden in het rijgen en de daarbij behorende drempels meet. Daarnaast is pas met een studie empirisch bewijs geleverd voor het drempelmodel, wat aangeeft dat dit nog een vrij nieuw idee is wat meer onderzoek vergt.

Met de resultaten van dit onderzoek kan geen vergelijking worden gemaakt met onderwijs uit andere perioden, aangezien het maar een testmoment is. Zo is het onder andere niet mogelijk om deze resultaten te vergelijken met resultaten bij mechanistische rekenmethoden. Het is dus niet duidelijk of er ook al een groot verschil was tussen zwakke en rekenaars en een verschil in scores tussen seksen. Ook kan er op basis van deze gegevens geen goede handreiking worden gedaan naar beter realistisch onderwijs, omdat het niet afgezet kan worden naar een basisstandaard voor rekenen.

Aanbevelingen

Uit dit onderzoek kunnen een aantal aanbevelingen worden gedaan voor verder onderzoek. Allereerst is meer onderzoek nodig naar het verschil tussen jongens en meisjes, omdat hierover zowel uit dit onderzoek, als uit de literatuur geen eenduidig beeld naar voren komt. De meest recente literatuur die gebruikt is voor dit onderzoek laat een kleiner verschil zien tussen jongens en meisjes dan oudere literatuur. Het is echter onduidelijk of deze ontwikkeling zich voort heeft gezet.

Aan de hand van de gegevens over de verschillen tussen zwakke en rekenaars, kunnen tevens aanbevelingen gedaan worden. Het grote verschil tussen zwakke en sterke rekenaars heeft implicaties voor het huidige rekenonderwijs, omdat er aan de hand

van dit onderzoek blijkt dat er in het huidige rekenonderwijs rekening moet worden gehouden met dit verschil. Er kan meer onderzoek gedaan worden naar de meest optimale manier om les te geven aan zwakke rekenaars, zonder sterke rekenaars te benadelen. Op deze manier zullen beiden groepen rekenaars zich voldoende kunnen ontwikkelen binnen het rekenonderwijs.

Over het algemeen moet nog meer onderzoek worden gedaan naar het drempelmodel. Om de ontwikkeling van het drempelmodel in kaart te brengen, is er meer longitudinaal onderzoek nodig naar het drempelmodel. Met een duidelijker beeld van het drempelmodel, kunnen de resultaten beter geïnterpreteerd worden. Aanbevolen wordt om bij longitudinaal onderzoek meer gebruik te maken van grotere aselect gekozen groepen leerlingen, verspreid door heel Nederland.

Tot slot zal meer onderzoek gedaan moeten worden naar de validiteit en betrouwbaarheid van de toetsen behorend bij het drempelmodel van Danhof en collega's (2008). Aanbevolen wordt om te onderzoeken of de toetsen werkelijk de ontwikkeling van het rijgen en de daarbij behorende drempels meten. Ook moet worden nagegaan of de opgaven binnen de toets goed aangepast zijn aan het niveau van onderwijs binnen de groepen. Dit kan frustraties van leerlingen in het vervolg verminderen, waardoor de scores een betere weergave van de werkelijkheid zullen zijn.

De resultaten van dit onderzoek met betrekking tot algemene bevestiging van het drempelmodel dragen bij bestaande wetenschappelijke kennis van de rekenontwikkeling van kinderen op de basisschool. Van deze kennis kunnen nieuwe rekenmethodeontwikkelaars gebruik maken.

Literatuurlijst

- Butler, F. M., Miller, S. P., Lee, K., & Pierce, T. (2001). Teaching mathematics to students with mild-to-moderate mental retardation: A review of literature. *Mental Retardation, 39*, 20-31.
- Braams, T., & Milikowski, M. (2008). Hoe wordt een kind een gelukkige rekenaar? In T. Braam & M. Milikowski (Eds.), *De gelukkige rekenklas* (pp 13-24). Amsterdam: Boom.
- Danhof, W., Bandstra, P., Milo, B., Mushati-Hamadani, E., Minnaert, A., & Ruijssenaars, W. (2008). Onderzoeksproject leerbaarheid van hoofdrekenen –naar criteria voor differentiatie en/of planning-. *Panama-post, 27*, 24-28.
- Danhof, W., Bandstra, P., Milo, B., Mushati-Hamadani, E., Minnaert, A., Ruijssenaars, W., ... Van Vliet, L. (2009). *Onderzoeksproject leerbaarheid van het hoofdrekenen. Rapport van een eerste verkenning*. Drachten: Rijksuniversiteit Groningen.
- DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in the elementary grades. *Journal of School Psychology, 43*, 397-392.
- Doorman, M., Drijvers, P., Dekker, T., Van den Heuvel-Panhuizen, M., De Lange, J., & Wijers, M. (2007). Problem solving as a challenge for mathematics education in the Netherlands. *Mathematics Education, 39*, 405-418.
- Gelderblom, G. (2007). Elk kind kan rekenen. Effectieve zorg in de rekenles en de rol van de schoolleider. *Basisschool Management, 20*, 1-6.
- Goffree, F., Munk, F., Markusse, A., & Olofsen, N. (2005). *Gids voor rekenen en wiskunde groep 3/4*. Groningen: Noordhoff Uitgevers.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). Science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1-51.
- Houtveen, T., Van de Grift, W., Kuijpers, J., Boot, M., Groot, F., & Kooijman, H. (2007). Improving underperforming schools. *Journal of Education for Students Placed at Risk, 12*, 361-381.
- Inspectie van het Onderwijs (2008). *Basisvaardigheden rekenen-wiskunde in het basisonderwijs. Een onderzoek naar het niveau van rekenen-wiskunde in het basisonderwijs en naar verschillen tussen scholen met lage, gemiddelde en goede reken-wiskunderesultaten*. Utrecht: Inspectie van het Onderwijs.
- Jacobs, J. E., Lanze, S., Osgood, D. W., Eccles, S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509-527.

- Janssen, J., Van der Schoot, F., Hemker, B., & Verhelst, N. (1999). *Balans van het rekenwiskundeonderwijs aan het einde van de basisschool 3. Uitkomsten van de derde peiling in 1997*. Arnhem: Citogroep.
- Lynn, R., & Irwing, P. (2008). Sex differences in mental arithmetic, digit span, and *g* defined as working memory capacity. *Intelligence*, 36, 226-235.
- Meelissen, M. R. M., & Drent, M. (2008). *TIMSS 2007. Trends in leerprestaties in exacte vakken in het basisonderwijs*. Enschede: Universiteit Twente.
- Meelissen, M., & Luyten, H. (2008). The Dutch gender gap in mathematics: Small for achievement, substantial for beliefs and attitudes. *Studies in Educational Evaluation*, 34, 82-93.
- Meijerink, H. (2008). *Over drempels met taal en rekenen*. Enschede: SLO.
- Milo, B. F., & Ruijsenaars, A. J. J. M. (2003a). Instructie en leerlingkenmerken. Onmogelijkheden van realistische instructie in het SBO. *Tijdschrift voor Nascholing en Onderzoek van het Reken-wiskundeonderwijs*, 22, 27-33.
- Milo, B. F., & Ruijsenaars, A. J. J. M. (2003b). Rekeninstructie op scholen voor speciaal basisonderwijs, wat is realistisch? *Tijdschrift voor Orthopedagogiek*, 42, 423-435.
- Nelissen, J., & Van Oers, B. (2000). *Reken maar! Reflecties op de praktijk*. Baarn: Bekadidact.
- Ruijsenaars, A. J. J. M., Van Luit, J. E. H., & Van Lieshout, E. C. D. M. (2004). *Rekenproblemen en dyscalculie. Theorie, onderzoek, diagnostiek en behandeling*. Rotterdam: Lemniscaat
- Skaalvik, S., & Skaalvik, E. M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles*, 50, 241-252.
- Van den Heuvel – Panhuizen, M., & Vermeer, H. J. (1999). *Verschillen tussen meisjes en jongens bij het vak rekenen-wiskunde op de basisschool. Eindrapport MOOJ-onderzoek*. Culemborg: Technimedia.
- Woodward, J., Monroe, K., & Baxter, J. (2001). Enhancing student achievement on performanceassessments in mathematics. *Learning Disabilities Quarterly*, 24, 33-45.