



Universiteit Utrecht

De predictieve validiteit van de UGT-R

**Een onderzoek naar de voorspellende waarde van testcores op de Utrechtse
Getalbegrip Toets - Revised met betrekking tot voorbereidende rekenprestaties**

Utrecht, juni 2010

Master Orthopedagogiek, werkveld Leerlingenzorg

Universiteit Utrecht

Student: J. C. E. M. Bessler (0414603)

Thesisbegeleider: Prof. Dr. J. E. H. Van Luit

Voorwoord

Dit onderzoeksverslag betreft het afstudeeronderzoek naar de predictieve validiteit van de Utrechtse Getalbegrip Toets – Revised. Onderzocht is of de voorbereidende rekenprestaties van leerlingen uit groep 2, 3 en 4 van het reguliere basisonderwijs voorspeld kunnen worden op basis van hun eerder behaalde testcores op de UGT-R, afgenomen in groep 1, 2 en 3. Mijn voorkeur is naar dit onderwerp uitgegaan vanwege mijn interesse voor het rekenen bij kinderen in het basisonderwijs.

Ik wil graag mijn dank uitspreken naar iedereen die een bijdrage geleverd heeft aan de totstandkoming van dit onderzoeksverslag. In de eerste plaats wil ik mijn begeleider Hans van Luit bedanken voor het geven van heldere en deskundige feedback en voor zijn enorme geduld en prettige begeleiding tijdens alle fasen van dit onderzoek. Ook dank ik Bernadette de Rijt voor haar ondersteuning bij alle statistische analyses.

Graag dank ik de directie, intern begeleiders, leerkrachten en leerlingen van de School op de Berg in Amersfoort, de Damiaanschool te Zeist en de Valkenheuvel te Driebergen-Rijsenburg voor hun inzet gedurende de dataverzameling.

Daarnaast bedank ik mijn ouders en vriend voor hun liefde, vertrouwen en geduld. Tot slot bedank ik mijn vriendinnen voor hun steun, bemoedigende woorden en afleiding.

Hanneke Bessler,

Utrecht, juni 2010

Inhoudsopgave

| | |
|---|-----------|
| Abstract | 4 |
| Inleiding | 5 |
| <i>Getalbegrip en voorbereidende rekenvaardigheden</i> | 5 |
| <i>Belang voorbereidende rekenvaardigheid en verband latere rekenprestaties</i> | 7 |
| <i>Het onderzoek</i> | 10 |
| Methode | 11 |
| <i>Participanten</i> | 11 |
| <i>Meetinstrumenten</i> | 12 |
| <i>Procedure</i> | 13 |
| Resultaten | 13 |
| Conclusie en discussie | 16 |
| Literatuurlijst | 20 |

Abstract

Background: Mathematics is an ability used in every day life and starts developing at a young age. A lot of research has been done regarding early numeracy and it's relevance towards mathematical competence and performance. Researchers have demonstrated that difficulties with early numeracy in kindergarten can interfere with the acquisition of mathematical skills. This means that the problems children in kindergarten have in acquiring early numeracy also have consequences for their mathematical skills later on. Research has also shown that an insufficient development of early mathematical competence can lead to arithmetic difficulties later in life. Early math intervention is therefore important. **Aim:** The purpose of this study is to investigate whether the ENT-R testscores of low-performing young children in kindergarten and first grade are predictive their mathematical performance one year later. Differences between ENT-R comprehension and testscores are examined, as well as the correlation between these two. **Method:** The Utrecht Early Mathematical Competence Test-Revised has been used to assess the hypothesis mentioned above. Out of 2137 young pupils, aged 4 to 7 years old, 41 pupils were selected on a bases of availability and 34 pupils actually participated in the present study. Children in kindergarten and grade 1 and 2 were tested with the ENT-R during three assessments to examine their early mathematical competence. **Results:** In the first year of pre-primary education significant differences where found between ENT-R comprehension scores during three testing cycles. In the last year of pre-primary education significant differences were found in ENT-R comprehension scores between assessment 2 and 3 and between assessment 1 and 3. In grade 1 no significant differences where found between ENT-R comprehension scores during the three testing cycles. ENT-R comprehension scores during the first testing cycle were predictive for the following ENT-R scores, more on short-term than on long-term. **Conclusion:** The results indicate that the ENT-R is a moderate predictive valid instrument. The ENT-R testscores of children in kindergarten and first grade are indeed predictive for their mathematical performance one year later, whereas the mathematical performance of children in the last year of pre-primary education could not be predicted accurately based on the ENT-R testscores measured in the first year of pre-primary education. In kindergarten and grade 1 the ENT-R is more predictive on short-term than on long-term. Since the design and sample size of this research contains several limitations, caution in interpreting the results is highly recommended. Nevertheless, pupils with extremely low testscores should be carefully monitored during their mathematical learning process in order to perform early intervention and prevent the development of mathematical difficulties.

Keywords: *Early numeracy, mathematical performance, predictive validity, ENT-R*

Inleiding

Getalbegrip en voorbereidende rekenvaardigheden

Getallen spelen een grote rol in het alledaagse leven (Rousselle & Noe, 2007). Zelfs op jonge leeftijd zijn kinderen al geïnteresseerd in getallen, getalfeiten en de manier waarop getallen gebruikt kunnen worden (Torbeyns et al., 2002). Zij zien hoe volwassenen bezig zijn met getallen, bijvoorbeeld tijdens het afrekenen in de winkel of bij het aangeven van de tijd. Het niet effectief kunnen tellen of het niet begrijpen van de betekenis van getallen wordt al snel een 'handicap'. Dit geldt niet alleen voor op school, maar ook met betrekking tot het functioneren in de samenleving (Rousselle & Noe, 2007). Daarom is het van belang dat kinderen van jongs af aan goed leren rekenen. Vroege signalering van problemen is belangrijk (Van de Rijt & Van Luit, 1999) en eventuele interventie heeft tijdens de vroege ontwikkeling de beste kans van slagen (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Melhuish et al., 2008; Sophian, 2004).

Volgens Torbeyns et al. (2002) moet het ontwikkelen van rekenkennis en rekenvaardigheden gezien worden als een ontwikkelingsproces dat reeds voor aanvang van het primaire, formele onderwijs begint. Het blijkt dat jonge kinderen in de leeftijd van 3 tot 5 jaar al een klein aantal concrete objecten kunnen tellen, groeperen, onderscheiden en vergelijken. Ze beschikken al over uiterst eenvoudige rekenvaardigheden, concepten en strategieën als ze naar de basisschool gaan (Aubrey, 1993; 1994; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Neuman & Roskos, 2005). Deze vroege vaardigheden ontwikkelen zich geleidelijk door de ervaring met getallen die kinderen opdoen in situaties binnen en buiten school en vormen de basis voor het formele rekenen zoals dat in het verdere basisonderwijs wordt geleerd (Aunio, Hautamäki, Heiskari, & Van Luit, 2006) en worden tezamen getalbegrip genoemd (Torbeyns et al., 2002). Ruijsenaars, Van Luit en Van Lieshout (2006, p. 42) hanteren de volgende definitie van getalbegrip; "Wanneer een kind op ieder moment tijdens het aftellen van losse elementen elk telwoord zowel opgevat als aanduiding van het hoeveelste getelde element als van het totale aantal tot dan toe getelde elementen, spreken we van getalbegrip". De belangrijkste kenmerken van getalbegrip zijn conservatie, correspondentie, classificatie, seriatie, tellen, rekentaal en maatbegrip (Ruijsenaars et al., 2006). In de literatuur (Baltussen, 2000; Torbeyns et al., 2000) worden de begrippen *getalbegrip*, *voorbereidende rekenvaardigheid*, *ontluikende gecijferdheid* als synoniemen gebruikt.

De ontwikkeling van voorbereidende rekenvaardigheden bij kinderen is een onderwerp dat veel onderzoekers en betrokkenen in het werkveld bezighoudt. Verschillende onderzoeken

naar de ontwikkeling van rekenvaardigheid (Aunio, Hautamäki, Heiskari, & Van Luit, 2006, Geary, 1994; Neuman & Roskos, 2005; Torbeyns et al., 2002; Van de Rijt, 1996) lijken tot dezelfde conclusie te komen: de verwerving van rekenkennis en -vaardigheden kan gezien worden als een ontwikkelingsproces dat start al ruim voor het formele rekenwiskundeonderwijs begint. Ondanks deze overeenstemming over het bestaan van deze vroege rekenkennis en rekenvaardigheden zijn onderzoekers minder eenstemmig wat betreft de manier waarop deze zich ontwikkelt en welke invloed het bezit van deze kennis en vaardigheden heeft op het verdere leren (Torbeyns et al., 2002).

Er bestaan drie verschillende inzichten ten aanzien van de manier waarop de ontwikkeling van voorbereidende rekenvaardigheden onderverdeeld kan worden (Torbeyns et al., 2002). Een eerste inzicht is gebaseerd op het werk van Piaget (1965). Volgens zijn cognitief ontwikkelingstheoretische visie liggen logische principes aan de basis van de ontwikkeling van getalbegrip. De vier traditionele ‘piagetiaanse’ voorwaarden kunnen gezien worden als kenmerken van het logisch leren denken zoals dat doorgaans bij kleuters tot ontwikkeling komt: conserveren, classificeren, corresponderen en seriëren (Ruijsenaars et al., 2006). Getalbegrip berust op de synthese van deze logische operaties (Van Luit & Van de Rijt, 2005). Aangezien het logisch denken ook het kwantificeren van logische relaties mogelijk maakt, worden deze voorwaarden over het algemeen opgevat als voorwaarden voor het ontstaan van getalbegrip en daarmee als rekenvoorwaarden. De ontwikkeling van de logische operaties wordt gestimuleerd door de interactie van het kind met concrete materialen in zijn omgeving. Wanneer een kind deze vaardigheden beheerst, heeft het een feitelijk begrip van getallen en is het kind klaar om rekenvaardige operaties uit te voeren. Het kennen van de namen van getallen en het kunnen tellen ziet Piaget (1965) niet als noodzakelijke voorwaarden voor de ontwikkeling van getalbegrip (Torbeyns et al., 2002).

De tweede visie beschrijft de ontwikkeling van rekenvaardigheden daarentegen grotendeels in termen van de ontwikkeling van telvaardigheden. Frequente en gevarieerde ervaringen met tellen worden als noodzakelijk beschouwd voor de ontwikkeling van getalbegrip en tevens voor de vaardigheden optellen en aftrekken (Torbeyns et al., 2002). Wat betreft de fasering en de hiermee verbonden leeftijden is veel overeenstemming tussen de verschillende onderzoekers (Van Luit & Van de Rijt, 2005). Binnen deze benadering wordt naast het kijken naar de fasering van ontwikkeling ook gekeken naar de context waarbinnen deze ontwikkeling plaatsvindt. Fuson (1988) beargumenteerde dat getallen in verschillende contexten verscheidene betekenissen hebben. In de eerste plaats leren kinderen getallen als contextgebonden woorden, later pas worden de verschillende betekenissen van deze

getalwoorden aan elkaar verbonden (Wynn, 1992). Uit onderzoek blijkt dat telvaardigheid een goede voorspeller is voor getalbegrip (Passolunghi, Verelloni, & Schadee, 2007; Van de Rijt, 1996).

Een derde benadering combineert de ideeën van de eerste twee inzichten. In deze benadering worden beide inzichten als cruciaal gezien voor de ontwikkeling van vroeg getalbegrip (Torbeyns et al., 2002). Diverse studies (Aunio, 2006; Fuson, 1988; Geary, 1995; Steffe & Cobb, 1988) leiden tot de conclusie dat de traditionele voorwaarden, zoals door Piaget (1967) uitgewerkt, niet zozeer voorwaardelijk zijn voor de telvaardigheden, maar dat ze er wel aan gerelateerd zijn en gezamenlijk met de telvaardigheden de voorbereidende rekenvaardigheid vormen. Deze combinatie van beide wordt voorbereidende rekenvaardigheid genoemd (Ruijssenaars et al., 2006; Torbeyns et al., 2002; Van de Rijt, 1996; Van de Rijt, Van Luit, & Pennings, 1999).

Belang voorbereidende rekenvaardigheid en verband met latere rekenprestaties

Van de Rijt en Van Luit (1998) beschrijven voorbereidende rekenvaardigheid als een cognitieve structuur die het hele domein van vroege rekenvaardigheid bedekt en waarin verschillende vaardigheden en kennis zich ontwikkelen. Deze voorbereidende rekenvaardigheid, met name telvaardigheid, blijkt een uitstekende voorspeller voor de rekenvaardigheid in groep 3 en 4 (Aubrey, Dahl, & Godfrey, 2006; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Desoete & Grégoire, 2006; Jordan, Kaplan, Locuniak, & Ramineni, 2007; Koponen, Aunola, Ahonen, & Nurmi, 2008; Kurdek & Sinclair, 2001).

Deze vroege rekenvaardigheid bestaat uit acht componenten. De eerste vier componenten zijn de logische principes van Piaget, te weten ‘vergelijken’, ‘classificeren’, ‘corresponderen’ en ‘seriëren’. De laatste vier componenten zijn de telvaardigheden, te weten het ‘gebruiken van telwoorden’, ‘synchroon tellen’, ‘resultatief tellen’ en ‘kennis van getallen’. Deze acht componenten van vroege rekenvaardigheid vormden de basis voor de ‘schalen’ van voorbereidende rekenvaardigheid van de Utrechtse Getalbegrip Toets (Van de Rijt & Van Luit, 1998). De Utrechtse Getalbegrip Toets (UGT; Van Luit, Van de Rijt & Pennings, 1994) is een instrument dat voorbereidende rekenvaardigheid meet bij kinderen (van 4.0 tot 7.6 jaar) van de eerste drie groepen van de basisschool (Van de Rijt en Van Luit, 1998). Een voldoende niveau van getalbegrip wordt immers noodzakelijk geacht wanneer een leerling in groep 3 het formele rekenonderwijs wil volgen. Aan de hand van scores op de UGT is het mogelijk de ontwikkeling van de voorbereidende rekenvaardigheden te meten en te volgen (Van Luit & Van de Rijt, 2005). Hierdoor kunnen eventuele achterstanden al vroeg

in de ontwikkeling opgemerkt worden en leerlingen adequaat begeleid worden, zodat de achterstand niet groter zal worden (Bosker, 2002).

Inmiddels zijn binnen het veld van de leerproblemen een aantal valide instrumenten ontwikkeld die relatief goed kunnen voorspellen welke leerlingen waarschijnlijk moeilijkheden met lezen zullen ondervinden. Deze instrumenten worden in toenemende mate gebruikt voor screeningsdoeleinden. Daarnaast geven ze scholen de mogelijkheid vroegtijdig te interveniëren en extra hulp en ondersteuning te bieden aan kinderen in kleuterklassen en in groep 3 (Gersten, Jordan, & Flojo, 2005).

Er is veel onderzoek gedaan naar de ontwikkeling van vroege rekenvaardigheden (Carpenter, Ansell, Franke, Fennema, & Weisbeck, 1993; Fuson, 1988; Geary, 1995; Sophian, 1987), maar er is weinig bekend over de effecten van vroege interventies (Van de Rijt & van Luit, 1998). De ontwikkeling van valide vroege interventies voor latere wiskundige bekwaamheid staat nog in de kinderschoenen. Wel is er in toenemende mate (h)erkenning voor het belang van de vroege jaren onderwijs, net als een uitgebreide interesse in het lange termijn effect van voor/vroeg schoolse educatie op latere prestaties (Aubrey & Godfrey, 2003). Resultaten gevonden in een longitudinaal onderzoek (Aubrey & Godfrey, 2003) naar het getalbegrip van jonge kinderen, tonen aan dat de scores van kinderen in de leeftijd van 5 jaar voorspellend waren voor de latere rekenprestaties wanneer de kinderen 7 jaar oud waren.

In het onderzoek van Locuniak en Jordan (2008) werd het getalbegrip van kinderen in kleuterklassen gebruikt om hun 'calculation fluency', vrij vertaald als rekenvloeiendheid, in groep 4 te voorspellen. Getalbegripmetingen verklaarden een groot deel (26 tot 41%) van de variantie boven de meer algemene voorspellers, zoals leeftijd, lezen en geheugen. Unieke voorspellende subgebieden waren een actief geheugen (werkgeheugen) voor getallen, kennis van getallen en getalcombinaties, waarbij de factor 'getalcombinaties' naar voren kwam als de sterkste voorspeller. De relatie tussen vroege rekenvaardigheden en latere rekenvloeiendheid heeft belangrijke gevolgen voor de screening en interventie van mogelijke rekenproblemen. Uit andere onderzoeken is gebleken dat zwak ontwikkeld getalbegrip bijdraagt aan rekenproblemen (Gersten, Jordan, & Flojo, 2005; Mazzocco & Thompson, 2005). Volgens Malofeeva, Day, Saco, Young, en Ciancio (2004) refereert getalbegrip in de meest brede zin van het woord naar het 'begrijpen van getallen' en de 'relaties tussen getallen'. Geoperationaliseerde definities bevatten de mogelijkheid om kleine hoeveelheden direct te herkennen zonder te tellen, om numerieke grootheden te vergelijken, te tellen en om simpele wiskundige berekeningen uit te voeren. Jordan, Kaplan, Olah, en Locuniak (2006) ontwikkelden een 'getalbegriptestbatterij' om kleuterklassen te screenen. Met deze

onderzoeksinstrumenten screenden zij kleuters met een verhoogd risico op rekenproblemen. De kern van deze testbatterij (Jordan, Kaplan, Locuniak, & Ramineni, 2007) bevat opgaven met tellen, kennis van getallen, nonverbale berekeningen, verhaaltjesproblemen en getalcombinaties. Jordan, Kaplan, Loconiak, en Ramineni (2007) vonden dat getalbegrip zowel aan het begin als aan het einde van de kleuterklas sterk samenhang met rekenprestaties in groep 3.

Onderzoeken van Deheane (1992) en Geary (1990) hebben verder aangetoond dat de meeste kinderen die problemen hebben met rekenen op de basisschool, al slecht presteerden tijdens voorbereidend rekenen in de kleuterklas. Onderzoek naar hoe kinderen in de leeftijd van 5 tot 7 jaar leren rekenen is belangrijk, want veel van deze kinderen vertonen gebreken in de vroege rekenvaardigheden (Schopman & Van Luit, 1996), die het verwerven van latere rekenvaardigheden kunnen belemmeren (Van Luit & Schopman, 2000).

Onderzoek toont verder aan dat rekenprestaties verbeteren dankzij het verzorgen van een effectieve basisinstructie aan alle leerlingen in de onderbouw om zodoende een goed fundament te leggen voor het begrijpen van rekenen en tevens vroegtijdige rekenproblemen te voorkomen (Clarke, Baker, & Chard, 2007). Het voorkomen van leerproblemen door middel van gerichte vroege interventie krijgt steeds meer aandacht in zowel het regulier als speciaal onderwijs (Fuson, Smith, & LoCicero, 1997; Griffin, 2004). Overeenkomstige bevindingen illustreren dat als rekenproblemen eenmaal zijn ontstaan het steeds moeilijker wordt deze problemen te herstellen naarmate ze langer onopgelost en onbehandeld blijven (Chard et al., 2008).

Geconcludeerd kan worden dat er tot op heden betrekkelijk weinig onderzoek is gedaan naar het vroegtijdig signaleren van rekenproblemen en naar mogelijke interventies voor jonge leerlingen (Gersten, Jordan, & Flojo, 2005). Daarnaast kan gesteld worden dat de vroege ontwikkeling van rekenvaardigheid een goede voorspeller is voor latere rekenvaardigheden (Arnold, Fisher, Doctoroff, & Dobbs, 2002) en het daarom van belang is om in een zo vroeg mogelijk stadium in te grijpen (Varol & Farran, 2006). Om te zorgen dat de basis goed gelegd kan worden, dienen problemen in de ontwikkeling van de voorwaarden daartoe vroeg opgespoord te worden om daarmee hopelijk een mogelijke rekenachterstand in de loop van de basisschool te voorkomen (Van Luit & Van de Rijt, 2009). Een vroegtijdige onderkenning van mogelijke leerproblemen vraagt dus om systematische signalering en toetsing, ook in de kleuterleeftijd (Ruijssenaars et al., 2006). De voorbereidende rekenvaardigheden vormen immers de basis voor het latere rekenen. Veel rekenproblemen op de basisschool blijken terug te voeren op onvoldoende ontwikkelde voorbereidende

rekenvaardigheden (Geary, 1990; Van Luit, 1987). Uit onderzoek is gebleken dat een rekeninterventie een positieve uitwerking heeft op de rekenvaardigheid van basisschoolleerlingen (Arnold, Fisher, Doctoroff, & Dobbs, 2002). Het is daarom begrijpelijk dat onderzoekers zich meer en meer richten op het identificeren van cruciale variabelen die voorspellen welke leerlingen het risico lopen later rekenproblemen te ontwikkelen (Chard et al., 2005; Clarke & Shinn, 2004) en op het voorkomen van deze moeilijkheden voordat ze hardnekkige problemen worden (Fuchs et al., 2005).

Het onderzoek

In het navolgende wordt nader ingegaan op de mogelijkheden om een achterstand in de ontwikkeling van voorbereidende rekenvaardigheid te signaleren. Dit wordt gedaan aan de hand van de Utrechtse Getalbegrip Toets-Revised (UGT-R, Van Luit & Van de Rijt, 2009). In dit onderzoek wordt nagegaan wat het verband is tussen voorbereidende rekenvaardigheden (gemeten met de UGT-R) van kinderen met minder goede rekenprestaties uit groep 1, 2 en 3, en hun latere rekenprestaties in groep 2, 3 en 4. Onder kinderen met minder goede rekenprestaties worden kinderen verstaan die behoren tot de groep van 25% laagst scorende kinderen op de UGT-R, afgenomen in groep 1, 2 en 3. Deze 25% laagst scorende kinderen kunnen worden verdeeld in E-kinderen; de kinderen die behoren tot de 10% laagst scorenden en de D-kinderen; de kinderen die behoren tot de 15% die daar net boven presteert.

Er wordt nagegaan op welke testonderdelen (voorbereidende rekenvaardigheden) van de UGT-R kinderen met minder goede rekenprestaties (scores op D/E niveau) uitvallen. Dit is van belang, omdat het rekenkundig inzicht zich al vroeg ontwikkelt (Geary, Hoard & Hamson, 1999) en om extra ondersteuning en interventie bij vroege rekenproblemen mogelijk te maken (Wilson & Swanson, 2001).

De vraag die centraal staat in dit onderzoek is of de UGT-R voldoende predictief valide is. In het bijzonder: Kunnen de voorbereidende rekenprestaties van leerlingen uit groep 2, 3 en 4 worden voorspeld op basis van hun eerder behaalde toetsscores op de UGT-R, afgenomen in groep 1, 2 en 3? Aan de hand van deze vraagstelling komen in dit onderzoek een aantal meer specifieke onderzoeksvragen aan de orde namelijk:

- Zijn de eerder behaalde UGT-R scores van kinderen uit groep 1 verschillend in vergelijking met de door hen behaalde UGT-R scores in groep 2.
- Zijn de eerder behaalde UGT-R scores van kinderen uit groep 2 verschillend in vergelijking met de door hen behaalde UGT-R scores in groep 3.

- Zijn de eerder behaalde UGT-R scores van kinderen uit groep 3 verschillend in vergelijking met de door hen behaalde UGT-R scores in groep 4.

Verwacht wordt dat de UGT-R predictieve waarde heeft voor latere rekenvaardigheid.

Methode

Participanten

Uit een aselechte steekproef van 2137 basisschoolkinderen, waarbij in het schooljaar 2007-2008 op twee momenten de Utrechtse Getalbegrip Toets-Revised ten behoeve van hernormeringsonderzoek is afgenomen, zijn 41 leerlingen uit drie jaargroepen (groep 2, 3 en 4) op basis van Cito niveauscores D en E geselecteerd om deel te nemen aan het voorliggende valideringsonderzoek.

Aan het eerder uitgevoerde normeringsonderzoek hebben in totaal 45 scholen, verspreid over twaalf provincies, deelgenomen. Praktische overwegingen hebben ertoe geleid dat voor het huidige onderzoek vijf scholen uit de provincie Utrecht zijn benaderd. Van de vijf scholen hebben drie reguliere basisscholen hun medewerking aan het valideringsonderzoek verleend. Deze basisscholen bevinden zich in Amersfoort, Zeist en Driebergen-Rijsenburg. De steekproeftrekking van dit onderzoek betreft dus een selecte, toevallige steekproef en kan vanwege de verdeling van de populatie in deelgroepen als quotasteekproef worden aangeduid. Het is derhalve niet mogelijk om vergaande generaliserende uitspraken over de onderzoeksuitkomsten te doen.

In dit huidige onderzoek zijn uitsluitend die kinderen geselecteerd uit de groepen 2, 3 en 4, die op meetmoment 2 van het hernormeringsonderzoek in mei/juni 2008 tot de 25 procent laagst scorende kinderen behoorden. Deze groep kinderen kan worden verdeeld in E-kinderen (kinderen die behoren tot de 10% laagst scorenden) en D-kinderen (kinderen die behoren tot de 15% die daar net boven presteert).

Wegens verhuizing, ziekte of andere redenen hebben een aantal leerlingen, die vorig schooljaar wel hebben meegewerkt aan het hernormeringsonderzoek, niet kunnen deelnemen aan het huidige onderzoek. Daarnaast zijn de leeftijdsgroepen van de UGT-R genormeerd tot en met 99 maanden. Op basis hiervan bleken uiteindelijk vier geselecteerde leerlingen te oud. Dit reduceert de onderzoekspopulatie (zie tabel 1) tot 34 leerlingen (18 jongens, 16 meisjes). Het onderzoek (3^e meetmoment) vond plaats in het voorjaar van 2009. De leeftijd van de kinderen varieerde op dat meetmoment van 5 jaar en 4 maanden tot 8 jaar en 2 maanden.

Tabel 1

Onderzoekspopulatie

| | Geslacht | | Leeftijd (in maanden) | |
|---------------|----------|--------|-----------------------|------|
| | Jongen | Meisje | M | SD |
| Groep 2 | 7 | 7 | 67,9 | 3,1 |
| Groep 3 | 5 | 6 | 83,1 | 2,9 |
| Groep 4 | 6 | 3 | 93,0 | 2,8 |
| Totaal (N=34) | 18 | 16 | 79,5 | 10,9 |

Meetinstrumenten

De Utrechtse Getalbegrip Toets-Revised (UGT-R) is een methodeonafhankelijke taakgerichte toets en is afgenomen om getalbegrip te meten. De toets bestaat uit negen verschillende onderdelen welke betrekking hebben op vergelijken, classificeren, correspondentie leggen, seriëren, telwoorden gebruiken, synchroon en verkort tellen, resultaatief tellen, toepassen van kennis van getallen en schatten. Scores op deze onderdelen geven een beeld van de sterke en zwakke kanten van het voorbereidende rekenvaardigheidniveau (niveau A t/m E, vergelijkbaar met de Cito vaardigheidsscores). Ook kan worden nagegaan welke onderdelen al wel of nog niet (voldoende) worden beheerst (Van Luit & Van de Rijt, 2009).

Tijdens de eerdere onderzoeksmomenten en het huidige onderzoek is gebruik gemaakt van een pilotversie van de UGT-R. Deze bestaat uit zes verschillende testboekjes, ieder voorzien van 45 items. Dit houdt in dat er per toetsonderdeel vijf items aan iedere leerling zijn aangeboden. Bij elk van de onderzochte kinderen is drie keer één van de zes versies van de toets afgenomen. De kinderen zijn toegewezen aan een bepaalde versie van de toets door per groep ongeveer twee jongens en twee meisjes dezelfde versie te laten maken. Uit onderzoek is gebleken dat de Cronbach's alpha en betrouwbaarheidscoëfficiënt van de gebruikte boekjes hoog is. De gemiddelde betrouwbaarheidscoëfficiënt is .94. Verschillen in getalbegrip kunnen derhalve betrouwbaar gemeten worden (Van Luit & Van de Rijt, 2009).

Procedure

De Utrechtse Getalbegrip Toets - Revised (Van Luit & Van de Rijt, 2009) meet het niveau van beheersing van getalbegrip van kinderen uit groep 1, 2 en 3 van het basisonderwijs. Voor het huidige valideringsonderzoek is gebruik gemaakt van zowel bestaande als nieuw verzamelde onderzoeksgegevens. Deze gegevens zijn op drie verschillende meetmomenten verzameld en kunnen ook worden gebruikt ten behoeve van een follow-up onderzoek. Het eerste meetmoment vond plaats in januari/februari 2008 (T1) en het tweede meetmoment in mei/juni 2008 (T2). Het derde meetmoment (T3) heeft in januari/februari 2009

plaatsgevonden, in het kader van het huidige onderzoek naar de predictieve validiteit van de UGT-R.

Bij de leerlingen die tijdens de eerste twee meetmomenten in groep 1, 2 en 3 de UGT-R is afgenomen, is in januari/februari 2009 bij kinderen die een D of E score behaalden op het tweede meetmoment, opnieuw de UGT-R individueel afgenomen. Hierbij is gebruik gemaakt van dezelfde versie van de werkboekjes waaruit deze leerlingen tijdens de eerder genoemde meetmomenten hebben gewerkt. Tijdens de afname werden de antwoorden genoteerd op het scoreformulier. Door de kinderen in een aparte ruimte te toetsen, is gestreefd naar een minimale invloed van omgevingsfactoren. Het aantal juist beantwoorde opgaven vormt de totaalscore op de toets, die maximaal 45 is. De afnameduur van de UGT-R bedraagt ongeveer 30 minuten per persoon.

Resultaten

De gemiddelde vaardigheidsscores op de UGT-R en de verschillen hiertussen over de drie meetmomenten genomen, zijn per basisschoolgroep weergegeven in de tabellen 2 en 3. Met behulp van de Paired-samples T-test is nagegaan of de UGT-R vaardigheidsscores van kinderen uit groep 1, 2 en 3 in 2008 en de door hen behaalde testcores een jaar later in groep 2, 3 en 4 verschillend zijn.

Tabel 2

Gemiddelde vaardigheidsscores op de UGT-R per groep op de drie verschillende meetmomenten

| | Voormeting (T1) | | | Nameting (T2) | | Uitgestelde nameting (T3) | |
|---------|-----------------|-------|------|---------------|------|---------------------------|------|
| | N | M | SD | M | SD | M | SD |
| Groep 1 | 14 | 40.62 | 6.19 | 47.00 | 7.65 | 55.34 | 5.75 |
| Groep 2 | 11 | 55.98 | 8.54 | 55.74 | 7.46 | 63.72 | 3.86 |
| Groep 3 | 9 | 62.40 | 6.90 | 64.57 | 4.31 | 67.32 | 7.61 |

Resultaten wijzen uit dat de gemiddelde vaardigheidsscores van leerlingen uit groep 1, 2 en 3 tussen T1 en T2, respectievelijk T3 toenemen, zie tabel 2. De enige uitzondering hierop vormt de gemiddelde vaardigheidsscore van de kinderen uit groep 2 op meetmoment 2 in vergelijking met de gemiddelde vaardigheidsscore van deze kinderen op meetmoment 1. In groep 1 zijn de gevonden toenames in vaardigheidsscore significant ($t(13) = 5.79, p = .000$ en $t(13) = 11.09, p = .000$ en $t(13) = 5.56, p = .000$), dit houdt in dat de scores over de meetmomenten genomen, toenemen (zie tabel 3). In groep 2 is het verschil in vaardigheids-

score tussen meetmoment 1 en 2 (T1-T2) niet significant ($t(10) = -.13, p = .903$), dit betekent dat de score, zoals verwacht, op meetmoment 2 niet verschilt van de score op meetmoment 1. De verschillen in vaardigheidsscore tussen meetmoment 1 en 3 (T1-T3) en tussen meetmoment 2 en 3 (T2-T3) zijn echter wel significant ($t(10) = 3.19, p = .010$ en $t(10) = 4.65, p = .001$), dit wil zeggen dat scores over meetmomenten toenemen. Binnen groep 3 zijn de vaardigheidsscores, zoals verwacht, op alle drie de meetmomenten niet verschillend ($t(8) = 1.26, p = .245$ en $t(8) = 2.25, p = .055$ en $t(8) = 1.88, p = .097$).

Tabel 3

Vershil in vaardigheidsscores op de UGT-R per groep op de drie verschillende meetmomenten

| | Vershil T1-T2 | | | | Vershil T1-T3 | | | Vershil T2-T3 | | |
|---------|---------------|------|----|------|---------------|----|------|---------------|----|------|
| | N | t | df | p | t | Df | P | t | Df | p |
| Groep 1 | 14 | 5.79 | 13 | .000 | 11.09 | 13 | .000 | 5.56 | 13 | .000 |
| Groep 2 | 11 | -.13 | 10 | .903 | 3.19 | 10 | .010 | 4.65 | 10 | .001 |
| Groep 3 | 9 | 1.26 | 8 | .245 | 2.25 | 8 | .055 | 1.88 | 8 | .097 |

Verschillen zijn significant bij $p < .05$

De Wilcoxon Signed-ranks-toets is uitgevoerd om te kunnen bepalen of de UGT-R vaardigheidsscores van kinderen uit groep 1, 2 en 3 in 2008 verschillen van de door hen behaalde vaardigheidsscores in groep 2, 3 en 4 van het schooljaar 2009, zie tabel 4. Deze Wilcoxon Signed-ranks-toets is bij kleine steekproeven, zoals ook in dit onderzoek het geval is, minder gevoelig voor extreme scores dan de t-toets.

Tabel 4

Vershil in vaardigheidsscores op de UGT-R per groep op de drie verschillende meetmomenten gemeten middels de Wilcoxon Signed-ranks-toets

| | Vershil T1-T2 | | | Vershil T1-T3 | | Vershil T2-T3 | |
|---------|---------------|-------|------|---------------|------|---------------|------|
| | N | Z | p | Z | p | Z | p |
| Groep 1 | 14 | -3.17 | .002 | -3.30 | .001 | -3.17 | .002 |
| Groep 2 | 11 | -.05 | .959 | -2.22 | .026 | -2.67 | .008 |
| Groep 3 | 9 | -1.13 | .260 | -1.84 | .066 | -1.60 | .110 |

Verschillen zijn significant bij $p < .05$

Er is binnen groep 1 een significant verschil ($Z = -3.17; p = .002$) gevonden tussen de vaardigheidsscores op het eerste (T1) en tweede meetmoment (T2). Van de veertien

onderzochte leerlingen was er één keer sprake van een lagere ruwe score op T2 dan op T1. Bij dertien leerlingen was er sprake van een hogere ruwe score op T2 dan op T1. Tussen de vaardigheidsscores van het eerste (T1) en het derde meetmoment (T3) blijkt eveneens een significant verschil ($Z = -3.30$; $p = .001$) te bestaan. Er blijken geen leerlingen een lagere ruwe score op T3 dan op T1 te halen; alle veertien leerlingen behalen een hogere ruwe score op T3 dan op T1. Van T2 naar T3 blijken de scores ook significant te verschillen ($Z = -3.17$; $p = .002$). Slechts één leerling behaalt op T3 een lagere ruwe score dan op T2, de resterende dertien leerlingen halen allen een hogere ruwe score op T3 vergeleken met T2.

Binnen groep 2 wordt geen significant gevonden tussen de vaardigheidsscores op T1 en T2 ($Z = -0.05$; $p = .959$). Van T1 naar T3 blijken de vaardigheidsscores wel significant te verschillen ($Z = -2.22$; $p = .026$). Twee van de elf leerlingen behalen een lagere ruwe score op T3 dan op T1; de overige negen leerlingen scoren hoger op T3 dan op T1. Ook de scores van T2 naar T3 verschillen significant ($Z = -2.67$; $p = .008$). Van de elf leerlingen blijken er twee een lagere ruwe score te hebben op T3 dan op T2 en de resterende negen leerlingen halen allemaal een hogere ruwe score op T3 dan op T2.

Er is binnen groep 3 geen significant verschil ($Z = -1.13$; $p = .260$) gevonden tussen de vaardigheidsscores van negen leerlingen op T1 en T2. Ook tussen T1 en T3 is geen significant verschil gevonden ($Z = -1.84$; $p = .066$). Ditzelfde is het geval tussen T2 en T3 ($Z = -1.60$; $p = .110$).

Wanneer de resultaten van de Paired-samples T-test worden vergeleken met de resultaten van de Wilcoxon Signed-ranks-toets is te zien dat deze resultaten overeenkomen. Gesteld kan worden dat binnen groep 1 en 2 sprake is van significante verschillen in vaardigheidsscores tussen meetmoment 2 en 3 en tussen meetmoment 1 en 3. Dit geldt evenwel niet voor groep 3. Er is wel een trend zichtbaar voor groep 3: de verschillen in scores tussen meetmoment 1 en 3 tendeert naar een verschil en gezien de geringe omvang van de steekproef kan dit als betekenisvol worden beschouwd.

De samenhang tussen de vaardigheidsscores is met behulp van de Spearman's rangcorrelatie onderzocht, zie tabel 5. Er blijkt in groep 1 een zeer sterke samenhang te bestaan tussen de vaardigheidsscores op T1 en de scores op T2 ($r_s = .96$; $p < .001$, tweezijdig). De samenhang tussen de vaardigheidsscores op T2 en op T3 is matig, maar deze correlatiecoëfficiënt is wel significant bij een alpha van 0.05 ($r_s = .54$; $p = .047$, tweezijdig). Hieruit kan worden opgemaakt dat de score op meetmoment 2 voorspeld kan worden aan de hand van de behaalde score op meetmoment 1 (T1-T2). De score op meetmoment 3 kan deels worden verklaard door de behaalde score op meetmoment 2 (T2-T3).

Tabel 5

Correlatie tussen de vaardigheidsscores op de UGT-R binnen de verschillende groepen op de drie meetmomenten

| | N | Correlatie T1-T2 | Correlatie T1-T3 | Correlatie T2-T3 |
|---------|----|---------------------|---------------------|---------------------|
| Groep 1 | 14 | .96** | .52 | .54* |
| Groep 2 | 11 | .76** | .54 | .80** |
| Groep 3 | 9 | .47 | .47 | .82** |

** Correlatie is significant bij een alpha van 0.01

* Correlatie is significant bij een alpha van 0.05

Binnen groep 2 blijkt, net als in groep 1, een sterke samenhang te bestaan tussen de vaardigheidsscores op T1 en T2 ($r_s = .76$; $p = .007$, tweezijdig). De samenhang op T2 en T3 is eveneens sterk ($r_s = .80$; $p = .003$, tweezijdig). De vaardigheidsscores binnen groep 3 blijken voor T1 naar T2 slechts matig te correleren. Bij de overgang naar de volgende jaargroep (T2-T3) blijkt echter wel weer sprake van een sterke correlatie te zijn tussen de vaardigheidsscores ($r_s = .82$; $p = .007$, tweezijdig).

Conclusie en discussie

Middels statistische analyses is onderzocht of de rekenprestaties van leerlingen uit groep 2, 3 en 4 voorspeld kunnen worden op basis van hun eerdere behaalde toetsscores op de UGT-R in de groepen 1, 2 en 3. Met behulp van de Paired-samples T-test en de Wilcoxon Signed-ranks-toets is onderzocht of de behaalde vaardigheidsscores op de UGT-R van de kinderen uit groep 1, 2 en 3 in 2008 significant verschillen van de door hen behaalde test scores een jaar later in groep 2, 3 en 4.

Resultaten wijzen uit dat de UGT-R vaardigheidsscores van de leerlingen uit groep 1 significant verschillen van de door hen behaalde scores in groep 2. In tegenstelling tot de verwachting met oog op de predictieve validiteit van de UGT-R dat deze scores gelijk zouden blijven, kan worden gesteld dat deze scores van elkaar verschillen. De vaardigheidsscores van deze leerlingen zijn, over de drie meetmomenten gezien, toegenomen.

Uit beide uitgevoerde testen blijkt dat de vaardigheidsscores van de leerlingen in groep 2 op het derde meetmoment hoger zijn dan op het tweede meetmoment (T2-T3). De scores op meetmoment 3 zijn ook hoger dan op meetmoment 1 (T1-T3). De vaardigheidsscores op het eerste en het tweede meetmoment verschillen echter niet (T1-T2).

In de vaardigheidsscores van de leerlingen uit groep 3 zijn geen verschillen gevonden tussen het eerste en tweede meetmoment (T1-T2) en tussen meetmoment 2 en 3 (T2-T3).

Tussen meetmoment 1 en 3 (T1-T3) is wel een trend zichtbaar: de scores nemen in deze tijdspanne wel toe, maar deze verschillen zijn vanwege de geringe omvang van de steekproef niet significant.

Er worden grote individuele verschillen gevonden in de ontwikkeling van het getalbegrip gedurende de kleutertijd. Daarnaast kunnen kinderen in verschillende mate profiteren van het informele rekenonderwijs dat geboden wordt (Van de Rijt, 1996; Van Luit, Van de Rijt, & Pennings, 1994). Deze factoren kunnen binnen het onderzoek hebben geleid tot een significante toename, van de (gemiddeld genomen) vaardigheidsscores in een jaar tijd voor de leerjaargroep 1.

Een mogelijke verklaring voor de toename in vaardigheidsscores van de leerlingen in groep 2 is dat de kinderen mogelijk geprofiteerd hebben van het formele geboden rekenwiskundeonderwijs begin groep 3. In groep 3 krijgen de leerlingen meer gestructureerde rekeninstructies, waardoor zowel het getalbegrip als de rekenvaardigheden extra worden beïnvloed (Torbeyns et al., 2002). Hierbij dient te worden opgemerkt dat de kinderen in groep 3 geen significante toename in vaardigheidsscore laten zien. De resultaten van deze leerlingen laten echter wel een trend in vergelijkbare richting zien.

Binnen dit onderzoek is daarnaast de samenhang van de vaardigheidsscores door middel van de Spearman's rangcorrelatie onderzocht. Verwacht werd dat de UGT-R predictieve waarde heeft voor latere rekenvaardigheid. Hiermee is tevens verondersteld dat de prestaties van de rekenzwakke kinderen op de eerste twee meetmomenten ook op het derde meetmoment van zwak niveau zullen zijn. Uit het onderzoek blijkt dat de samenhang van de UGT-R vaardigheidsscores voor groep 1 tussen meetmoment 1 en 2 zeer sterk is (T1-T2) en matig is tussen meetmoment 2 en 3 (T2-T3). Hieruit kan worden opgemaakt dat de score op meetmoment 2 voorspeld kan worden aan de hand van de behaalde score op meetmoment 1 (T1-T2). Het verschil in tijd tussen het eerste en tweede meetmoment is om en nabij 4 maanden. De UGT-R blijkt op basis van dit onderzoek een goed instrument om resultaten van kinderen van groep 1 binnen de jaargroep op korte termijn te voorspellen. De score op meetmoment 3 kan slechts deels worden verklaard door de behaalde score op meetmoment 2 (T2-T3), maar er zijn ook nog andere factoren die mogelijk invloed hebben. Hierbij kan worden gedacht aan de invloed van het informele rekenonderwijs, kind- en omgevingsfactoren. Uit het onderzoek komt naar voren dat rekenprestaties op lange termijn middels de UGT-R relatief goed voorspeld kunnen worden. Bij de overgang van leerlingen uit groep 1 naar de volgende jaargroep neemt de voorspellende waarde iets af. In gedachten moet

worden gehouden dat een correlatie van .50 in absolute zin matig is, maar op basis van de zeer kleine steekproef (veertien leerlingen) als relatief sterk beschouwd kan worden.

Binnen groep 2 blijkt, net als in groep 1, een sterke samenhang te bestaan tussen de vaardigheidsscores op meetmoment 1 en meetmoment 2 (T1-T2). De samenhang tussen de scores op meetmoment 2 en meetmoment 3 is eveneens sterk (T2-T3). Op grond van deze onderzoeksresultaten kan worden geconcludeerd dat de UGT-R ook voor deze jaargroep een goed instrument blijkt te zijn om voorspellingen op korte termijn te doen. Ook op de lange termijn blijkt de UGT-R een goed middel om de prestaties van kinderen, bij de overgang van groep 2 naar de volgende jaargroep, te voorspellen.

De vaardigheidsscores blijken in groep 3 tussen meetmoment 1 en 2 (T1-T2) zwak samen te hangen, wat betekent dat de scores op meetmoment 1 in geringe mate de scores op meetmoment 2 voorspellen. In een onderzoeksgroep van slechts negen leerlingen is de gevoeligheid voor enkele uitschieters zeer groot. Om deze reden kan een samenhang van .47 als relatief groot worden gezien. Bij de overgang naar de volgende jaargroep blijkt de samenhang zowel in absolute als relatieve zin sterk te zijn (T2-T3). Een verklaring hiervoor zou kunnen zijn dat rekenzwakke kinderen in een jaar tijd (van midden groep 3 tot halverwege groep 4), waarin ze formeel rekenwiskundeonderwijs ontvangen en mogelijk expliciet voorbereidende rekenvaardigheden aangeleerd krijgen, veel bijleren op het gebied van voorbereidend rekenen. Dit komt overeen met eerder verricht onderzoek waarin werd aangetoond dat rekenprestaties verbeteren dankzij het verzorgen van een effectieve basisinstructie aan leerlingen in de onderbouw om zodoende een goed fundament te leggen voor het begrijpen van rekenen en tevens vroegtijdige rekenproblemen te voorkomen dan wel te verminderen (Clarke, Baker, & Chard, 2007).

Opvallend is dat voor alle drie groepen de samenhang tussen de scores op meetmoment 1 en 3 (T1-T3) gering is, dit wil zeggen dat de score van een kind op meetmoment 1 in mindere mate voorspellende waarde heeft voor de score op meetmoment 3. Dit zou verklaard kunnen worden door het feit dat tussen meetmoment 1 en 3 de meeste (onderwijs)tijd zit. Hoogstwaarschijnlijk ligt de verklaring echter in een van de beperkingen van dit onderzoek, namelijk dat er sprake is van een zeer kleine onderzoeksgroep. Hierdoor is interpretatie van de resultaten lastig en is grote voorzichtigheid geboden.

Dit onderzoek wijst uit dat de UGT-R vaardigheidsscores positief met elkaar samenhangen. Dit wil zeggen dat een bepaalde vaardigheidsscore zal leiden tot een overeenkomstige score op een volgend meetmoment; een vaardigheidsscore is op basis van een eerder behaalde score dus op deze manier te voorspellen. Op basis van de

onderzoeksresultaten blijkt de UGT-R een instrument te zijn dat over het algemeen redelijk in staat is om voorspellende uitspraken te doen over latere rekenprestaties. De predictieve validiteit van de UGT-R is zodoende voldoende tot ruim voldoende te noemen. Dit houdt in dat de UGT-R een betrouwbaar en valide instrument is waarmee kinderen gesignaleerd kunnen worden met een verhoogde kans op het vertonen van uitval in rekenprestaties of zelfs van het ontwikkelen van een rekenprobleem dan wel rekenstoornis. De (reken)ontwikkeling van kinderen die in groep 1, 2 en 3 lage scores op de UGT-R behalen, moet daarom bijzonder goed worden gevolgd. Uit literatuur (Arnold et al., 2002; Ruijssenaars et al., 2006; Van Luit & Van de Rijt, 2009) blijkt dat middels vroegtijdige screening al vroeg in de basisschoolperiode kan worden geïntervenieerd, zodat de mogelijke ontwikkeling van rekenproblemen voorkomen dan wel verminderd kan worden.

Binnen dit onderzoek zijn een aantal beperkingen te noemen waardoor voorzichtigheid geboden is bij de interpretatie en generalisatie van de gevonden resultaten. Om mee te beginnen is er in het huidige onderzoek sprake geweest van een selecte steekproef. In acht moet worden genomen dat de conclusies die getrokken zijn ten aanzien van de predictieve validiteit van de UGT-R vooralsnog alleen betrekking hebben op de populatie van het huidige onderzoek. De conclusies zijn gesteld op basis van leerlingen uit drie verschillende groepen van drie basisscholen en kunnen om deze reden niet gegeneraliseerd worden naar de Nederlandse populatie. Een andere beperking die hierbij moet worden vermeld is dat aan het huidige onderzoek veel leerlingen hebben deelgenomen, die tijdens het derde meetmoment bijna de maximale leeftijd (7.6 jaar) bereikten. Volgens Van Luit en Van de Rijt (2009) is, gezien het doel van de toets: het meten van voorbereidende rekenvaardigheid, een oudere leeftijdsrange (>7.6 jaar) niet noodzakelijk en zelfs niet mogelijk, vanwege overduidelijk plafondeffecten en gebrek aan differentiatie bij oudere leerlingen. Mogelijk speelt bij een aantal leerlingen, die ten tijde van het derde meetmoment in groep 4 zaten, dit zogenoemde plafondeffect een rol. Een aanbeveling met het oog op de generaliseerbaarheid van de bevindingen is om bij vervolgonderzoek naar de predictieve validiteit van de UGT-R, een groter aantal scholen, die in het verleden deel hebben genomen aan het hernormeringsonderzoek, te laten participeren. Deze scholen zijn afkomstig uit verschillende provincies en bevinden zich in zowel grote, middelgrote als kleine plaatsen in Nederland. Dit komt de generaliseerbaarheid van de bevindingen ten goede.

Een volgende beperking die hierop nauw aansluit is dat het huidige onderzoek een zeer kleine steekproef betrof, met als gevolg dat het niet mogelijk was een aantal wenselijke analyses uit te voeren ten aanzien van de ontwikkeling van rekenvaardigheden. Een grotere

steekproef had de waarschijnlijkheid en generaliseerbaarheid van de gevonden onderzoeksresultaten verhoogd en de gevonden onderzoeksuitkomsten sterker kunnen onderbouwen. Naast een grotere steekproef was het voor het huidige onderzoek van meerwaarde geweest als naast de UGT-R ook de Cito-toetsen van de desbetreffende leerjaren waren afgenomen. Deze toetsen hebben een controlefunctie en verhogen, in geval van overeenstemming tussen beide onderzoeksresultaten, de waarschijnlijkheid van de gevonden onderzoeksuitkomsten op de UGT-R. Een positief punt aan dit onderzoek (naast al deze beperkingen) is dat er tussen meetmoment 1 en 2 én meetmoment 3 geen onderzoekersbias kan hebben plaatsgevonden. De onderzoeker op meetmoment 3 was namelijk niet dezelfde persoon als de onderzoeker op meetmoment 1 en 2.

Aanbeveling is dan ook om bovengenoemde beperkingen van dit onderzoek mee te nemen. Een volgende aanbeveling komt voort uit nieuwsgierigheid en interesse van de onderzoekster en betreft vervolgonderzoek verderop in de schoolperiode (bijvoorbeeld in de bovenbouw) om het lange termijneffect van de onderzoeksresultaten na te kunnen gaan. Bovengenoemde beperkingen en aanbevelingen kunnen wellicht in grootschaliger onderzoek naar de predictieve validiteit meegenomen worden.

Literatuur

- Arnold, D. H., Fisher, P. H., Doctoroff, G. L., & Dobbs, J. (2002). Accelerating math development in head start classrooms. *Journal of Educational Psychology, 94*, 762-770.
- Aubrey, C. (1993). An investigation of the mathematical knowledge and competencies which young children bring into school. *British Educational Research Journal, 19*, 27-41.
- Aubrey, C. (1994). An investigation of children's knowledge of mathematics at school entry and the knowledge their teaching and learning mathematics, about young learners and mathematical subject knowledge. *British Educational Research Journal, 20*, 105-120.
- Aubrey C., & Godfrey, R. (2003). The development of children's early numeracy through key stage 1. *British Educational Research Journal, 29*, 821-839.
- Aunio, P., Hautamäki, J., Heiskari, P., & Van Luit, J. E. H. (2004). Young children's number sense in Finland, Hong Kong and Singapore. *International Journal of Early Years Education, 12*, 195-216.
- Aunio, P., Hautamäki, J., Heiskari, P., & Van Luit, J. E. H. (2006). The early numeracy test in Finnish: Children's norms. *Scandinavian Journal of Psychology, 47*, 369-378.

- Aunola, K., Leskinen, E., Lerkkanen, M. K., & Nurmi, J. E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*, 699-713.
- Baltussen, M. (2000). Stagnaties in beginnend rekenen. In G. M. Van der Aalsvoort & A. J. J. M. Ruijsenaars (Red.), *Jonge risicokinderen. Achtergronden, onderkenning, aanpak en praktijk* (pp. 141-164). Rotterdam: Lemniscaat.
- Baroody, A. J. (1992). Remediating common counting difficulties. In J. Bideaud, C. Meljac, & J. – P. Fischer (Eds.), *Pathways to number. Children's developing numerical abilities* (307-324). Hillsdale, NJ: Erlbaum.
- Bosker, R. (2002). Elke leerling telt. In Onderwijsraad (Red.), *Over leerlinggewichten en schoolgewichten* (pp. 1-55). Den Haag: Onderwijsraad.
- Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37*, 231–242.
- Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., & Weisbeck, L. (1993). Models of problem solving: A study of kindergarten children's problem-solving processes. *Journal for Research in Mathematics Education, 24*, 428-441.
- Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten mathematics curriculum. *Learning Disability Quarterly, 31*, 11-20.
- Chard, D. J., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention, 30*, 3-14.
- Clarke, B., Baker S. K., & Chard, D.J. (2007). Measuring number sense development in young children: A summary of early research. *Leadership to Math Success for All, 5*, 1-11.
- Clarke, B., & Shinn, M. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33*, 234-248.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.
- Fuchs, L. S., Compton, D. L., Fuchs, D., Paulson, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97*, 493-513.

- Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.
- Fuson, K. C., Smith, S. T., & LoCicero, A. (1997). Supporting Latino first graders' ten-structured thinking in urban classrooms. *Journal for Research in Mathematics Education*, 28, 738-760.
- Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychologist*, 49, 363-383.
- Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.
- Geary, D. C. (1995). Reflections of evolution and culture in children's cognition. *American Psychologist*, 50, 24-37.
- Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology*, 74, 213-239.
- Gersten, R., Jordan, N., & Flojo, J. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, 38, 293-304.
- Griffin, S. (2004). Number worlds: A research-based mathematics program for young children. In D. H. Clements & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 325-340). Mahwah, NJ: Erlbaum.
- Jordan, N. C., Kaplan, D., Olah, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77, 153-175.
- Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities: Research and Practice*, 21, 37-47.
- Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities*, 41, 451-459.
- Malofeeva, E., Day, J., Saco, X., Young, L., & Ciancio, D. (2004). Construction and evaluation of a number sense test with head start children. *Journal of Educational Psychology*, 96, 648-659.
- Mazzocco, M. M., & Thompsons, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research and Practice*, 20, 142-155.

- Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64*, 95-114.
- Neuman, S. B., & Roskos, K. (2005). The state of state pre-kindergarten standards. *Early Childhood Research Quarterly, 20*, 125-145.
- Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*, 348-367.
- Piaget, J. (1965). *The child's concept of number*. Norwich, UK: Fletcher & Son.
- Rousselle, L., & Noe, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition, 102*, 361–395.
- Ruijsenaars, A. J. J. M., Van Luit, J. E. H., & Van Lieshout, E. C. D. M. (2006). *Rekenproblemen en dyscalculie. Theorie, onderzoek, diagnostiek en behandeling*. Rotterdam: Lemniscaat.
- Schopman, E. A. M., & Van Luit, J. E. H. (1996). Learning and transfer of preparatory arithmetic strategies among young children with a developmental lag. *Journal of Cognitive Education, 5*, 117-131.
- Sophian, C. (1987). Early developments in children's use of counting to solve quantitative problems. *Cognition and Instruction, 4*, 61-90.
- Sophian, C. (1992). Learning about numbers: Lessons for mathematics education from preschool number development. In J. Bideaud, C. Meljac, & J.-P. Fischer (Eds.), *Pathways to number. Children's developing numerical abilities* (pp. 19-40). Hillsdale, NJ: Erlbaum.
- Sophian, C. (2004). Mathematics for the future: Developing a Head Start curriculum to support mathematics learning. *Early Childhood Research Quarterly, 19*, 59–81.
- Torbeyns, J., Van de Noortgate, W., Ghesquiere, P., Verschaffel, L., Van de Rijt, B. A. M., & Van Luit, J. E. H. (2002). The development of early mathematical competence of 5- to 7- year-old children: A comparison between Flanders and The Netherlands. *Educational Research and Evaluation, 8*, 249-275.
- Torbeyns, J., Van de Rijt, B. A. M., Van den Noortgate, W., Van Luit, J. E. H., Ghesquiere, P., & Verschaffel, L. (2000). Ontwikkeling van getalbegrip bij vijf- tot zevenjarigen.

- Een vergelijking tussen Vlaanderen en Nederland. *Tijdschrift voor Orthopedagogiek*, 39, 118-131.
- Van de Rijt, B. A. M., (1996). *Vorbereidende rekenvaardigheid bij kleuters*. Doetinchem: Graviant.
- Van de Rijt, B. A. M., & Van Luit, J. E. H. (1998). Effectiveness of the Additional Early Mathematics program for teaching children early mathematics. *Instructional Science*, 26, 337-358.
- Van de Rijt, B. A. M., & Van Luit, J. E. H. (1999). Milestones in the development of infant numeracy. *Scandinavian Journal of Psychology*, 40, 65-71.
- Van de Rijt, B. A. M., Van Luit, J. E. H., & Pennings, A. H. (1999). The construction of the Utrecht Early Mathematical Competence Scales. *Educational and Psychological Measurement*, 59, 289-309.
- Van Luit, J. E. H. (1987). *Rekenproblemen in het speciaal onderwijs*. Nijmegen: Katholieke Universiteit (dissertatie).
- Van Luit, J. E. H., & Schopman, E. A. M. (2000). Improving early numeracy of young children with special educational needs. *Remedial and Special Education*, 21, 27-40.
- Van Luit, J. E. H., & Van de Rijt, B. A. M. (2005). *Utrechtse Getalbegrip Toets (3^e herziene druk)*. Doetinchem: Graviant.
- Van Luit, J. E. H., & Van de Rijt, B. A. M. (2009). *Utrechtse Getalbegrip Toets-Revised*. Doetinchem: Graviant.
- Van Luit, J. E. H., Van de Rijt, B. A. M., & Pennings, A. H. (1994). *Utrechtse Getalbegrip Toets*. Doetinchem: Graviant.
- Varol, F., & Farran, D. C. (2006). Early mathematical growth: How to support young children's mathematical development. *Early Childhood Education Journal*, 33, 381-387.
- Wilson, K. M., & Swanson, H. L. (2001). Are mathematics disabilities due to a domain-general or a domain-specific working memory deficit? *Journal of Learning Disabilities*, 34, 237-248.
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24, 220-251.