

**Twee gradaties van *cognitive load* in een educatief
computerspel: de prestaties van zwakke rekenaars
op getalbegrip**

Anneke Grootenhuis (4002881)

Bachelor thesis

Universiteit Utrecht

7496 woorden

Abstract

In deze studie is het effect van een educatief computerspel onderzocht bij zwakke en niet-zwakke rekenaars in twee condities. Uit eerder onderzoek blijkt dat educatieve computerspellen een positief effect hebben op rekenonderwijs. Er dient echter wel rekening te worden gehouden met overbelasting van het werkgeheugen. Zwakke rekenaars hebben namelijk een verminderde werkgeheugencapaciteit. In conditie 2 was sprake van veel entertainende elementen, terwijl deze elementen in conditie 1 afwezig waren. Daarnaast blijkt uit eerder onderzoek dat leerlingen met een verminderde werkgeheugencapaciteit een langere reactietijd hebben. Daarom is onderzocht of de speeltijd van zwakke rekenaars in conditie 2 langer is dan in conditie 1. Tot slot is onderzocht of de speeltijden van de verschillende spellen verschillen en worden naar aanleiding hiervan suggesties voor vervolgonderzoek gegeven. Uit de resultaten blijkt dat het spel een significant positief effect op zwakke en niet-zwakke rekenaars, maar dat hierbij geen onderscheid in condities is. Ook blijkt dat zwakke rekenaars geen langere speeltijd hebben in conditie 2. Tot slot blijkt dat er significante verschillen bestaan tussen speeltijden van de verschillende spellen.

Introductie

Tegenwoordig richten computerspelontwikkelaars zich niet enkel meer op computerspellen die alleen bedoeld zijn voor vermaak. Met de steeds verder vorderende integratie van ICT in het onderwijs hebben scholen de mogelijkheid om kinderen te laten leren met behulp van ICT (Tobias & Fletcher, 2011). Er worden dan ook computerspellen ontwikkeld voor onderwijssituaties waarbij het doel zowel onderwijzen als vermaken is (Itō, 2009).

Educatieve computerspellen worden in het onderwijs ingezet als ondersteuning, uitbreiding of vervanging van de lesmethode. Volgens een meta-analyse van Wouters, Van Nimwegen, Van Oostendorp en Van der Spek (2013) zijn *serious games* op het gebied van kennis en cognitieve vaardigheden meer effectief dan conventionele instructiemethodes. Uit deze meta-analyse bleek dat door *serious games* meer werd geleerd en kennis langere tijd werd onthouden dan bij conventionele instructiemethodes. Zij geven overigens wel aan dat het niet duidelijk is welke instructie- en context-factoren een bijdrage leveren aan dit positieve effect. Echter, uit een literatuurstudie van Mitchell en Savill-Smith (2004) blijkt dat deze positieve resultaten bij een beperkt aantal vakken naar voren komen. Wiskunde en taal zouden zich het beste lenen voor educatieve computerspellen, in tegenstelling tot zaakvakken. Het is niet bekend waarom deze vakken meer geschikt zijn voor educatieve computerspellen. Een grootschalige casestudie met meer dan 10.000 leerlingen van Wastiau, Kearney en Van den Berghe (2009) bevestigt deze positieve resultaten voor wiskunde en taal. Daarnaast laten de testcores op wiskunde en taal uit deze studie zien dat leerlingen meer profiteerden van educatieve computerspellen dan van klassikale lesmethoden.

Er zijn echter ook elementen in educatieve computerspellen die een negatief effect kunnen hebben op leerresultaten. Volgens de *multimedia learning theory* (MLT) van Mayer en Moreno (2003) kan bij de aanbidding van instructie *cognitive overload* ontstaan. In het werkgeheugen worden drie informatieverwerkingsprocessen onderscheiden: essentiële

verwerking, incidentele verwerking en vasthouden van representaties. Bij de essentiële verwerking wordt betekenis gegeven aan de aangeboden informatie door het selecteren, organiseren en integreren van woorden, en het selecteren, organiseren en integreren van afbeeldingen. Incidentele verwerking betreft het verwerken van niet-essentiële delen van de aangeboden instructie. Vasthouden van representaties is het vasthouden van verbale of visuele representaties in het werkgeheugen. Het werkgeheugen heeft twee kanalen voor het verwerken van informatie: een verbaal en een visueel kanaal. Deze kanalen hebben beperkte capaciteit en leren vereist actieve informatieverwerking in beide kanalen. In combinatie met de drie manieren van informatieverwerking geeft dit vijf mogelijke oorzaken voor het ontstaan van *cognitive overload*: 1) overbelasting van één kanaal, 2) overbelasting van beide kanalen, 3) overbodig materiaal in de instructie, 4) onjuiste manier van materiaalpresentatie en 5) te veel informatie in het werkgeheugen moeten vasthouden (Mayer & Moreno, 2003).

In het onderwijs moet bij het ontwerpen van instructie rekening worden gehouden met het mogelijk optreden van *cognitive overload* (Sweller, 2011). Door tegelijkertijd aandacht te besteden aan de manier waarop informatie gepresenteerd wordt en aan de cognitieve structuren die het de lerende mogelijk maken deze informatie te verwerken, kan *evidence-based* instructie worden ontworpen (Paas, Renkl, & Sweller, 2003). Richtlijnen hierbij zijn dat de verbale en visuele presentatie van informatie in evenwicht zijn; de informatie in delen wordt aangeboden; interessant, maar overbodig en afleidend materiaal niet aanwezig is; aanwijzingen worden gegeven voor het verwerken van informatie; en verbale en visuele informatie op dezelfde tijd en plaats wordt aangeboden. Voor dit laatste moet de lerende informatie kunnen vasthouden in het werkgeheugen (Mayer & Moreno, 2003).

Een studie van Raghobar, Barnes en Hecht (2010) toont aan dat de capaciteit van het werkgeheugen gerelateerd is aan rekenprestaties. Een meta-analyse van Friso-Van den Bos, Van der Ven, Kroesbergen en Van Luit (2013) laat zien dat deze relatie tevens voor

basisschoolkinderen geldt. De prestatie op eenvoudige rekentaken hangt sterk samen met het maximaal aantal eenheden dat in het werkgeheugen kan worden vastgehouden (Klein & Bisanz, 2000). Onderzoek van Bull en Scerif (2001) toont aan dat bij kinderen het vermogen om overbodige informatie uit het werkgeheugen te weren tijdens het uitvoeren van een taak samenhangt met de werkgeheugencapaciteit. Kinderen met rekenproblemen scoren lager en maken meer fouten bij werkgeheugentaken waarbij blokkering van irrelevante informatie nodig is (Passolunghi & Siegel, 2001; Rasmussen & Bisanz, 2005). Werkgeheugencapaciteit helpt bij het reguleren van aandacht, het weerstaan van afleiding en het eenvoudiger verkennen van het probleem in probleemoplossingstaken (Wiley & Jarosz, 2012). Een verminderde werkgeheugencapaciteit leidt tot een toename in reactietijd en fouten (Ashcraft & Kirk, 2001). Het werkgeheugen kan statistische significante variantie in het getal-schrijven, schatten en één-cijferig rekenen verklaren (Simmons, Willis, & Adams, 2012). Uit onderzoek van Bull en Scerif (2001) en Passolunghi en Siegel (2004) blijkt dat een verminderde werkgeheugencapaciteit bij kinderen gecorreleerd is aan een verminderd rekenvermogen. Andersom vonden Kytälä, Aunio en Hautamäki (2010) dat kinderen met vroege rekenproblemen, voordat zij formeel rekenonderwijs ontvangen, slecht presteren op verbale en visueel-ruimtelijke werkgeheugentaken. De verschillende cognitieve routes dragen onafhankelijk van elkaar bij aan vroege rekenvaardigheden in kleuterschool en kinderopvang en zijn verschillend gerelateerd aan rekenprestaties twee jaar later (LeFevre et al., 2010). Het werkgeheugen voorspelt de voortgang van kinderen ten opzichte van vroege leerdoelen aan het begin van formeel onderwijs (Alloway et al., 2005). Friso-Van den Bos (2014) vond dat het werkgeheugen een voorspeller is van rekenvaardigheden en tevens van getalbegrip.

Reeds in het kleuteronderwijs wordt gestart met het bevorderen van getalbegrip (Ruijsenaars, Van Luit, & Van Lieshout, 2004). Stichting Leerplan Ontwikkeling (SLO) heeft een leerlijn ontwikkeld, bestaande uit de gebieden getallen, meten en meetkunde, ter

bevordering van getalbegrip bij kleuters (SLO, 2011). Met getalbegrip wordt het effectief kunnen omgaan met getallen en aantallen bedoeld, uitgewerkt in het voldoen aan zeven rekenvoorwaarden, te weten conservatie, correspondentie, classificatie, seriatie, tellen, kennis van rekentaal en maatbegrip (Ruijsenaars et al., 2004). Van de Rijt (1996) heeft deze zeven voorwaarden uitgewerkt in acht componenten van het voorbereidend rekenen, te weten vergelijken, classificeren, correspondentie leggen, ordenen, gebruiken van telwoorden, gestructureerd tellen, resultaatief tellen en toepassen van algemene kennis van getallen. Het beheersen van deze componenten lijkt essentieel om een goede rekenstart in de eerste klas van het basisonderwijs te kunnen maken. Het blijkt namelijk uit onderzoek dat er een positieve sterke correlatie is tussen de voorbereidende rekenprestaties aan het begin van de kleuterklas en het einde van de eerste klas van de basisschool (Jordan, Kaplan, Locuniak, & Ramineni, 2007). Met name telvaardigheden in de kleutertijd hebben een voorspellende waarde: Aunola, Leskinen, Lerkkanen en Nurmi (2004) en Passolunghi, Vercelloni en Schadee (2007) vonden dat telvaardigheden in de kleutertijd een significante voorspellende waarde hebben op de rekenprestaties in de eerste en tweede klas van het basisonderwijs. Ook de vaardigheid met betrekking tot het koppelen tussen het cijfersymbool, het cijferwoord en de correcte hoeveelheid blijkt een significante voorspeller te zijn van rekenprestaties in de tweede klas van het basisonderwijs (De Smedt, Verschaffel, & Ghesquière, 2009). Goed getalbegrip in de kleutertijd blijkt zelfs van belang voor rekenprestaties op latere leeftijd: Kavkler, Tancig en Magajna (2003) vonden een sterk verband tussen de score op de Utrechtse Getalbegrip Toets en rekenprestaties in de vierde klas van het basisonderwijs. Uit onderzoek van Siegler (2009) blijkt zelfs dat het niveau van voorbereidend rekenen in de kleutertijd een zeer groot deel van de verdere rekenprestaties in het basis- en voortgezet onderwijs bepaalt.

Gedurende de kleuterjaren ontstaan grote verschillen tussen kinderen wat betreft voorbereidende rekenvaardigheden. Het ontwikkelen van voorbereidende rekenvaardigheden

begint namelijk al in de thuissituatie (Ginsburg, Lee, & Boyd, 2008). Onderzoek heeft uitgewezen dat deze vorm van informeel leren van groot belang is voor de cognitieve ontwikkeling van kinderen (Belsky et al., 2007). Anders et al. (2011) en Melhuish et al. (2008) vonden dat met name sociaal-economische status en het stimuleren van getalbegrip middels oefeningen met getallen en het aanleren van versjes en rijmpjes met getallen door de ouders, van invloed is op het niveau van voorbereidende rekenvaardigheden. Wanneer kinderen reeds met achterstanden het kleuteronderwijs instromen, is het noodzakelijk dat deze achterstand aangepakt wordt om met voldoende niveau het formele voorbereidend rekenonderwijs te volgen. Lage prestaties op voorbereidend rekenen kunnen immers grote gevolgen hebben voor rekenprestaties op latere leeftijd.

Getalbegrip bij voorbereidende rekenaars kan worden gemeten aan de hand van de Utrechtse Getalbegrip Toets-Revised (UGT-R) (Van Luit & Van de Rijt, 2009a). Deze toets, bestaande uit 45 items, meet de vaardigheden van kleuters op de gebieden vergelijken, classificeren, corresponderen, ordenen, gebruik van telwoorden, synchroon en verkort tellen, resultatief tellen, toepassen van kennis van getallen en schatten. Per drie maanden gelden andere normen, aansluitend bij het ontwikkelingsniveau. De UGT-R sluit aan op de tussendoelen omgaan met de telrij, hoeveelheden en getallen in het domein 'getallen' van voorbereidend rekenen zoals beschreven door Noteboom en Klep (2005) en SLO en Freudenthal Instituut (Treffers, Van den Heuvel-Panhuizen, & Buijs, 1999) in de werkgroep Tussendoelen Annex Leerlijnen. Hier zijn tevens de huidige doelen van SLO (2011) op gebaseerd. De prestaties kunnen ingedeeld worden over vijf niveaus, te weten A, B, C, D en E. Niveau A geeft aan dat het kind tot de 25% hoogst scorende leerlingen behoort, niveau B geeft aan dat het kind tot de 25% leerlingen behoort die net boven het gemiddelde scoren, niveau C geeft aan dat het kind tot de 25% leerlingen behoort die net onder het gemiddelde scoren, niveau D geeft aan dat het kind tot de 15% leerlingen behoort die ruim onder het gemiddelde scoren en niveau E geeft aan

dat het kind tot de 10% laagst scorende leerlingen behoort. De leerlingen met niveau D of E worden geselecteerd voor interventie om hun achterstand te verhelpen (Van Luit & Van de Rijt, 2009b).

Getalbegrip is een vaardigheid die door middel van remediëring verbeterd kan worden (Ruijsenaars, Van Luit, & Van Lieshout, 2004). Op dit moment zijn in Nederland verschillende interventies in gebruik om het voorbereidend rekenen bij laagscorende kleuters te bevorderen. Deze interventies hebben een bewezen significant effect. Veelgebruikte effectieve programma's zijn 'De Rekenhulp voor Kleuters' (Van de Rijt & Van Luit, 1998), 'Als speciale kleuter tel je ook mee!' met een effect van ($\Delta = 1.44$) (Van Luit & Schopman, 2000). Een nieuw programma is 'Op weg naar rekenen' met een effect van ($\eta^2 = .33$) (Toll & Van Luit, 2013). Echter, een nadeel dat deze programma's gemeenschappelijk hebben, is dat zij docentintensief zijn. Reguliere methodes in de vorm van klassikaal onderwijs hebben dit nadeel niet, maar zijn gericht op het algemeen voorbereidend rekenen, en niet op het verhelpen van grotere achterstanden bij individuele leerlingen.

Concluderend kan gesteld worden dat uit onderzoek blijkt dat educatieve computerspellen het leren kunnen bevorderen op het gebied van wiskunde en dat achterstand in getalbegrip geremedieerd kan worden. De eerste onderzoeksvraag luidt daarom als volgt: 1) Is het spel in staat om de achterstand van zwakke rekenaars te remediëren? Wat is het effect bij niet-zwakke rekenaars? Omdat uit de literatuur tevens blijkt dat getalbegrip een cognitief ontwikkelingsproces is, is gecorrigeerd voor leeftijd in maanden. De verwachting is dat het spel getalbegrip kan bevorderen.

Echter, bij het spelen van een educatief computerspel kan *cognitive overload* optreden. Doordat uit de literatuur blijkt dat kinderen met een verminderd rekenvermogen, in dit onderzoek aangeduid als zwakke rekenaars, een verminderde werkgeheugencapaciteit hebben, hebben zij meer kans op *cognitive overload* bij het spelen van een educatief computerspel. De

tweede onderzoeksvraag luidt daarom als volgt: 2) Gaan zwakke rekenaars meer vooruit op de verkorte UGT-R in een van beide condities? Hierbij zijn aan conditie 1 geen extra entertainende elementen toegevoegd anders dan de standaard gekleurde spelomgeving, en zijn aan conditie 2 wel extra entertainende elementen toegevoegd, anders dan de standaard gekleurde spelomgeving. In het vervolg zal conditie 1 aangeduid worden als hebbende geen entertainende elementen en zal conditie 2 aangeduid worden als hebbende wel entertainende elementen. Doordat entertainende elementen bij kunnen dragen aan het optreden van *cognitive overload*, wordt verwacht dat leerlingen in conditie 1 meer vooruitgaan dan in conditie 2. Ook bij deze onderzoeksvraag is gecorrigeerd voor leeftijd in maanden.

Daarnaast blijkt uit de literatuurstudie dat zwakke rekenaars meer tijd nodig hebben om rekentaken te maken. Dit leidt tot de volgende onderzoeksvraag: 3) Is de speeltijd van zwakke rekenaars verschillend in twee condities van een educatief computerspel? Hierbij zijn aan conditie 1 geen extra entertainende elementen toegevoegd anders dan de standaard gekleurde spelomgeving, en zijn aan conditie 2 wel extra entertainende elementen toegevoegd anders dan de standaard gekleurde spelomgeving. In het vervolg zal conditie 1 aangeduid worden als hebbende geen entertainende elementen en zal conditie 2 aangeduid worden als hebbende wel entertainende elementen. Doordat *cognitive overload* door entertainende elementen zorgt voor een langere reactietijd, is de verwachting dat de speeltijd van zwakke rekenaars langer is bij het spelen in een conditie met entertainende elementen. Als aanvulling op deze vraag is tevens onderzocht of er significante verschillen bestaan tussen speeltijden van de verschillende spellen. Omdat in elk spel hetzelfde aspect van voorbereidend rekenen geoefend wordt, is de verwachting dat de speeltijden niet significant van elkaar verschillen.

Helaas was het niet mogelijk om te corrigeren voor de scores op de werkgeheugentaken *Dot Matrix* en *Word Recall*. De reden hiervoor is dat er sprake is van veel missende data op

beide taken waardoor de onderzoeksgroepen te klein zouden worden, wat de betrouwbaarheid van het onderzoek in gevaar brengt.

Het wetenschappelijk belang van dit onderzoek ligt het gegeven dat uit verschillende studies blijkt dat zwakke rekenaars een verminderde werkgeheugencapaciteit hebben, zodat er eerder *cognitive overload* optreedt. Als blijkt dat de conditie met entertainende elementen inderdaad tot lagere prestaties leidt, kan verder onderzocht worden welke specifieke elementen in het spel deze *cognitive overload* veroorzaken. Mogelijk doet *cognitive load* zich in computerspellen op een andere manier voor dan in conventionele lesmethodes.

Methode

Deelnemers

In totaal participeerden 382 leerlingen in het onderzoek. De participanten zijn overgenomen van het onderzoek van Kirschner et al. (2015). Van deze leerlingen behoorden 255 leerlingen tot conditie 1 en 3 van de studie van Kirschner et al. (2015), hierna genoemd conditie 1 en 2. Deze leerlingen varieerden in de leeftijd van 3 jaar en 8 maanden tot en met 6 jaar en 9 maanden ($M = 4.88$, $SD = 0.66$). De leerlingen waren verdeeld over 24 klassen van 11 Nederlandse scholen voor het primair onderwijs. De steekproef bevatte 121 meisjes en 113 jongens. Van 21 leerlingen is het geslacht niet bekend. Alle leerlingen zaten in groep 1 of 2 van de basisschool en nemen sinds september 2014 deel aan dit onderzoek.

De totale steekproef is op basis van een pre-analyse geclassificeerd op rekenniveau. Dit is gedaan door de leerlingen in zes leeftijdscategorieën te verdelen en per leeftijdscategorie het gemiddelde van de verkorte UGT-R op meetmoment 1, voorafgaand aan het spelen van het spel, te bepalen. Omdat het gebruik van categorieën per drie maanden zoals in de officiële UGT-R leeftijdscategorieën met te weinig deelnemers zou geven, is gekozen om leeftijdscategorieën per zes maanden op te stellen. Om een betrouwbaar gemiddelde te vinden, zijn ook de scores

gebruikt van de leerlingen uit de conditie die niet is meegenomen in dit onderzoek. In tabel 1 zijn de descriptieve resultaten per leeftijdscategorie weergegeven.

Tabel 1

Frequenties en gemiddelden per leeftijdscategorie op meetmoment 1 verkorte UGT-R.

Leeftijdscategorie	<i>N</i> totaal	<i>N</i> valide	<i>N</i> missend	<i>M</i>	<i>SD</i>
1. 44-49 maanden	60	26	34	4.23	2.76
2. 50-55 maanden	94	58	36	4.62	3.46
3. 56-61 maanden	88	58	30	7.98	4.23
4. 62-67 maanden	87	57	30	9.88	4.36
5. 68-73 maanden	43	29	14	10.14	3.76
6. 74-80 maanden	10	5	5	9.00	4.07

Noot. *M* = gemiddeld aantal goede antwoorden op meetmoment 1 verkorte UGT-R.

De leerlingen die in hun leeftijdscategorie onder het gemiddelde in hun leeftijdscategorie scoorden, zijn als zwakke rekenaars gedefinieerd. Deze definitie zal voorafgaand aan onderzoeksvraag 1 onderzocht worden. Uit deze gegevens bleek dat 82 leerlingen onder het gemiddelde van hun leeftijdscategorie scoorden. 73 leerlingen scoorden boven het gemiddelde en van 100 leerlingen is geen score bekend. Beide groepen leerlingen zijn verdeeld over twee condities. In tabel 2 is de frequentie van deze zwakke en niet-zwakke rekenaars per conditie en leeftijdscategorie weergegeven.

Tabel 2

Frequentie zwakke en niet-zwakke rekenaars en criteria per conditie en leeftijdscategorie.

			Zwakke		Niet-zwakke	
			rekenaars		rekenaars	
Leeftijdscategorie	Criterium \leq	<i>N</i>		<i>N</i>		
		conditie 1	conditie 2	conditie 1	conditie 2	
1.	44-49 maanden	3	4	3	5	2
2.	50-55 maanden	4	8	14	11	5
3.	56-61 maanden	7	14	8	6	13
4.	62-67 maanden	9	8	10	12	6
5.	68-73 maanden	9	4	7	7	4
6.	74-80 maanden	8	2	0	1	1
Totaal			40	42	42	31

Noot: Criterium = maximaal toelaatbare score om tot zwakke rekenaars te behoren.

Instrumenten

Computerspel. Het door HELMGAMES ontwikkeld computerspel ‘Tel je zoo!’ is bedoeld om kinderen de basale vaardigheden van getalbegrip te leren. Het doel van het spel is om leerlingen met en zonder rekenachterstand op een vergelijkbaar niveau te laten starten in groep 3. Het computerspel begon met een opdracht die het niveau van de leerling bepaalt. Vervolgens speelde de leerling gedurende elk niveau drie individuele spellen: een aap-, olifant- en kikker-variant. Deze drie individuele spellen waren inhoudelijk hetzelfde, maar de speelwijze was anders. Daarnaast wisselde het spel de leerinhoud af door de leerling het spel in twee richtingen aan te bieden: symbool naar non-symbool en van non-symbool naar symbool. Er waren vier niveaus beginnend bij getalbegrip tot en met vijf en eindigend bij getalbegrip tot

en met twintig. Elke leerling had zijn eigen account om het computerspel te kunnen spelen. Het computerspel werd in het onderzoek van Kirschner, Janssen, Te Pas en Van Tartwijk (2015) toegepast in drie verschillende condities; in dit onderzoek werd gebruik gemaakt van twee van deze condities.

Verkorte UGT-R. Om de leerlingen tussentijds te kunnen toetsen op hun vooruitgang op het gebied van getalbegrip zijn de essentiële onderdelen de Utrechtse Getalbegrip Toets Revised (UGT-R; Van Luit & Van de Rijt, 2009a) gebruikt. Deze verkorte UGT-R bevatte 20 opgaven van de volgende onderdelen: omgaan met de telrij, omgaan met hoeveelheden en omgaan met getallen. De opgaven zijn toegevoegd als bijlage A. De UGT-R is bewerkt door Kirschner et al. (2015) zodat de getalbegrip-onderdelen die in het computerspel terugkomen tussentijds getoetst kunnen worden. De verkorte UGT-R is in dit onderzoek op twee momenten gebruikt: de test is in oktober 2014 afgenomen om het beginniveau te bepalen en in maart 2015 om de vooruitgang te meten. Op beide momenten is deze toets via internet afgenomen.

Design en procedure

De leerlingen zijn willekeurig toegewezen aan conditie 1 of 2. In conditie 2 was sprake van entertainende elementen en in conditie 1 waren deze elementen afwezig. Deze entertainende elementen in conditie 2 bestonden uit achtergrondmuziek, bewegende beelden en het ontwijken van zogenoemde ‘*bad guys*’. In het spel met de aap bestonden de bewegende beelden uit loodrecht vallende cijfers en kokosnoten. De *bad guys* in dit spel waren de vallende kokosnoten. De kokosnoten vielen sneller dan de cijfers. Indien een kokosnoot aangeklikt werd, verdwenen de cijfers en overige kokosnoten die tot dan toe in beeld waren, en verschenen nieuwe, andere cijfers en kokosnoten. Het doel van het spel met de aap was het opvangen van de juiste cijfers of het juiste aantal bananen. In het spel met de olifant waren de *bad guys* drie willekeurig bewegende insecten. Wanneer een insect werd aangeklikt, verdwenen de appels die reeds in de ton lagen zodat opnieuw begonnen moest worden met de opgave. De bewegende

beelden bestonden verder uit het heen-en-weer bewegen van de appels. Het doel van het spel met de olifant was om evenveel appels in de ton te krijgen als aangegeven met het cijfersymbool. In het spel met de kikker waren de *bad guys* bewegende insecten. Indien een insect aangeklikt werd, werd deze ‘opgegeten’ door de kikker. Het doel van het spel met de kikker was om de juiste cijfers aan te klikken. In conditie 1 was geen sprake van achtergrondmuziek, bewegende beelden en het ontwijken van *bad guys*.

Voor het beantwoorden van de eerste en tweede onderzoeksvraag zijn de variabelen ‘score verkorte UGT-R meetmoment 1’ en ‘score verkorte UGT-R meetmoment 2’ gemeten. Voorafgaand aan en zes maanden na de start van het onderzoek is bij alle leerlingen een verkorte versie van de UGT-R afgenomen. Voor onderzoeksvraag 2 is de verschilscore tussen de beide meetmomenten berekend. Voor het beantwoorden van de derde onderzoeksvraag is de variabele speeltijd, dat is de totale tijd die het kind nodig heeft om één individueel spel van de olifant, kikker of aap te spelen, gemeten.

Het volledige spel, te weten de aap-, kikker- en olifant-variant, werd elke week één keer uitgespeeld op de computer. Het verschilde per school of deze computers zich binnen of buiten de klas bevonden. Indien de computers zich buiten de klas bevonden, dan was er toezicht van een hulpouder. De leerkracht en de hulpouder mochten de leerlingen niet helpen met de opgaven en voorafgaand aan het spelen van het spel werd ook geen instructie gegeven. Op het moment van de metingen had de leerkracht geen inzicht in de scores van de leerlingen. Het is niet bekend of de scholen zelf ook voorbereidende rekenprogramma’s in hun curriculum aanbieden. Voor alle direct betrokkenen binnen de school is aan het begin van het onderzoek van Kirschner et al. (2015) een informatiemiddag geweest waarin het spel werd toegelicht en de werkwijze uitgelegd.

Analyse

Voor het beantwoorden van de eerste onderzoeksvraag zijn de scores van zwakke en niet-zwakke rekenaars op de verkorte UGT-R op meetmoment 1 en 2 vergeleken. De scores op de voormeting van de verkorte UGT-R geven het beginniveau aan van de leerlingen bij de start van dit onderzoek. De scores op de nameting geven het niveau aan na zes maanden. De scores zijn vergeleken met behulp van RM-ANCOVA (repeated measures-analysis of covariance) waarbij gecorrigeerd is voor leeftijd in maanden. Met behulp van een ANCOVA waarbij gecorrigeerd is voor leeftijd in maanden zijn de scores van zwakke en niet-zwakke rekenaars op meetmoment 1 vergeleken om te onderzoeken of deze groepen goed gedefinieerd zijn.

Voor het beantwoorden van de tweede onderzoeksvraag werden de verschillen op de verkorte UGT-R van de zwakke rekenaars tussen meetmoment 1 en 2 in conditie 1 en 2 vergeleken met behulp van een een-weg ANCOVA (analysis of covariance) waarbij gecorrigeerd is voor leeftijd in maanden. Hetzelfde is gedaan voor niet-zwakke rekenaars.

Om de derde onderzoeksvraag te beantwoorden werden de speeltijden van zwakke rekenaars in conditie 1 en 2 vergeleken met behulp van een een-weg ANCOVA waarbij gecorrigeerd is voor leeftijd in maanden. Van de deelnemers zijn de tijden gelogd waarop zij elk spel startten en eindigden. In de analyse is de gemiddelde speeltijd genomen van de eerste 10 keer dat de leerling een spel gespeeld heeft. Dit gemiddelde is per spel, de aap-, olifant- en kikker-variant, berekend. Dit resulteert per conditie in drie gemiddelden. De gemiddelde speeltijd wordt uitgedrukt in minuten en seconden. Daarnaast is met behulp van een t-toets onderzocht of de speeltijden van de spellen onderling significant van elkaar verschillen.

In de analyse is niet gecorrigeerd voor de covariabelen *DM*- en *WR*-score. De reden hiervoor is dat er sprake is van veel missende data waardoor de onderzoeksgroepen te klein zouden worden, wat de betrouwbaarheid van het onderzoek in gevaar brengt. Om deze beperking op te vangen is wel een Pearson-correlatie uitgevoerd voor onderzoeksvraag 2 en 3.

Voor dit onderzoek is gekozen voor een betrouwbaarheidsinterval van 95% met $\alpha = .05$.

Resultaten

Onderzoeksvraag 1

Assumpties. Om een betrouwbare (RM)-ANCOVA uit te kunnen voeren, is over alle 255 deelnemers een Shapiro-Wilk test uitgevoerd om te onderzoeken of de scores op de verkorte UGT-R op meetmoment 1 en 2 normaal verdeeld zijn. Deze scores zijn niet normaal verdeeld, met respectievelijk $p < .001$ en $p = .004$ op meetmoment 1 en 2. Echter, de covariabele leeftijd in maanden van de zwakke rekenaars en de niet-zwakke rekenaars is wel normaal verdeeld met respectievelijk $p = .091$ en $p = .214$, zodat toch een RM-ANCOVA uitgevoerd kan worden. De Levene's test was niet-significant voor zwakke rekenaars en niet-zwakke rekenaars, met respectievelijk $F(1, 80) = 3.204, p = .077$ en $F(1, 71) = 0.085, p = .771$ op meetmoment 1, en $F(1, 56) = 0.087, p = .769$ en $F(1, 46) = 0.271, p = .605$ op meetmoment 2. Er is voor beide groepen op beide momenten dus sprake van homoscedasticiteit. Tevens voldoet de data voor zwakke en niet-zwakke rekenaars aan de aanname van homogene regressie met respectievelijk $F(1, 78) = 0.181, p = .672$ en $F(1, 69) = 0.555, p = .459$ op meetmoment 1 en $F(1, 54) = 0.293, p = .590$ en $F(1, 44) = 2.162, p = .149$ op meetmoment 2.

Vergelijking zwakke en niet-zwakke rekenaars. Een ANCOVA is uitgevoerd om te bepalen of er een significant verschil is tussen de scores van zwakke en niet-zwakke rekenaars op de verkorte UGT-R wanneer gecorrigeerd wordt voor leeftijd in maanden. Het blijkt dat niet-zwakke rekenaars significant hoger scoren dan zwakke rekenaars, op zowel meetmoment 1, $F(1, 152) = 247.84, p < .001, \eta^2_p = .620$ als meetmoment 2, $F(2, 108) = 19.76, p < .001, \eta^2_p = .268$.

Effect spel op score. Daarnaast blijkt uit de RM-ANCOVA dat niet-zwakke rekenaars significant hoger scoren op meetmoment 2 vergeleken met meetmoment 1 wanneer gecorrigeerd wordt voor leeftijd, $F(1, 46) = 6.665, p = .013, \eta^2_p = .127$, evenals de zwakke

rekenaars, $F(1, 56) = 6.280, p = .015, \eta^2_p = .101$. Er bestaat een significant verband tussen de covariabele leeftijd in maanden en deze vooruitgang zowel bij zwakke rekenaars, $F(1, 56) = 19.579, p < .001, \eta^2_p = .259$, als bij niet-zwakke rekenaars, $F(1, 46) = 14.938, p < .001, \eta^2_p = .245$.

Relatie werkgeheugen en rekenprestatie. Zoals aangegeven is niet gecorrigeerd voor *DM*- en *WR*-score. Omdat uit de literatuur blijkt dat er een verband bestaat tussen het werkgeheugen en de prestaties op voorbereidend rekenen, is onder alle deelnemers een Pearson-correlatie uitgevoerd om dit verband toch te onderzoeken. Eerst is onderzocht of voldaan is aan de assumpties. Het blijkt dat de *DM*-scores wel normaal verdeeld zijn, $p = .066$, in tegenstelling tot de *WR*-scores, $p < .001$. De scores op meetmoment 1 en 2 van de verkorte UGT-R zijn echter niet normaal verdeeld, met respectievelijk $p < .001$ en $p = .004$. Tevens bleek uit de *scatter plot* dat is voldaan aan de assumpties van lineariteit en homoscedasticiteit.

Het blijkt dat er een significante relatie bestaat tussen *DM*-score en score op meetmoment 1 van de verkorte UGT-R, $r(74) = .441, p < .001$, maar niet tussen *WR*-score en score op meetmoment 1 van de verkorte UGT-R, $r(74) = .203, p = .079$. Ook op meetmoment 2 bestaat er een significant verband tussen score op UGT-R en *DM*-score, $r(68) = .329, p = .005$, maar niet tussen score op UGT-R en *WR*-score, $r(74) = .039, p = .739$.

Onderzoeksvraag 2

Assumpties. Om een betrouwbare ANCOVA uit te kunnen voeren, is over alle 255 deelnemers een Shapiro-Wilk test uitgevoerd om te onderzoeken of de verschillscores normaal verdeeld zijn. De verschillscores zijn normaal verdeeld met $p = .361$. Daarnaast is met de Levene's test onderzocht of sprake is van homoscedasticiteit in de twee condities bij beide groepen. De Levene's test voor de zwakke rekenaars en de niet-zwakke rekenaars was niet-significant, met respectievelijk $F(1, 56) = 0.198, p = .658$ en $F(1, 46) = 0.802, p = .375$. Er is voor beide groepen dus sprake van homoscedasticiteit. Tevens voldoet de data aan de aanname

van homogene regressie voor zwakke en niet-zwakke rekenaars met respectievelijk $F(1, 54) = 0.285, p = .596$ en $F(1, 44) = 0.893, p = .350$. Tot slot bleek uit een *scatter plot* dat is voldaan aan de assumptie van lineariteit voor beide groepen.

Vergelijking condities. Voor het beantwoorden van de tweede onderzoeksvraag is een ANCOVA uitgevoerd waarbij gecorrigeerd is voor leeftijd in maanden. Het blijkt dat zowel zwakke als niet-zwakke rekenaars niet significant meer vooruitgaan in een van beide condities, met respectievelijk $F(1, 55) = 0.115, p = .736, \eta^2_p = .002$ en $F(1, 45) = 0.011, p = .915, \eta^2_p < .001$. De verschillcores zijn weergegeven in tabel 3. Voor de zwakke rekenaars blijkt dat er geen significant verband bestaat tussen leeftijd in maanden en verschilscore, $F(1, 55) = 1.500, p = .226, \eta^2_p = .027$. Dit geldt echter niet voor de niet-zwakke rekenaars, $F(1, 45) = 4.360, p = .042, \eta^2_p = .088$.

Effect covariabele. Naar aanleiding van de significante verbanden tussen de covariabele en de afhankelijke variabele uit onderzoeksvraag 1 en 2, is een een-weg ANOVA uitgevoerd om te bepalen of er significante verschillen bestaan tussen leeftijdsgroepen. Op meetmoment 1 blijkt voor niet-zwakke rekenaars in conditie 1 en 2 een significant verschil te bestaan tussen de verschillende leeftijdsgroepen, met respectievelijk $F(5, 36) = 15.780, p < .001$ en $F(5, 25) = 3.343, p = .019$. Op meetmoment 2 bestaat dit verschil eveneens. Op meetmoment 2 blijkt voor niet-zwakke rekenaars in conditie 1 en 2 een significant verschil te bestaan tussen de verschillende leeftijdsgroepen, met respectievelijk $F(5, 23) = 3.321, p = .021$ en $F(4, 14) = 5.544, p = .007$.

Voor zwakke rekenaars geldt eenzelfde verschil op meetmoment 1. In conditie 1 en 2 blijkt een significant verschil te bestaan tussen de verschillende leeftijdsgroepen, met respectievelijk $F(5, 34) = 7.257, p < .001$ en $F(4, 37) = 12.695, p < .001$. Op meetmoment 2 geldt dit verschil alleen voor conditie 2. In conditie 1 blijkt geen significant verschil te bestaan

tussen de verschillende leeftijdsgroepen met zwakke rekenaars, $F(5,22) = 1.287, p = .305$. In

conditie 2 bestaat dit verschil wel tussen leeftijdsgroepen, $F(4,25) = 4.108, p = .011$.

In tabel 3 zijn de scores op beide meetmomenten weergegeven van de zwakke en de niet-zwakke leerlingen in beide condities.

TWEE GRADATIES VAN *COGNITIVE LOAD* IN EEN EDUCATIEF COMPUTERSPEL;
DE PRESTATIES VAN ZWAKKE REKENAARS OP GETALBEGRIP

Tabel 3

Scores op verkorte UGT-R van zwakke en niet-zwakke rekenaars in conditie 1 en 2 op meetmoment 1 en 2.

		Meetmoment 1			Meetmoment 2			Verschilscore		
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Niet-zwakke	1. 44-49 maanden	5	6.00	2.00	2	9.50	6.36	2	1.50	7.78
rekenaars conditie 1										
	2. 50-55 maanden	11	6.64	1.50	7	10.86	2.55	7	4.00	3.06
	3. 56-61 maanden	6	12.33	3.93	4	13.50	4.44	4	1.00	4.55
	4. 62-67 maanden	12	13.75	2.49	9	15.22	2.95	9	0.89	3.66
	5. 68-73 maanden	7	13.00	2.00	6	16.00	1.41	6	3.00	1.10
	6. 74-80 maanden	1	11.00		1	10.00		1	-1.00	
Totaal		42	10.57	4.03	29	13.52	3.66	29	2.07	5.54
Niet-zwakke										
rekenaars conditie 2	1. 44-49 maanden	2	5.00	1.41						
	2. 50-55 maanden	5	9.60	4.56	2	9.50	0.71	2	2.50	0.71

TWEE GRADATIES VAN *COGNITIVE LOAD* IN EEN EDUCATIEF COMPUTERSPEL;
DE PRESTATIES VAN ZWAKKE REKENAARS OP GETALBEGRIIP

21

	3. 56-61 maanden	13	11.77	3.37	9	15.22	2.54	9	2.67	3.24
	4. 62-67 maanden	6	14.50	2.07	4	15.75	2.06	4	1.25	1.71
	5. 68-73 maanden	4	12.25	2.06	3	15.33	1.16	3	3.67	1.53
	6. 74-80 maanden	1	15.00		1	8.00		1	-7.00	
Totaal		31	11.68	3.77	19	14.37	3.10	19	2.00	3.28
<hr/>										
Zwakke rekenaars										
conditie 1										
	1. 44-49 maanden	4	2.50	1.00	2	12.00	2.83	2	10.00	4.24
	2. 50-55 maanden	8	2.13	1.46	1	6.00		1	3.00	
	3. 56-61 maanden	14	4.50	1.91	11	9.09	3.62	11	4.73	3.44
	4. 62-67 maanden	8	6.00	2.93	8	9.75	4.06	8	3.75	2.55
	5. 68-73 maanden	4	8.25	0.96	4	13.50	3.11	4	5.25	2.22
	6. 74-80 maanden	2	5.50	0.71	2	9.50	0.71	2	4.00	0.00
Totaal		40	4.55	2.61	28	10.04	3.69	28	4.79	3.13
<hr/>										
Zwakke rekenaars										
conditie 2										
	1. 44-49 maanden	3	1.67	1.53	2	6.00	5.66	2	5.00	7.07
	2. 50-55 maanden	14	2.43	1.51	8	6.50	2.98	8	4.75	2.96

TWEE GRADATIES VAN *COGNITIVE LOAD* IN EEN EDUCATIEF COMPUTERSPEL;
 DE PRESTATIES VAN ZWAKKE REKENAARS OP GETALBEGRIP

22

3. 56-61 maanden	8	4.75	1.04	5	8.20	2.59	5	3.00	2.00
4. 62-67 maanden	10	5.70	1.25	8	12.38	2.72	8	6.63	2.56
5. 68-73 maanden	7	6.43	2.37	7	9.57	3.60	7	3.14	2.48
6. 74-80 maanden									
Totaal	42	4.26	2.28	30	9.03	3.78	30	4.60	3.06

Noot: *M* is het gemiddeld aantal goede antwoorden op verkorte UGT-R.

Relatie werkgeheugen en speeltijden. Zoals aangegeven is niet gecorrigeerd voor *DM*- en *WR*-scores. Echter, omdat uit de literatuur blijkt dat er toch een verband bestaat tussen het werkgeheugen en de prestaties op voorbereidend rekenen, is onder alle deelnemers een Pearson-correlatie uitgevoerd om dit verband toch te onderzoeken. Eerst is onderzocht of voldaan is aan de assumpties. Het blijkt dat de *DM*-scores wel normaal verdeeld zijn, $p = .066$, in tegenstelling tot de *WR*-scores, $p < .001$. De speeltijden in de aap-, olifant- en kikker-variant zijn niet normaal verdeeld met $p < .001$, $p < .001$ en $p < .001$. Tevens bleek uit de *scatter plot* dat is voldaan aan de assumpties van lineariteit en homoscedasticiteit.

Het blijkt dat er geen significante correlatie bestaat tussen *DM*-score en speeltijden in de aap-, olifant- en kikker-variant, met respectievelijk $r(74) = .046$, $p = .698$, $r(74) = -.096$, $p = .417$ en $r(74) = .205$, $p = .080$. Tevens blijkt dat er geen significante correlatie bestaat tussen *WR*-score en speeltijden in de aap-, olifant- en kikker-variant, met respectievelijk $r(74) = .047$, $p = .692$, $r(74) = -.094$, $p = .423$ en $r(74) = .205$, $p = .080$.

Onderzoeksvraag 3

Assumpties. Om betrouwbare een-weg ANCOVA's uit te kunnen voeren, is onder de 74 zwakke rekenaars die elk spel tien keer hebben gespeeld, een Shapiro-Wilk test uitgevoerd om te onderzoeken of de gemiddelde speeltijden normaal verdeeld zijn in de aap-, olifant- en kikker-variant. Voor de aap-, olifant- en kikker-variant is de Shapiro-Wilk test significant met $p < .001$, $p < .001$ en $p < .001$. Er is dus geen sprake van een normale verdeling van gemiddelde speeltijd in elk spel. Echter, de covariabele leeftijd in maanden van deze zwakke deelnemers is wel normaal verdeeld, $p = .271$, zodat een ANCOVA uitgevoerd kan worden. Daarnaast is bij de aap-, olifant- en kikker-variant sprake van homoscedasticiteit met respectievelijk $F(1, 72) = 0.088$, $p = .767$, $F(1, 72) = 1.332$, $p = .252$ en $F(1, 72) = 2.445$, $p = .122$. Tevens is voor de aap-, olifant- en kikker-variant voldaan aan de assumptie van homogene regressie wat betreft de covariabele leeftijd in maanden, met respectievelijk $F(1, 70) = 0.229$, $p = .633$, $F(1, 70) =$

0.045, $p = 0.833$ en $F(1, 70) < 0.001$, $p = .988$. Tot slot voldoet de data tevens aan de assumptie van lineariteit, zo bleek uit een *scatter plot*.

Vergelijking condities. Voor het beantwoorden van de derde onderzoeksvraag is een ANCOVA uitgevoerd waarbij gecorrigeerd is voor leeftijd in maanden. Voor de aap-variant geldt dat zwakke rekenaars geen significant kortere of langere speeltijd hebben in een van beide condities, $F(1, 71) = 0.447$, $p = .506$, $\eta^2_p = .006$. Daarnaast blijkt dat er geen significant verband bestaat tussen leeftijd in maanden en speeltijd in beide condities. Voor leeftijd in maanden geldt in de aap-variant in conditie 1 $F(1, 33) = 0.365$, $p = .550$, $\eta^2_p = .011$. Voor leeftijd in maanden geldt in de aap-variant in conditie 2 $F(1, 37) = 2.111$, $p = .154$, $\eta^2_p = .054$.

Ook in de olifant-variant bestaat geen significant verschil tussen de twee condities, $F(1, 71) = 0.044$, $p = .834$, $\eta^2_p = .001$. Tevens blijkt dat er geen significant verband bestaat tussen de leeftijd in maanden en de speeltijd in beide condities. Voor leeftijd in maanden geldt in de olifant-variant in conditie 1 $F(1, 33) = 0.217$, $p = .644$, $\eta^2_p = .007$. Voor leeftijd in maanden geldt in de olifant-variant in conditie 2 $F(1, 37) = 0.001$, $p = .973$, $\eta^2_p < .001$.

Ook in de kikker-variant blijken zwakke rekenaars geen significant kortere of langere speeltijd in een van beide condities te hebben, $F(1, 71) = 3.259$, $p = .075$, $\eta^2_p = .044$. Tot slot blijkt dat er geen significant verband bestaat tussen leeftijd in maanden en speeltijd in beide condities. Voor leeftijd in maanden geldt in de kikker-variant in conditie 1 $F(1, 33) = 0.228$, $p = .636$, $\eta^2_p = .007$. Voor leeftijd in maanden geldt in de kikker-variant in conditie 2 $F(1, 37) = 3.860$, $p = .057$, $\eta^2_p = .094$.

In tabel 4 zijn de gemiddelde speeltijden per spel weergegeven.

Vergelijking spellen. Daarnaast blijkt wel dat zwakke rekenaars in conditie 1 in de aap-variant een significant langere speeltijd van gemiddeld 55 seconden hebben vergeleken met de olifant-variant in conditie 1, $t(34) = 7.705$, $p < .001$, 95% CI [0.671, 1.170]. Dit significante verschil tussen de aap- en olifant-variant is ook zichtbaar in conditie 2, waarbij de speeltijd

gemiddeld 54 seconden langer is in de aap-variant, $t(38) = 7.428$, $p < .001$, 95% CI [0.652, 1.141].

Vergeleken met de olifant-variant hebben zwakke rekenaars een significant kortere speeltijd van gemiddeld 1 minuut en 45 seconden in de kikker-variant in conditie 1, $t(34) = 10.764$, $p < .001$, 95% CI [1.414, 2.072] en gemiddeld 3 minuten en 10 seconden in conditie 2 in de kikker-variant, $t(38) = 3.995$, $p < .001$, 95% CI [1.1.560, 4.765].

Tot slot blijkt dat ook de speeltijden van de aap- en kikker-variant significant van elkaar verschillen. In conditie 1 hebben zwakke rekenaars een significant kortere speeltijd van gemiddeld 2 minuten en 16 seconden in de kikker-variant, vergeleken met de aap-variant $t(34) = 22.179$, $p < .001$, 95% CI [2.471, 2.970]. Dit geldt ook voor conditie 2 met een kortere speeltijd van gemiddeld 3 minuten en 12 seconden in de kikker-variant, $t(38) = 26.487$, $p < .001$, 95% CI [2.952, 3.441]. In tabel 4 zijn de gemiddelde speeltijden per spel weergegeven.

Tabel 4

Gemiddelde speeltijden per spel per conditie in minuten en seconden.

	<i>N</i>	<i>M</i>	<i>SD</i>
Aap conditie 1	35	0:05:27	0:00:43
Aap conditie 2	39	0:05:21	0:00:45
Olifant conditie 1	35	0:04:27	0:00:58
Olifant conditie 2	39	0:04:32	0:01:46
Kikker conditie 1	35	0:02:44	0:02:00
Kikker conditie 2	39	0:02:09	0:00:37

Discussie

In dit onderzoek is het effect van een educatief computerspel onderzocht op getalbegrip. De onderzoeksvragen luiden als volgt: 1) Is het spel in staat om de achterstand van zwakke rekenaars te remediëren? Wat is het effect bij niet-zwakke rekenaars? Is de gebruikte definitie van zwakke rekenaars een goede definitie? 2) Gaan zwakke rekenaars meer vooruit op de verkorte UGT-R in een van beide condities? 3) Is de speeltijd van zwakke rekenaars verschillend in twee condities van een educatief computerspel? Verschillen de speeltijden van de spellen onderling?

Relatie werkgeheugen en rekenprestaties. Uit de resultaten blijkt dat er een significante correlatie bestaat tussen de score op de UGT-R op meetmoment 1 en 2, en de score op de *DM*-taak. Deze relatie bestaat niet tussen de score op de UGT-R op meetmoment 1 en 2, en de score op de *WR*-taak. Dit zijn respectievelijk visueel-ruimtelijke en verbale werkgeheugentaken. Dit komt niet overeen met onderzoek van Kyttälä et al. (2010) die vonden dat kinderen met vroege rekenproblemen, voordat zij formeel rekenonderwijs ontvangen, slecht presteerden op visueel-ruimtelijke en verbale werkgeheugentaken. Mogelijk doet de verkorte UGT-R minder een beroep op het verbale werkgeheugen.

Vergelijking zwakke en niet-zwakke rekenaars. Het blijkt dat op meetmoment 1 significante verschillen bestaan tussen zwakke en niet-zwakke rekenaars. Hieruit volgt dat de gebruikte definitie van zwakke rekenaars, namelijk ‘onder het gemiddelde scorend’ een valide definitie is. Echter, van 100 leerlingen misten de gegevens op meetmoment 1, wat negatieve gevolgen kan hebben voor de betrouwbaarheid en validiteit van deze definitie. Ook op meetmoment 2 is het onderscheid tussen zwakke en niet-zwakke rekenaars nog zichtbaar, maar wanneer de criteria op basis van de score op meetmoment 1 als uitgangspunt worden genomen, kan gesteld worden dat zwakke rekenaars niet langer meer gedefinieerd worden als zwak. Zij hebben overigens niet hetzelfde niveau als hun niet-zwakke leeftijdsgenoten. Dit heeft gevolgen

voor het beoogde doel van het spel. De intentie van het spel is om zwakke en niet-zwakke rekenaars op hetzelfde rekenniveau te laten starten in groep 3. Echter, omdat zowel zwakke als niet-zwakke rekenaars profiteren van het spel, blijft dit verschil tussen beide groepen bestaan, waardoor de ex-zwakke rekenaars relatief zwak blijven, maar absoluut niet.

Effect spel. De resultaten geven aan dat het spel een groot effect heeft op de vooruitgang bij zowel zwakke als niet-zwakke rekenaars. De hypothese dat het spel getalbegrip kan bevorderen, wordt aangenomen. Het gegeven dat het spel effectief is voor zowel zwakke als niet-zwakke rekenaars, sluit aan bij de resultaten van Wastiau et al. (2009) en Mitchell en Savill-Smith (2004). Zij geven aan dat instructie met behulp van een educatief computerspel effectief is bij wiskunde. Ook sluit dit aan bij het gegeven dat getalbegrip verbeterd kan worden middels remediëring (Ruijssenaars et al., 2004). Belangrijk is om op te merken dat het niet bekend was of de scholen gebruikmaken van andere methoden om voorbereidend rekenen te bevorderen. Als scholen wel gebruikmaken van dergelijke methoden kan de vooruitgang ook hierdoor veroorzaakt worden in plaats van door het spel, wat de betrouwbaarheid en validiteit van dit onderzoek in gevaar brengt.

Daarnaast blijkt dat er een significant groot verband bestaat tussen leeftijd in maanden en de vooruitgang. Dit wordt ondersteund door het gegeven dat voorbereidend rekenen een cognitief ontwikkelingsproces is, waarbij per leeftijdscategorie een norm gesteld dient te worden, aansluitend bij het ontwikkelingsniveau (Ruijssenaars et al., 2004; Van Luit & Van de Rijt, 2009a).

Vergelijking condities. Uit de analyses blijkt dat zowel zwakke als niet-zwakke rekenaars niet significant meer vooruitgaan in één van beide condities. De hypothese dat een conditie met entertainende elementen zorgt voor *cognitive overload*, wordt daarom verworpen. Dit is in tegenspraak met het ontwerpprincipe dat het toevoegen van overbodige elementen aan het educatieve materiaal zorgt voor *cognitive overload* en lagere prestaties (Mayer & Moreno,

2003). De conclusie van Sweller (2011) en Paas et al. (2003) dat bij het ontwerpen van instructie rekening moet worden gehouden met *cognitive overload*, lijkt dus niet van toepassing bij de mate van entertainende elementen in een educatief computerspel. Een mogelijke verklaring is dat de leerlingen in hun vrije tijd veel computerspellen spelen en daardoor geen hinder ondervinden van entertainende elementen in dit educatieve spel. Deze verklaring brengt een gevaar voor de interne validiteit van het onderzoek met zich mee: omdat geen rekening is gehouden met de tijd die leerlingen in hun vrije tijd aan computerspellen besteden, kan het zijn dat er wel sprake is van *cognitive overload*, maar dat deze door het ontbreken van de correctie voor computerspeeltijd in vrije tijd niet naar voren komt. Wanneer dit onderzoek herhaald zou worden, dient hiervoor gecorrigeerd te worden.

Voor de zwakke rekenaars blijkt er geen significant verband te zijn tussen de covariabele en de verschilscore, in tegenstelling tot de niet-zwakke rekenaars. Dit zou erop kunnen wijzen dat zwakke rekenaars meer profiteren van het spel dan niet-zwakke rekenaars: bij laatstgenoemde zou het cognitieve ontwikkelingsproces een grotere rol kunnen spelen.

Effect covariabele. Uit de extra analyse naar aanleiding van de resultaten uit onderzoeksvraag 1 en 2, blijkt dat er significante verschillen bestaan tussen leeftijdsgroepen wat betreft de scores op de verkorte UGT-R. Dit geldt voor zwakke en niet-zwakke rekenaars. Deze verschillen tussen leeftijdsgroepen worden bevestigd door Van Luit en Van de Rijt (2009a), die per leeftijdscategorie een norm stellen, aansluitend bij het ontwikkelingsniveau. Van belang is om op te merken dat de leeftijdsgroepen echter kleine aantallen of geen leerlingen bevatten, wat de betrouwbaarheid van dit resultaat in gevaar kan brengen. Dit verklaart mogelijk ook waarom bij zwakke rekenaars op meetmoment 2 in conditie 1 deze significante leeftijdsverschillen niet zijn gevonden.

Relatie werkgeheugen en speeltijden. Uit de resultaten blijkt dat er geen significante correlatie bestaat tussen speeltijd in de aap-, olifant- en kikker-variant, en de score op de *DM-*

taak. Ook bestaat er geen significante correlatie tussen de score op de speeltijd in de aap-, olifant- en kikker-variant, en de score op de *WR*-taak. Dit komt niet overeen met de resultaten van Ashcraft en Kirk (2001) dat een verminderde werkgeheugencapaciteit leidt tot een toename in reactietijd. Dit kan te maken hebben met het gegeven dat het spelen van het spel niet in een gestandaardiseerde omgeving plaatsvindt. Het blijkt namelijk dat sommige leerlingen het spel niet in een keer uitspelen maar het spel opnieuw starten, enige tijd geen actie ondernemen waardoor er opnieuw instructie verschijnt en het spel na verloop van tijd uit zichzelf eindigt, of regelmatig op pauze klikken. Dit geeft geen representatief beeld van de daadwerkelijke speeltijd.

Vergelijking condities. Uit onderzoek blijkt dat er een significante relatie bestaat tussen werkgeheugen en voorbereidende rekenprestaties (Raghubar et al., 2010; Friso-Van den Bos et al., 2013). Daarnaast stellen Ashcraft en Kirk (2001) dat een verminderde werkgeheugencapaciteit leidt tot een toename in reactietijd en fouten. De resultaten uit dit onderzoek weerspreken de hieruit afgeleide hypothese dat zwakke rekenaars een langere reactietijd hebben in een conditie met entertainende elementen door *cognitive overload*. Een verklaring hiervoor kan opnieuw liggen in het gegeven dat de leerlingen in hun vrije tijd veel computerspellen spelen en daardoor geen hinder ondervinden van entertainende elementen in dit educatieve spel. Deze mogelijke verklaring leidt tot hetzelfde gevaar voor de interne validiteit van het onderzoek zoals eerder aangegeven.

Daarnaast vindt het oefenen van het spel niet plaats in een gestandaardiseerde setting wat de eerder aangegeven problemen met zich meebrengt. Voor de ineffectieve speeltijd is niet voor gecorrigeerd, zodat dit mogelijk geen representatief beeld is van de speeltijden. Wanneer dit onderzoek herhaald zou worden, dienen de leerlingen elk spel volledig in een keer uit te spelen zonder onderbrekingen.

Er bestaat geen significant verband tussen leeftijd in maanden en speeltijd. Een verklaring hiervoor ligt in het gegeven dat het spel zich aanpast het cognitieve niveau van de leerling zodat leeftijd geen invloed heeft op de prestaties.

Op basis van deze resultaten is het niet duidelijk of het spel ook een diagnostische werking kan hebben. Hier is vervolgonderzoek voor nodig waarbij *in-game*-data van zwakke en niet-zwakke rekenaars vergeleken wordt.

Vergelijking spellen. Tussen de verschillende spellen bestaan echter wel significante verschillen in speeltijd. Hiervoor is geen theoretische verklaring omdat de spellen inhoudelijk hetzelfde zijn. Na zorgvuldige *in-game*-analyse blijkt dat de oorzaak mogelijk ligt in *in-game*-elementen. Wanneer een leerling in conditie 2 van de aap-variant klikt op een *bad guy*, verdwijnen alle cijfers en kokosnoten die tot dan toe in beeld waren, en verschijnen nieuwe cijfers en kokosnoten, waardoor de speeltijd verlengd wordt. Wanneer een leerling in conditie 2 van de olifant-variant de appels heeft aangeklikt, en vervolgens het insect aanklikt, ‘eet’ het insect de appels waarna de leerling de opgave opnieuw moet maken. Ook hierdoor wordt de totale speeltijd verlengd. In de kikker-variant komt een dergelijke verlenging van speeltijd niet voor. De kikker ‘eet’ het insect, maar dit heeft geen verdere gevolgen voor de voortgang van het spel. In conditie 1 van de aap- en olifant-variant moet de leerling meer klikken dan in conditie 1 van de kikker-variant waardoor de speeltijd verlengd wordt. Uit deze *in-game*-analyse wordt echter niet duidelijk waarom de tijd tussen de aap- en olifant-variant significant verschilt: de leerling moet evenveel klikken.

Als gevolg van deze verschillen tussen spellen kan in vervolgonderzoek worden onderzocht of een bepaald spel een groter positief effect heeft dan een ander spel. Mogelijkerwijs geeft een spel meer *cognitive load* wanneer de leerling meer moet klikken.

Opmerkingen. Tenslotte zijn enkele opmerkingen te plaatsen bij dit onderzoek. Ten eerste is geen enkele afhankelijke variabele normaal verdeeld. Ook de variabele *WR*-score,

gebruikt in de correlatie, is niet normaal verdeeld. De resultaten dienen daarom met voorzichtigheid te worden geïnterpreteerd. Ook is vooraf niet onderzocht of de verkorte UGT-R een betrouwbare en valide weergave is van de officiële UGT-R.

Tevens is niet gecorrigeerd voor *DM*- en *WR*-score. Dit is gedaan omdat er sprake is van veel missende gegevens waardoor de onderzoeksgroepen te klein zouden worden. Er bestaat echter wel een verband tussen het verbale en visueel-ruimtelijke werkgeheugen en getalbegrip (Kytälä et al., 2010). In een herhaling van dit onderzoek is het van belang om te corrigeren voor *DM*-score en *WR*-score.

Verder wordt de classificatie ‘zwakke rekenaar’ in dit onderzoek gedefinieerd als ‘onder het gemiddelde scorend’. De UGT-R echter definieert een zwakke rekenaar als ‘behorende tot de laagste scorende 15%’. In een herhaling van dit onderzoek dient men een zwakke rekenaar te definiëren als ‘één standaarddeviatie onder het gemiddelde scorend’. In het huidige onderzoek is daar niet voor gekozen, omdat de onderzoeksgroepen dan te klein zouden zijn.

Het is van belang om in vervolgonderzoek te onderzoeken of dit spel effectiever is dan de huidige remediëringsprogramma’s, omdat uit de literatuur duidelijk wordt dat educatieve computerspellen effectiever zijn dan conventionele instructiemethoden (Wastiau et al., 2009; Mitchell & Savill-Smith, 2004). Daarnaast kan onderzocht worden of de scores van zwakke en niet-zwakke rekenaars van elkaar verschillen wanneer alleen zwakke rekenaars met het spel oefenen, omdat blijkt dat het spel ook bevorderend is voor niet-zwakke rekenaars. Het doel van het spel is immers om alle leerlingen in groep 3 op hetzelfde niveau te laten starten.

Conclusie

De implicatie van de resultaten is dat het spel gebruikt kan worden om getalbegrip te bevorderen en te remediëren, maar dat er geen onderscheid is in condities. Ook lijkt een conditie met entertainende elementen niet voor *cognitive overload* te zorgen. Er is echter vervolgonderzoek nodig om de conclusies verder te onderbouwen.

Referenties

- Anders, Y., Rossbach, H. G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & Von Maurice, J. (2011). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*, 231–244. DOI: 10.1016/j.ecresq.2011.08.003.
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of experimental psychology: General, 130*(2), 224.
- 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*(4), 699–713. DOI: 10.1037/0022-0663.96.4.699
- Belsky, J., Vandell, D., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., & Owen, M. (2007). Are there long-term effects of early child care? *Child Development, 78*, 681–701. DOI: 10.1111/j.1467-8624.2007.01021.x.
- Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental neuropsychology, 19*(3), 273-293.
- De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology, 103*, 469-479. DOI: 10.1016/j.jecp.2009.01.010.
- Friso-Van den Bos, I. (2014). *Making sense of numbers: early mathematics achievement and working memory in primary school children*. Utrecht University.
- Friso-Van den Bos, I., van der Ven, S. H., Kroesbergen, E. H., & van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational research review, 10*, 29-44.
- Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics Education for Young Children:

- What It is and How to Promote It. *Social Policy Report*, 22(1), 3-22.
- Itō, M. (2009). *Engineering play: A cultural history of children's software*. MIT Press.
- Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories. *Learning Disabilities Research en Practice*, 22(1), 36–46.
- Kavkler, M., Tancig, S., & Magajna, L. (2003). *Follow-up study of children with very low and very high mathematical competence in preschool years*. Paper presented at the Tenth European Conference for Research on Learning and Instruction, Padua, Italy.
- Kirschner, F. C., Janssen, J. J. H. M., Te Pas, S. F., & Van Tartwijk, J. (2015). Serious games to support the development of number sense: regulating learners curiosity and cognitive load. Manuscript in preparation.
- Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 54(2), 105.
- Kyttälä, M., Aunio, P., & Hautamäki, J. (2010). Working memory resources in young children with mathematical difficulties. *Scandinavian journal of psychology*, 51(1), 1-15.
- LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child development*, 81(6), 1753-1767.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, 38(1), 43-52.
- Melhuish, E. C., Phan, M. B., Sylva, K., & Sammons, P. (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(1), 95-114.

- Mitchell, A., & Savill-Smith, C. (2004). The use of computer and video games for learning: *A review of the literature*. London: Learning and Skills Development Agency.
- Noteboom, A., & Klep, J. (2005). *Als kleuters leren tellen: Peilen en stimuleren van getalbegrip bij jonge kinderen*. Amersfoort/Enschede: CPS/SLO.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, *38*(1), 1-4.
- Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of experimental child psychology*, *80*(1), 44-57.
- 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*(4), 348-367.
- Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development*, *22*, 165-184. DOI: 10.1016/j.cogdev.2006.09.001.
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110-122.
- Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of experimental child psychology*, *91*(2), 137-157.
- 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.
- Siegler, R. S. (2009). Improving the Numerical Understanding of Children From Low-Income

- Families. *Child Development Perspectives*, 3(2), 118-124. DOI: 10.1111/j.1750-8606.2009.00090.x.
- Simmons, F. R., Willis, C., & Adams, A. M. (2012). Different components of working memory have different relationships with different mathematical skills. *Journal of experimental child psychology*, 111(2), 139-155.
- Stichting Leerplanontwikkeling. (2011). Rekenontwikkeling van het jonge kind: de doelen. In *SLO: nationaal expertisecentrum leerplanontwikkeling*. Retrieved March 3, 2015, from <http://www.slo.nl/jongekind/doelen>.
- Sweller, J. (2011). Cognitive Load Theory [Electronic Version]. *The Psychology of Learning and Motivation: Cognition in Education*, 55, 27-76.
- Tobias, S., & Fletcher, J. D. (Eds.). (2011). *Computer games and instruction*. IAP.
- Toll, S. W. M., & Van Luit, J. E. H. (2013). Accelerating the early numeracy development of kindergartners with limited working memory skills through remedial education. *Research in Developmental Disabilities*, 34, 745-755. DOI: 10.1016/j.ridd.2012.09.003.
- Treffers, A., Van den Heuvel-Panhuizen, M., & Buijs, K. (1999). *Jonge kinderen leren rekenen. Tussendoelen annex leerlijnen. Hele getallen onderbouw*. Groningen: Wolters Noordhoff.
- Van de Rijt, B. A. M. (1996). *Voorbereidende rekenvaardigheid bij kleuters. De ontwikkeling van rekenvaardigheidsschalen en een onderzoek naar de invloed van een programma*. Doetinchem: Graviant (Dissertatie).
- Van de Rijt, B. A. M., en Van Luit, J. E. H. (1998). Effectiveness of the Additional Early Mathematics program for teaching children early mathematics. *Instructional Science*, 26, 337-358.
- 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(1), 27-

40.

Van Luit, J. E. H., & Van de Rijt, B. A. M. (2009a). *Utrechtse Getalbegrip Toets Revised*.

Doetinchem: Graviant Educatieve Uitgaven.

Van Luit, J. E. H., & Van de Rijt, B. A. M. (2009b). De Utrechtse Getalbegrip Toets Revised:

Het belang van vroegtijdige signalering. *Tijdschrift voor de Orthopedagogiek*, 48, 255-270.

Wastiau, P., Kearney, C., & Van den Berghe, W. (2009). How are digital games used in schools? Complete results of the study (Final report.). Brussels: European Schoolnet, EUN Partnership AISBL.

Werkgroep Tussendoelen Annex Leerlijnen (1999). Jonge kinderen leren rekenen: hele getallen – onderbouw basisschool. Groningen: Wolters-Noordhoff.

Wiley, J., & Jarosz, A. F. (2012). Working memory capacity, attentional focus, and problem solving. *Current Directions in Psychological Science*, 21(4), 258-262.

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249-265. doi:10.1037/a0031311.

Bijlage A. Opgaven Verkorte UGT-R

1. Tel eens tot 20.
2. Wijs de vijftiende slang aan.
3. Tel eens tot negentien, met telkens één overslaan: één, drie, vijf...
4. Wijs de achttiende bloem eens aan.
5. Tel eens terug met één overslaan vanaf veertien: veertien, twaalf, tien...
6. Hoeveel stippen staan er op de dobbelsteen?
7. Tel de blokjes eens.
8. Kun jij vertellen hoeveel vingers er zijn opgestoken?
9. In welk groepje staan tien stippen?
10. Tel de blokjes eens, je kunt ze ook verschuiven.
11. Ik doe 3 blokjes onder mijn hand. Dan leg ik er nog 2 bij. Hoeveel blokjes liggen er nu onder mijn hand?
12. Weet jij welk kaartje met welk cijfer mist?
13. Maak eens een rij van elf blokjes.
14. Hoeveel blokjes liggen hier?
15. Hier ligt een rij van 14 blokjes. Haal er eens 5 weg, en tel vanaf 14 terug?
16. In de gele doos zitten 8 dropjes, en in de paarse doos zitten er 5. In welke doos zitten de meeste dropjes?

17. Welk getal ligt precies tussen dertien en vijftien? Wijs het goede getal maar aan.

18. Kijk eens hoeveel punten je hebt gegooid en wijs aan waar de pion moet komen te staan.

19. Hoeveel stippen horen er precies tussen zeven en negen stippen? Wijs het goede antwoord aan.

20. Je hebt 12 gebakjes. 7 gebakjes worden opgegeten. Hoeveel gebakjes zijn er nu nog over?

Voor elke vraag kan 0 of 1 punt(en) worden gescoord. Als het kind alles goed doet, kan het maximaal 20 punten halen. De test wordt door een begeleider voorgelezen en het kind moet af en toe wat aanwijzen op de iPad of blokjes verslepen. Bij de vragen worden afbeeldingen getoond.