

# Het effect van illustraties bij rekenopgaven: hulp of hinder?

E. C. D. M. van Lieshout en I. E. Berends<sup>1</sup>

## Samenvatting

De cognitieve belastingstheorie (Sweller, 1988) gaat over de werkgeheugenbelasting die door diverse factoren tijdens het leren oplossen van vraagstukken ontstaat. Deze factoren kunnen betrekking hebben op gelijktijdige aanwezigheid zijn van twee informatiebronnen zoals een plaatje en tekst. Het onderzoek binnen het kader van deze theorie heeft zich tot nu toe weinig beziggehouden met primair onderwijs. Rekenvraagstukjes in het moderne Nederlandse rekenonderwijs bestaan echter ook vaak uit illustraties en tekst. Hoewel deze illustraties ter ondersteuning zijn bedoeld, zijn vanuit de cognitieve belastingstheorie negatieve effecten door overbelasting te voorspellen. Een aantal concepten van de theorie die relevant zijn voor het oplossen van geïllustreerde vraagstukjes zullen worden toegevoegd aan de hand van twee onderzoeken bij 130 oudere en 100 jongere leerlingen in het primair onderwijs. Bij de oudere leerlingen bleken illustraties inderdaad een nadelig effect te kunnen hebben. Bij de jongere leerlingen was dit vooralsnog onduidelijk. Verschillen in werkgeheugencapaciteit lijken ook van belang.

## 1 Inleiding

### 1.1 Cognitieve belastingstheorie

De cognitieve belastingstheorie (*cognitive load theory*) (Sweller, 1988, 1994) heeft betrekking op de beperkte verwerkingscapaciteit van het cognitieve systeem tijdens het oplossen van diverse soorten vraagstukken in de meest brede zin, zoals op het gebied van de mechanica, elektronica, geneeskunde en programmeervaardigheid. John Sweller is de grondlegger van deze theorie. De theorie sluit aan bij het Baddeley's model van het werkgeheugen (Baddeley, 2002). De cognitieve belastingstheorie is echter meer gericht op instructievariabelen die de cognitieve belasting

bepalen dan de structuur en het functioneren van het werkgeheugen op zich.

Sweller's theorie maakt een onderscheid tussen de intrinsieke belasting opgelegd door de taak en cognitieve belasting opgelegd door de vormgeving of instructie bij de taak. Dit laatste kan zowel ineffectief/irrelevant (*extraneous*) of effectief/relevant (*germane*) zijn. De intrinsieke belasting is de belasting die bij het wezen van de taak hoort en wordt bepaald door de wisselwerking van de informatie-elementen in de taak (Sweller, 1994). Het gaat om de elementen die minimaal noodzakelijk zijn om de relatie ertussen te kunnen ontdekken en daarmee de taak te begrijpen. Hoe hoger dit minimum is, hoe hoger de elementinteractiviteit en daarmee de intrinsieke belasting. Voorkennis verlaagt deze belasting. Irrelevante belasting is belasting die ontstaat door de vorm van de presentatie van de opgaven zonder dat deze bijdraagt aan het (leren) begrijpen van de taak (Sweller, 1994). Deze vorm van belasting moet zo laag mogelijk worden gehouden (Sweller & Chandler, 1991). Effectieve ofwel relevante (*germane*) belasting is belasting die wel belangrijk is voor het leerproces. Ze helpt bij de vorming van cognitieve schema's en de automatisering van de informatieverwerking. Deze schema's en automaticiteit verlagen de intrinsieke cognitieve belasting. De totale cognitieve belasting wordt gevormd door de som van de drie genoemde vormen van cognitieve belasting (Paas, Renkl, & Sweller, 2004). In dit artikel gaat het om onderzoek naar de onmiddellijke invloed van opgavenkenmerken op de prestaties en niet om de invloed van deze kenmerken op het mogelijke leereffect. De kenmerken werden zodanig gekozen dat verondersteld kon worden dat zij (in ieder geval) de irrelevante belasting zouden verhogen.

Een ander belangrijk begrip in de cognitieve belastingstheorie is het gesplitste aandachtseffect (*split-attention effect*). Dit effect ontstaat bij informatiebronnen die ruimtelijk gezien zodanig ver van elkaar liggen dat de

benodigde integratie van de informatie gehinderd wordt (Sweller, 1994; Sweller, van Merriënboer, & Paas, 1998). Het gaat daarbij om informatiebronnen die op zich zelf niet, maar wel in combinatie, voldoende zijn om de opgave op te lossen. Zoiets kan voorkomen bij plaatjes, waarbij de begeleidende tekst niet direct in de buurt staat van de picturale elementen waar de tekst betrekking op heeft. Gesplitste aandacht verhoogt de irrelevante cognitieve belasting. Een ander effect dat de irrelevante cognitieve belasting verhoogt is redundantie. Het gaat hierbij om (minstens) twee informatiebronnen waarvan er minstens een op zichzelf genoeg is om de opgave op te lossen (Sweller e.a., 1998).

De effecten die hier besproken zijn, hebben empirische ondersteuning gekregen in diverse onderzoeken van Sweller en collega's (bijv. Sweller e.a., 1998) en diverse anderen. Alleen de concepten uit de cognitieve belastingstheorie die voor dit artikel van belang zijn, zijn hier besproken. Een overzicht van alle effecten en concepten is te vinden in onder meer Sweller (2004) en Van Merriënboer en Sweller (2005). De vele publicaties op het gebied van de cognitieve belastingstheorie gaan voornamelijk over het leren van relatief complexe vaardigheden in het secundair en hoger onderwijs. Slechts een klein aantal heeft betrekking op onderzoek in het primair onderwijs. Een voorbeeld hiervan is het onderzoek van Mwangi en Sweller (1998) op het gebied van het oplossen van redactieopgaven ofwel rekenvraagstukjes in het primair onderwijs. Zij boden leerlingen onder andere opgaven aan waarin per zin een schema was toe gevoegd. Dit schema beeldde (met een rondje) ieder element van de verzameling af met de verandering in het aantal elementen. Dit leidde tot betere prestaties dan opgaven waar de schematische representaties onder de opgave waren afgebeeld. Zij interpreteerden de betere prestaties bij de per zin geïntegreerde schema's als het resultaat van de vermindering van het gesplitste aandachts-effect.

In feite is de cognitieve belastingstheorie een instructietheorie. Vrijwel alle publicaties rapporteren dan ook over de hierboven genoemde effecten op leerprestaties. Niettemin lijkt het ook nuttig concepten uit de cognitieve

belastingstheorie te gebruiken om te onderzoeken hoe opgavenkenmerken de directe prestaties beïnvloeden. In een testsituatie waar het bijvoorbeeld gaat om het schatten van de vaardigheid om een rekenkundige rekenoperatie uit te voeren, zou de testscore verlaagd kunnen worden door irrelevante of relevante cognitieve belasting als gevolg van extra informatie zoals een plaatje. Het aantonen van een dergelijk effect is de doelstelling van de twee hier gerapporteerde onderzoeken.

## **1.2 Tekst en plaatjes in contextopgaven**

In Nederland is het rekenonderwijs gebaseerd op de principes van het realistisch reken-wiskundeonderwijs. Dat betekent onder andere dat er in de verschillende rekenmethoden en het veelgebruikte Cito-leerlingvolgsysteem Rekenen-wiskunde (Janssen & Kraemer, 2002) veel gebruik wordt gemaakt van zogenaamde contextopgaven. Deze contextopgaven bestaan doorgaans uit plaatjes met weinig tekst in de lagere klassen of groepen van het primair onderwijs. In de hogere groepen komen ook veel opgaven voor die uit een combinatie van tekst en plaatjes bestaan. Deze contextopgaven moeten, door hun concreetheid en aansluiting op de ervaringen van het kind, informele oplossingsmethoden uitlokken die de basis vormen voor de verdere formalisering van de rekenoperaties (Gravemeijer & Doorman, 1999; Treffers, 1983, 1986). Toch kunnen er vanuit de cognitieve belastingstheorie vragen worden gesteld bij het gebruik van illustraties of plaatjes in rekenopgaven. In de eerste plaats zou een combinatie van een gedrukte tekst en een plaatje mogelijk tot een redundantie-effect kunnen leiden. Dat zou het geval kunnen zijn als de tekst of het plaatje of allebei voldoende informatie bevatten om de opgave op te lossen. Vooral in een situatie waarin de informatie in de plaatjes en de tekst soms wel en soms niet volledig is, moet de leerling moeite doen om na te gaan of een informatiebron nog additionele informatie geeft die nodig is om de opgave op te lossen.

In de tweede plaats kan in een opgave de afstand tussen tekst en plaatje tot een gesplitstaandachtseffect leiden, als de tekst en het plaatje geïntegreerd moeten worden om de oplossing te vinden. (Zie voor een wat andere

interpretatie van het gesplitste aandachts-effect Florax & Ploetzner, in druk). Bij het heen en weer kijken tussen tekst en plaatje om de corresponderende informatie te zoeken, kan door de beperkingen van het werkgeheugen informatie verloren gaan. In beide gevallen (redundantie en gesplitste aandacht) wordt de irrelevante cognitieve belasting verhoogd. Seufert, Jänen en Brünken (2007) beschrijven deze processen als onderdeel van de vorming van coherentie (*coherence formation*). De lokale coherentievorming betreft het begrijpen van de op zichzelf staande representaties van het probleem. Dat wil zeggen het begrijpen van alleen het plaatje en alleen de tekst. Globale coherentievorming gaat over het op elkaar betrekken van de overeenkomstige elementen in het plaatje en de tekst (*element-to-element mapping*) en, bij de ingewikkelder opgaven, het op elkaar betrekken van overeenkomstige relaties. Dit moet dan uiteindelijk met de nodige investering van werkgeheugencapaciteit leiden tot een coherente mentale representatie. Het is te verwachten dat deze processen tot meer fouten en langere responstijden leiden dan het geval is bij opgaven waarbij irrelevante cognitieve belastingsfactoren niet of minder aanwezig zijn. Voor het oplossen van de opgaven, die het kind krijgt aangeboden, betekent dat een hoge cognitieve belasting door irrelevante factoren de ruimte beperkt die nog gebruikt kan worden voor de intrinsieke factoren (het begrijpen van het wezen van de opgave) en de effectieve factoren (die het leerproces moeten ondersteunen).

Laten we uitgaan van een situatie dat, van de drie cognitieve belastingscomponenten, de irrelevante cognitieve belasting wordt verhoogd. De cognitieve belastingstheorie voorspelt bij een overschrijding van de werkgeheugencapaciteit een verlaging van de prestaties. Het ligt dan voor de hand te voorspellen dat leerlingen met een beperkte capaciteit van het werkgeheugen eerder verlies aan prestaties zullen laten zien dan leerlingen met een minder sterke beperking van het werkgeheugen. Juist de zwakkere rekenaars hebben namelijk vaak een beperktere capaciteit van het werkgeheugen (Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Passolunghi, Vercelloni, & Schadee, 2007; Swanson,

2004). Met andere woorden, als de toevoeging van plaatjes aan gedrukte tekst de irrelevante cognitieve belasting laat toenemen, zullen de zwakkere rekenaars daar last van hebben, terwijl zij juist ondersteuning zouden moeten krijgen. Als zij geregeld fouten maken, zal dat ook de opbouw van cognitieve schemata vertragen.

Voordat het onderzoek wordt gepresenteerd waarin de bovenstaande voorspellingen werden getoetst, is het van belang eerst nog eens nauwkeuriger naar de aard van de rekenopgaven te kijken die in het realistisch rekenonderwijs en het Cito-leerlingvolgsysteem Rekenen-Wiskunde in Nederland worden gebruikt. Een analyse van de rekenopgaven die in enkele rekenmethoden (onder meer Pluspunt en Wereld in Getallen) en de Cito-toets worden gebruikt, leverde op dat de plaatjes in ongeveer drie categorieën zijn in te delen. In de eerste plaats zijn dat illustraties die geen enkele informatie over de grootte van de bekende of onbekende verzameling of de relatie daartussen geven. Ze tonen slechts de algemene situatie waar het vraagstuk betrekking op heeft. Deze illustraties worden verder 'nutteloos' genoemd in de betekenis van het niet functioneel zijn voor het oplossingsproces. De tweede categorie wordt gevormd door plaatjes die dezelfde numerieke informatie geven als de tekst en ook de relatie tussen de verzamelingen uitbeelden. De numerieke informatie wordt soms uitgebeeld met telbare objecten (dat is vooral in de laagste jaargroepen het geval) en in andere gevallen met cijfers of een combinatie van beide. Deze categorie is een voorbeeld van opgaventypen waar het redundantie-effect zou kunnen optreden. Deze opgaven worden, gezien het doel dat de ontwikkelaars van deze opgaven waarschijnlijk nastreven, de 'behulpzame' illustraties genoemd. De derde categorie bestaat uit dezelfde illustraties als bij de behulpzame opgaven, maar de tekst mist de grootteaanwijzing van een van de verzamelingen die wel in het plaatje is te zien. Deze illustraties worden verder de 'noodzakelijke' illustraties genoemd. Wanneer het kind start met het lezen van de tekst, moet het om een de oplossing te kunnen vinden ook de illustratie bekijken. De tekstuele informatie staat echter meestal op enige afstand van het

plaatje, waardoor een gesplitstaandachtseffect kan optreden.

## 2 Rekenopgaven met tekst en illustraties in groep 7 / leerjaar 5

### 2.1 Inleiding

In een eerste onderzoek (Berends & Van Lieshout, 2009) werd nagegaan wat de invloed van verschillende typen illustraties (nutteloos, behulpzaam en noodzakelijk) op de prestaties van leerlingen in groep 7 (Nederland) / 5<sup>e</sup> leerjaar (België) van het primair onderwijs was. In lijn met de voorspellingen die hierboven werden gedaan, werd verwacht dat de verschillende opgaven zouden verschillen in de irrelevante cognitieve belasting die ze veroorzaken. Daarvoor zou met name het redundantie-effect een belangrijke oorzaak kunnen zijn. Daarnaast was het de vraag of zwakke rekenaars gevoeliger voor de toegevoegde irrelevante informatie zouden zijn dan de goede rekenaars.

De kinderen kregen vier typen opgaven (zie Figuur 1): 'kale illustraties' en 'nutteloze', 'behulpzame' en 'noodzakelijke' illustraties. De kale illustratie bestond uit een formele symbolische representatie van het vraagstukje. Deze illustratie bevatte geen grafische toevoeging in de vorm van een plaatje, maar alleen een somnotatie. Er kan eventueel sprake zijn van enige redundantie als het kind de tekst met de somnotatie gaat vergelijken, maar het is tevens een abstracte representatie van een deel van de oplossing, waardoor de opgave juist gemakkelijker wordt. De verwachting was dat deze opgave tot de meest efficiënte informatieverwerking zou leiden in termen van de hoogste accuratesse en de kortste responstijd bij het beantwoorden van de opgave.

Bij de nutteloze illustratie moet het kind de illustratie inspecteren om vast te stellen dat er geen bruikbare informatie in te vinden is. Ook mag het kind zich niet laten afleiden door het plaatje. Ten opzichte van de kale illustratie is er in dit geval geen hulp van een somnotatie. Het kind moet nu zelf afleiden welke berekening moeten worden uitgevoerd. Door de aanwezigheid van nutteloze informatie in het plaatje werd een hogere irrele-

vante cognitieve belasting verwacht dan bij de kale illustraties.

De behulpzame illustratie levert ten opzichte van de tekst een verdubbeling van de informatie op in termen van numerieke informatie en de relatie tussen de verzamelingen. Dit is een klassiek voorbeeld van redundantie. Vanuit de tekst gezien is de illustratie irrelevant. Het kind kan volstaan met het begrijpen van de tekst (of het plaatje), maar moet vaststellen of de tekst alleen (of het plaatje alleen) voldoende is om de oplossing te vinden. Dit vereist een proces van het een-op-een koppelen van de verschillende elementen (en hun onderlinge relaties) in het plaatje aan dezelfde elementen (of relaties) in tekst of omgekeerd. Door deze irrelevante cognitieve belasting werd een sterker negatief effect op de prestaties verwacht dan bij de nutteloze illustraties.

Bij de behulpzame illustratie kan eerst vastgesteld worden dat de informatie identiek is aan die van de tekst waarna een coherente representatie van de opgave als startpunt voor de berekening op alleen de tekst (of op alleen het plaatje) kan worden gebaseerd. Maar bij de vierde, noodzakelijke illustratie kan tegens het ontbreken van informatie in de tekst een coherente representatie van de opgave alleen opgebouwd worden door integratie van de informatie uit het plaatje en de tekst. Daarbij moet de irrelevante informatie van het plaatje (de elementen die ook al in de tekst zijn te vinden) genegeerd worden en is heen en weer kijken van tekst naar plaatje vrijwel onvermijdelijk (gesplitste aandacht). De verwachting was dan ook dat dit opgaventype de grootste cognitieve belasting oplevert.

Bovenstaande argumenten leidden tot de hypothese dat de kale illustraties het gemakkelijkst zouden zijn, gevolgd door de nutteloze, de behulpzame en de noodzakelijke, in die volgorde (hypothese 1). Tevens werd verwacht dat de het effect van illustratie meer op de voorgrond zou treden bij de zwakke rekenaars dan bij de goede (hypothese 2).

### 2.2 Methode

#### *Deelnemers*

Er namen 130 kinderen van groep 7 (5<sup>e</sup> leerjaar) van 17 scholen voor primair onderwijs

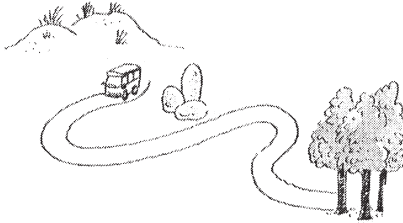
Er zaten 18 mensen in de bus. Er stapten 4 mensen uit. Hoeveel mensen zitten er nu in de bus?

$18 - 4 = ?$

Antwoord =.....

---

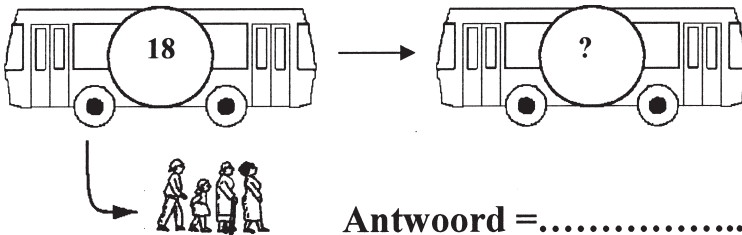
Er zaten 18 mensen in de bus. Er stapten 4 mensen uit. Hoeveel mensen zitten er nu in de bus?



Antwoord =.....

---

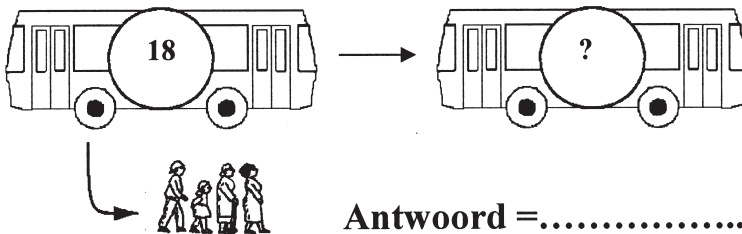
Er zaten 18 mensen in de bus. Er stapten 4 mensen uit. Hoeveel mensen zitten er nu in de bus?



Antwoord =.....

---

Er zaten 18 mensen in de bus. Er stapten een aantal mensen uit. Hoeveel mensen zitten er nu in de bus?



Antwoord =.....

Figuur 1. Voorbeelden van de vier typen illustraties. Van boven naar beneden: kaal, nutteloos, behulpzaam en noodzakelijk.

in Nederland deel aan het onderzoek. Op basis van hun score op de in Nederland gebruikelijke algemene schoolvorderingstoets voor rekenen (Cito-leerlingvolgsysteem Rekenen-Wiskunde 2002, Janssen & Kraemer, 2002) werden de kinderen met een score op of boven het normgemiddelde toegewezen aan de groep goede rekenaars en de leerlingen met een score die behoorde tot de 25% laagst scorenden aan de groep zwakke rekenaars. Leerlingen die tot de 25% laagst scorenden op een begrijpend leestoets (Cito Begrijpend Leestoets, Staphorsius & Krom, 1998) behoorden, werden van het onderzoek uitgesloten. Uiteindelijk resulteerde dit in een groep goede rekenaars bestaande uit 67 participanten (42 jongens, 25 meisjes,  $M = 9,2$  jaar,  $SD = 0,4$ ) en een groep zwakke rekenaars bestaande uit 63 participanten (23 jongens, 40 meisjes,  $M = 9,3$  jaar,  $SD = 0,4$  maanden). Het verschil in leeftijd was significant ( $t(128) = 3,55$ ,  $p < 0,001$ ), waarbij de zwakke rekenaars enigszins ouder waren dan de goede.

### Design

Er werd gebruik gemaakt van een 2 (groep: goed versus zwak) x 4 (type illustratie: kaal, nutteloos, behulpzaam, noodzakelijk) onderzoeksdesign met herhaalde metingen op de laatste factor. Accuratesse en responstijd waren de afhankelijke variabelen. Iedere deelnemer kreeg dezelfde opgaven en alle vier de typen illustraties. Welke tekst met welke illustratie samenging varieerde per deelnemer op basis van een rotatiesysteem. De score op de Cijferreeksentest (Wechsler, 2002) werd als covariaat gebruikt om te controleren voor mogelijke effecten van werkgeheugencapaciteit.

### Taken

Alle deelnemers kregen dezelfde 24 vraagstukken: 12 optel- en 12 aftrekopgaven die gelijk verdeeld waren over de vier illustratiecondities. De opgaven bevatten minstens één tweecijferig getal en bestonden zowel uit opgaven met als zonder tientalpassering. Getallen eindigend op 0 evenals dubbelen (bijv. 12 + 12) kwamen niet voor.

Er werden vier typen boekjes gemaakt waarin de volgorde van de sommen hetzelfde

was, maar de volgorde van de illustraties varieerde. De vraagstukjes konden gepresenteerd worden zonder plaatje (*kaal*), met een plaatje zonder functionele informatie (*nutteloos*), met een plaatje met informatie die gebruikt kon worden maar een dubbele van de informatie in de tekst was (*behelpzaam*) en tenslotte met een plaatje dat essentiële informatie voor de oplossing bezat (*noodzakelijk*). Om het laatste type te creëren werd bepaalde numerieke informatie uit de tekst verwijderd. Als de tekst in de behulpzame variant bijvoorbeeld luidde: "De boer heeft 48 eieren. Hij liet er 26 vallen. Hoeveel eieren heeft de boer over?" dan werd in de tekst met de noodzakelijke illustratie "26" vervangen door "een aantal". Om voorspelbaarheid te voorkomen werden de opgaven binnen een boekje in gemengde volgorde aangeboden. Die vier verschillende boekjes kwamen zoveel mogelijk in gelijke mate binnen iedere groep voor.

Om werkgeheugencapaciteit te meten werd de subtest Cijferreeksen van de WISC-III NL gebruikt. In het eerste deel van deze taak (Voorwaarts) wordt auditief een twaalf-tal cijferreeksen aangeboden, waarin het aantal cijfers per reeks oploopt van drie tot acht cijfers per reeks. Het kind wordt gevraagd de cijfers in de gepresenteerde volgorde na te zeggen. Het tweede deel van de taak (Achterwaarts) bestaat uit een twaalf-tal cijferreeksen, waarin het aantal cijfers per reeks oploopt van twee tot zeven. In dit geval moet het kind de cijfers in omgekeerde volgorde herhalen. De som van de scores (maximaal 24) werd als maat voor de capaciteit van het werkgeheugen beschouwd.

### Procedure

De kinderen kregen in een eerste individuele sessie (tijdens de schooluren) een van de opgavenboekjes aangeboden. Als start werd voor iedere illustratie één oefenopgave aangeboden. De proefleider gaf geen aanwijzingen over hoe opgaven opgelost moesten worden, dus ook niet over hoe de plaatjes erbij moesten worden gebruikt. Na de oefenopgaven moest het kind de experimentele opgaven maken. Iedere opgavetekst werd door de proefleider voorgelezen om het effect van verschillen in leesvaardigheid zoveel mogelijk te neutraliseren. Daarna mocht het kind

de opgave oplossen. De reactietijd werd gemeten vanaf het moment dat de proefleider klaar was met voorlezen totdat het kind een antwoord gaf. Als een kind binnen 2 minuten nog geen antwoord had gegeven, werd de trial afgebroken. Dit gebeurde slechts tweemaal op een totaal van 3.120 trials. In een tweede sessie werd de cijferreeksentaak aangeboden volgens de standaardinstructies van de WISC-III NL.

### 2.3 Resultaten

Tabel 1 geeft een overzicht van de prestaties in functie van groep en illustratietype. De accuratesse en responstijden werden afzonderlijk onderworpen aan een variantieanalyse met herhaalde metingen en de factoren groep (goed of zwak) als tussensubjectenfactor en type illustratie (kaal, nutteloos, behulpzaam of noodzakelijk) als binnensubjectenfactor.

#### Accuratesse

Het hoofdeffect groep op de accuratesse was significant ( $F(1, 128) = 120,38, p < 0,001, \eta_p^2 = 0,49$ ), waarbij de zwakke rekenaars lager scoorden dan de goede. Het hoofdeffect van type illustratie was eveneens significant ( $F(3, 126) = 14,99, p < 0,001, \eta_p^2 = 0,26$ ). Er was echter ook een marginaal significante interactie tussen groep en type illustratie ( $F(3, 126) = 2,51, p < 0,07, \eta_p^2 = 0,06$ ; zie Figuur 2). Geplande contrasten werden gebruikt om de voorspelde verschillen tussen de typen illustraties te toetsen. De contrasten waren *kaal* vs. *nutteloos*, *nutteloos* vs. *behulpzaam* en *behulpzaam* versus *noodzakelijk*. Zoals Figuur 2 laat zien was er geen verschil in accu-

ratesse tussen de kale en nutteloze illustraties ( $t(129) = 0,01, n.s.$ ), en dit gold zowel voor de goede als zwakke rekenaars ( $t(128) = 0,51, n.s.$ ). Anderzijds was het verschil tussen nutteloos en behulpzaam verschillend voor de twee groepen. Terwijl de goede rekenaars niet verschillend presteerden bij beide illustratietypen, presteerden de zwakke rekenaars wel lager bij de behulpzame dan de nutteloze illustraties ( $t(128) = 1,99, p < 0,05$ ). Na het toepassen van een Bonferronicorrectie met een  $\alpha$  van 0,02 was dit effect niet langer significant. Noodzakelijke illustraties leverden verdere daling van prestaties op ( $t(129) = 3,35, p < 0,001$ ).

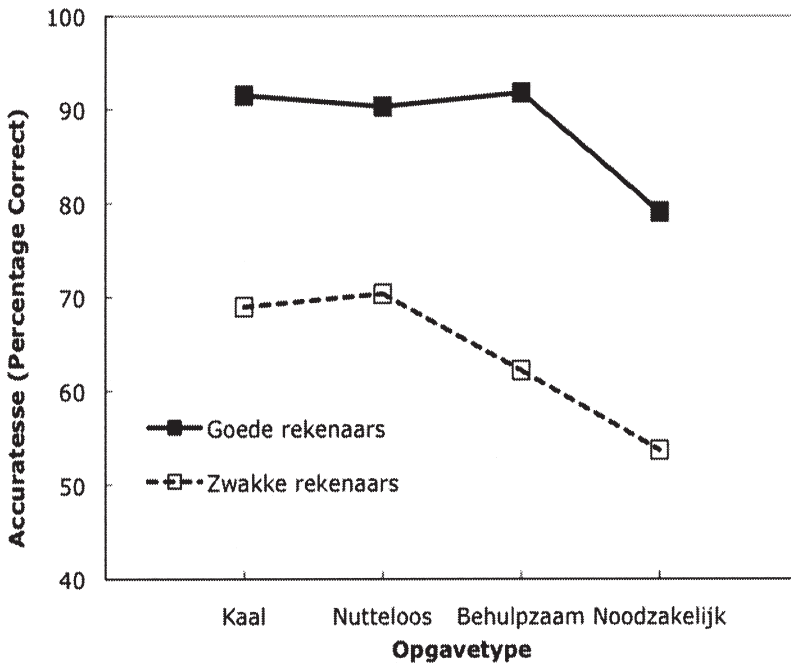
De zwakke rekenaars hadden in het algemeen lagere scores op de cijferreeksentaak dan de goede rekenaars ( $M = 7,8, SD = 2,3$  vs.  $M = 9,64, SD = 2,3$ , respectievelijk;  $t(128) = 4,67, p < 0,001$ ). Om het effect van deze maat voor werkgeheugencapaciteit op de rekenprestaties vast te stellen werd een covariantieanalyse uitgevoerd met hetzelfde factoriële design als de hierboven beschreven variantieanalyse, maar met de toevoeging van score op de cijferreeksentaak als covariaat. Het hoofdeffect groep bleef significant ( $F(1, 127) = 90,19, p < 0,001$ ), terwijl de effectgrootte ( $\eta_p^2$ ) afnam van 0,49 naar 0,42. De aanvankelijk marginaal significante interactie tussen groep en type illustratie was niet langer significant ( $F(3, 125) = 2,03, n.s.$ ). Ook het effect van illustratie was niet langer significant ( $F(3, 125) = 1,23, n.s.$ ).

Alle analyses werden ook uitgevoerd op een transformatie van de scores volgens de formule  $[X' = 2 \times \arcsin\sqrt{X}]$  om te corrigeren

Tabel 1

Gemiddelden (en SD) voor Accuratesse en Responstijd van de twee groepen en de vier typen Illustraties

Groep	Type Illustratie			
	Kaal	Nutteloos	Behulpzaam	Noodzakelijk
	Proportie correct			
Zwakke rekenaars	0,69 (0,25)	0,70 (0,25)	0,62 (0,28)	0,54 (0,31)
Goede rekenaars	0,91 (0,13)	0,90 (0,12)	0,92 (0,13)	0,79 (0,19)
	Responstijd			
Zwakke rekenaars	17,0 (5,13)	18,5 (8,09)	19,4 (7,39)	22,3 (7,98)
Goede rekenaars	11,2 (3,81)	11,8 (2,93)	13,3 (4,78)	18,2 (5,84)



Figuur 2. Accuratesse (percentage goed) van de twee rekengroepen (goed versus zwak) bij de vier typen illustraties.

voor de proportionele structuur van de scores (Winer, 1962). Het patroon van significanties veranderde hierdoor niet.

### Responstijd

De resultaten van de variantieanalyse van de responstijden week enigszins af van de resultaten van de accuratesse. Het hoofdeffect groep was significant ( $F(1, 128) = 56,76, p < 0,001, \eta_p^2 = 0,31$ ). De zwakke rekenaars hadden gemiddeld langer responstijden dan de goede rekenaars. Ook het hoofdeffect van type illustratie was significant ( $F(3, 126) = 48,62, p < 0,001, \eta_p^2 = 0,54$ ). Geplande contrasten lieten zien dat de kale opgaven significant sneller werden beantwoord dan de nutteloze ( $t(129) = 1,98, p < 0,05$ ), de nutteloze marginaal significant sneller dan de behulpzame ( $t(129) = 1,78, p < 0,08$ ), en de behulpzame significant sneller dan de noodzakelijke ( $t(129) = 5,32, p < 0,001$ ). Het groepseffect bleef significant na het toevoegen van de score op cijferreeksen als covariaat ( $F(1, 127) = 55,82, p < 0,001, \eta_p^2 = 0,31$ ). Ook het hoofdeffect illustratie bleef significant ( $F(3, 125) = 2,85, p < 0,04, \eta_p^2 = 0,06$ ).

### 2.4 Discussie

Dit onderzoek liet zien dat de accuratesse van de oplossing van een geïllustreerd vraagstuk omlaag ging, als de illustratie informatie bevatte die niet in de tekst stond maar wel noodzakelijk was voor het vinden van de oplossing. De responstijden werden langer als aan een opgavetekst een nutteloze illustratie werd toegevoegd. Deze illustratie zorgde kennelijk voor irrelevante cognitieve belasting. Er was geen sterke evidentie dat de behulpzame illustratie tot een langere responstijd leidde dan de nutteloze, hoewel door een doublure van de informatie in de tekst en de illustratie een redundantie-effect te verwachten was. Mogelijk waren de kinderen relatief snel in het vaststellen dat er sprake was van een doublure zodat geen moeite kostend proces van globale coherentievorming nodig was. Er was wel weer een duidelijke toename in responstijd van de nutteloze naar de noodzakelijke illustratie. Dit is mogelijk het gevolg van een gesplitstaandachtseffect, door de noodzaak zowel de tekst als het plaatje te doorzoeken op de voor de oplossing ontbrekende gegevens, naast een versterkte elementinteractivi-



teit die nodig was om vanuit twee informatiebronnen tot een coherente mentale representatie te komen.

Opgaven die een nutteloze of behulpzame illustratie bevatten leken de werkgeheugen capaciteit niet te boven te gaan omdat deze opgaven even accuraat werden opgelost als de kale opgaven, die alleen een somnotatie bevatten. De mate van redundantie was mogelijk niet groot genoeg om een overbelasting van het werkgeheugen te veroorzaken. Alleen de noodzakelijke illustratie leverde een lagere accuratesse op dan de andere illustraties. Het hierboven genoemde gesplitstaandachtseffect, het redundantie-effect en moeilijkheden bij coherentievorming, hebben bij deze opgave mogelijk tot een overbelasting van het werkgeheugen geleid. Met betrekking tot hypothese 1 (over het effect van verschillende illustraties op de prestaties) kan gezegd worden dat illustraties die de irrelevante cognitieve belasting bevorderden (de nutteloze en behulpzame illustraties) tot minder efficiënte informatieverwerking (langere responstijden) leidden en dat opgaven met vooral een gesplitstaandachtseffect en hoge elementinteractiviteit zowel tot langere responstijden als ook tot meer fouten leidden. Hoewel de evidentie voor een interactie tussen groep en type illustratie bij de accuratesse niet sterk was, bleek het controleren voor de invloed van het werkgeheugen tot het verdwijnen van het interactie-effect te leiden. Ook werd het hoofdeffect van groep kleiner. Dit suggereert dat een beperkter werkgeheugen eerder tot een overbelasting leidt bij verhoogde cognitieve belasting. De evidentie voor de rol van het werkgeheugen was evenwel niet sterk.

Een beperking van dit onderzoek is dat niet duidelijk is wat het effect van de somnotatie op de responstijd bij de kale redactieopgave was. De kortere responstijd van deze opgave in vergelijking met de redactieopgave met de nutteloze illustratie kan namelijk in plaats van het gevolg van het ontbreken van een illustratie ook het resultaat zijn geweest van de hulp die de leerlingen van de somnotatie ondervonden. Dit verklaart echter niet het negatieve effect van de 'noodzakelijke' illustraties ten opzichte van de 'behelpzame' illustraties. Niettemin zal in toekomstig onder-

zoek ook een vergelijking met redactieopgaven moeten worden gemaakt die noch een illustratie, noch een somnotatie bevatten.

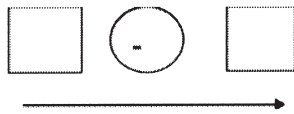
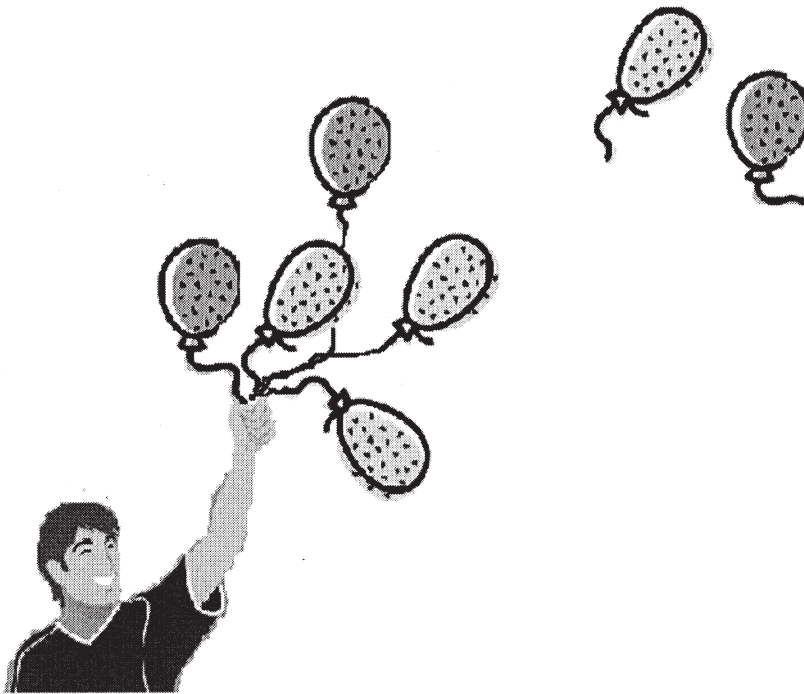
Samengevat, illustraties leiden niet tot meer goede, en bij een bepaalde illustratie zelfs tot minder goede antwoorden. Ook de efficiëntie van de verwerking in termen van responsnelheid van de informatie wordt eerder minder dan meer door de toevoeging van illustraties. Dit geldt zowel voor de goede als de zwakke rekenaars. Irrelevante cognitieve belasting lijkt hier vooral voor verantwoordelijk te zijn.

### 3 Plaatjes en tekst in groep 3 / leerjaar 1

#### 3.1 Inleiding

In een tweede onderzoek (Van Lieshout, 2009) ging het om de rekenopgaven voor de veel jongere kinderen in groep 3 (Nederland) / 1<sup>e</sup> leerjaar (België) van het primair onderwijs. In dit pilotonderzoek ging het om de vraag of, eveneens uitgaande van de cognitieve belastingstheorie, gelijksoortige effecten zouden zijn te vinden als in het eerste onderzoek. De opgaven die de rekenmethoden voor deze kinderen aanbieden wijken namelijk af van de opgaven voor oudere leerlingen zoals in het eerste onderzoek. De opgaven die in de rekenboekjes voorkomen zijn voor de jongere leerlingen vanzelfsprekend qua rekenkundige bewerking eenvoudiger. Maar belangrijker is dat ook de vorm van de opgaven anders is. De tekst bestaat doorgaans uit een enkel zinnetje (bijvoorbeeld "Hoeveel samen?"). De hoeveelheden worden meestal zodanig afgebeeld dat alle objecten in de verzamelingen telbaar zijn. Daarnaast bevatten de opgaven meestal een invulschema dat bedoeld is om kinderen te laten oefenen met het opstellen van een rekenzin. Veel opgaven hebben betrekking op een dynamische situatie: er vindt een verandering (toename of afname) in een hoeveelheid plaats. Anders geformuleerd: er is een beginverzameling, een veranderingsverzameling en een daaruit resulterende eindverzameling. Figuur 3 toont een voorbeeld van een opgave uit het onderzoek.

Het huidige onderzoek had betrekking op zulke dynamische opgaven. In dit verslag



Figuur 3. Voorbeeld van opgave in de vorm van een plaatje.

komen alleen analyses van het benoemen van de eindverzameling aan de orde. In de literatuur over tekstuele rekenvraagstukjes worden deze dynamische opgaven meestal aangeduid met de term oorzaak-veranderingsopgaven (De Corte & Verschaffel, 1986). Tekstuele oorzaak-veranderingsopgaven waarin sprake is van een gegeven beginhoeveelheid en een gegeven toename (oorzaak-verandering type 1) of een gegeven afname (oorzaak-verandering type 2) en een uit te rekenen eindhoeveelheid, behoren tot de gemakkelijkste redactieopgaven (Van Lieshout, Jaspers, & Landewé, 1994). Zulke opgaven vormen haast een handelingsrecept voor de oplosser omdat de beschreven actie min of meer door de oplosser in rekenhandelingen kan worden nagebootst (Briars & Larkin, 1984). Opgaven die bijvoorbeeld een onbekende beginverzameling hebben of meer statische opgaven lenen zich daar minder goed toe.

De opgaven die alleen uit plaatjes bestaan hebben niet de lineaire structuur van redactieopgaven van het type oorzaak-verandering. Het kind moet zelf de relaties tussen de objecten zien wat in termen van de cognitieve belastingstheorie een verhoogde elementinteractiviteit tot gevolg kan hebben. Het voorbeeld in Figuur 3 laat zien dat het kind moet begrijpen dat er een vermindering wordt afgebeeld, dat de beginhoeveelheid uit de som van beide getoonde deelverzamelingen bestaat, dat de wegvliegende ballonnen de veranderingsverzameling zijn en dat wat de jongen nog vast heeft de eindverzameling is. Het invulschema kan daarbij behulpzaam zijn, maar kan tevens door een gesplitstaandachtseffect extra geheugencapaciteit vragen: er moet heen en weer gekeken worden, er moet intussen geteld worden en het resultaat van het tellen moet onthouden worden. Het type plaatje zoals afgebeeld in Figuur 3 lijkt bo-

vendien eerder interpretatieproblemen te kunnen veroorzaken dan een plaatje dat een toename uitbeeldt. Het kind moet begrijpen dat (in tegenstelling tot het busmodel, zie Figuur 1) de ballonnen in de hand van de jongen niet de beginhoeveelheid uitbeelden, maar de eindhoeveelheid. Bij veranderingen bestaande uit een toename lijkt een dergelijk misverstand minder snel te ontstaan. Verder zijn opgaven waarin een optelbewerking wordt vereist gemakkelijker dan aftrekopgaven (Campbell, Fuchs-Lacelle, & Phenix, 2006) en daardoor mogelijk te gemakkelijk om verschillen te kunnen constateren. Daarom ging het in het huidige onderzoek om opgaven die een afname representeerden.

Maar ook een tekstuele aanbieding kan voor cognitieve belasting zorgen die niet bij het plaatje voorkomt. In de tekst worden de omvang van de begin- en veranderingverzameling met getallen aangeduid, waardoor een of andere rekenoperatie nodig is (tenzij het kind over voldoende rekenfeitenkennis beschikt). Een eenduidige hypothese over het effect van plaatjes tegenover tekst is daarom moeilijk op te stellen. Naast alleen plaatjes en alleen tekst werden in dit onderzoek ook de combinaties van plaatjes en tekst aangeboden. Omdat de numerieke informatie in de plaatjes en de teksten gelijk was, werd daarbij het optreden van een redundantie-effect verwacht wat zich zou moeten uiten in een verlaagde accuratesse en een verlengde responstijd. Evenals in het eerste onderzoek werd verwacht dat de zwakkere rekenaars een lagere werkgeheugencapaciteit zouden hebben, waardoor zij eerder zouden lijden onder een verhoogde cognitieve belasting.

Verwacht werd dat de combinaties van plaatje en tekst tot lagere prestaties zouden leiden dan plaatjes of tekst die afzonderlijk worden aangeboden (hypothese 1). Tevens werd verwacht dat dit effect sterker zou zijn bij de zwakke dan de goede (hypothese 2).

### 3.2 Methode

#### *Deelnemers*

Er namen 100 leerlingen uit groep 3 (1<sup>e</sup> leerjaar) van 9 scholen voor primair onderwijs in Nederland deel aan het onderzoek. Op basis van hun score op de algemene schoolvorde-

ringtoets voor rekenen (Cito-leerlingvolgsysteem Rekenen-Wiskunde 2002, Janssen & Kraemer, 2002) werden de kinderen met een score die volgens de norm behoorde tot de 25% hoogst scorenden toegewezen aan de groep goede rekenaars en de leerlingen met een score die behoorde tot de 25% laagst scorenden aan de groep zwakke rekenaars. In de groep goede rekenaars zaten 21 jongens en 29 meisjes en in de zwakke groep 19 jongens en 31 meisjes. De gemiddelde leeftijden van de twee groepen was nagenoeg gelijk, respectievelijk 7,10 ( $SD = 0,37$ ) en 7,11 ( $SD = 0,38$ ) jaar. Op een tempotoets met aftrekopgaven, met het grootste getal van de opgave onder de 10, haalden de goede rekenaars binnen 1 minuut een gemiddelde van 8,98 ( $SD = 4,06$ ) en de zwakke rekenaars 6,00 ( $SD = 2,86$ ). Dit verschil was significant ( $t(98) = 4,24, p < 0,001$ ). Het aantal correcte beantwoorde opgaven van deze toets moest minimaal 80% van de gemaakte opgaven zijn om aan het onderzoek te kunnen deelnemen.

#### *Design*

Er werd gebruik gemaakt van een 2 (groep: goed versus zwak) x 3 (type opgave: plaatje, tekst, combinatie van plaatje en tekst) onderzoeksdesign met herhaalde metingen op de laatste factor. Accuratesse en responstijd waren de afhankelijke variabelen. Iedere deelnemer kreeg dezelfde aantalcombinaties en alle drie de typen opgaven. Welke aantalcombinatie met welke opgave samenging varieerde per deelnemer op basis van een rotatiesysteem. De score op het onderdeel Achterwaartse Cijferreeksen van de Cijferreeksentest (Wechsler, 2002) werd als covariaat gebruikt om te controleren voor mogelijke effecten van (de fonologische lus en de centrale executieve) van het werkgeheugen (Gathercole, Pickering, Ambridge, & Wearing, 2004; Passolunghi e.a., 2007). De score op de Knox Blokkentest (Snijders & Snijders-Oomen, 1970) werd als covariaat gebruikt om te controleren voor mogelijke effecten van het visuospatieel schetsboek (Baddeley, 2002).

#### *Taken*

Alle deelnemers kregen 36 opgaven in drie sessies. Drie van 12 opgaven per sessie be-

stonden uit aftrekopgaven die ofwel alleen picturale, ofwel alleen tekstuele informatie, ofwel een combinatie daarvan bevatten. De overige opgaven zijn niet gebruikt voor de beantwoording van de huidige vraagstellingen. Het ging daarbij enerzijds om de 3 optelvarianten van de bovengenoemde aftrekopgaven. Deze optelopgaven werden (door de goede rekenaars) nagenoeg perfect beantwoord en waren daarom niet geschikt voor de analyses. Anderzijds ging het om statische optel- en aftrekopgaven (de andere helft van de 12) die om exploratieve redenen de structuur van de opgaven in rekenmethoden op school volgden en zich daardoor niet leenden voor een ongecontamineerde vergelijking op de hier van belang zijnde theoretische dimensies. Bij aanvang van de sessie werden vier oefenopgaven aangeboden. Een daarvan hoorde bij de opgaventypen waarvan hier verslag wordt gedaan.

In de opgaven ging het om beginhoeveelheden waarvan de grootte kleiner dan 10 was. Voor iedere picturale opgave bestond een variant in tekstuele of gecombineerde vorm, waarbij de aantallen en de soort objecten hetzelfde waren. Figuur 3 toont een voorbeeld van een picturale opgave. Een voorbeeld van de overeenkomstige tekst in de tekstuele en gecombineerde vorm is: "Jan had 7 ballonnen. Hij liet 2 ballonnen weg waaien. Hoeveel ballonnen heeft Jan nu?" In de gecombineerde vorm werd het invulschema (zie Figuur 3) bij dezelfde tekening als in de puur picturale vorm vervangen door dezelfde opgavetekst als in de tekstuele aanbiedingsvorm.

Om werkgeheugencapaciteit te meten werd de subtest Cijferreeksen van de WISC-III NL gebruikt. Verdere informatie staat bij de beschrijving van het eerste onderzoek. De som van de scores (maximaal 12) van de voorwaartse reeksen werd als maat voor de fonologische lus beschouwd, terwijl de som van de scores (maximaal 12) van de achterwaartse reeksen als maat voor de fonologische lus en de centrale executieve werd beschouwd. Als maat voor het visuospatieel schetsboek werd de Knox Blokken, een subtest van de Snijders-Oomen Niet-verbale Intelligentieschaal (Snijders & Snijders-Oomen, 1970) gebruikt, hoewel ook verbale

componenten een rol spelen (Richardson, 2005). In deze test tikt de onderzoeker vier blokjes aan in een vooraf vastgelegde volgorde aan (een blokje per seconde). Het kind moet het patroon van aanwijzen herhalen. De maximumscore was 20.

De (technische) rekenvaardigheid werd gemeten met een tempotoets bestaande uit 60 aftrekopgaven waarbij het eerste getal nooit groter was dan 9. De kinderen moesten binnen 1 minuut zoveel mogelijk opgaven correct oplossen.

#### *Procedure*

De kinderen kregen in drie individuele sessies (tijdens de schooluren) de opgaven een voor een op een computerscherm via Powerpoint aangeboden. Per sessie werd slechts een type opgave aangeboden (dus óf plaatjes, óf tekst óf de combinatie). De volgorde van de typen was over de sessies en apart binnen de twee groepen geroteerd. Als start werd voor iedere illustratie één oefenopgave aangeboden. De aanwijzingen van de proefleider waren hetzelfde als in het eerste onderzoek.

Na de oefenopgaven moest het kind de experimentele opgaven maken. Iedere opgavetekst werd, behalve wanneer het een plaatje alleen betrof, door de proefleider voorgelezen om het effect van verschillen in leesvaardigheid zoveel mogelijk te neutraliseren. Daarna mocht het kind de opgave oplossen en mondeling beantwoorden. De reactietijd werd bij de puur picturale opgaven gemeten vanaf het moment dat de opgave op het scherm verscheen totdat het kind een antwoord gaf. Dit antwoord moest bestaan uit het in de goede volgorde noemen van de drie getallen die in het invulschema moesten worden geplaatst. Bij de puur tekstuele en de combinatieopgave werd de tijdmeting pas gestart nadat de proefleider klaar was met het voorlezen van de opgave. Bij deze opgaven hoefde het kind alleen het aantal van de eindverzameling te noemen. De maximumscore per opgaventype was 3.

Voorafgaand aan de individuele sessies werd in een groepsessie de tempotoets met aftrekopgaven en de geheugentests (Cijferreeksen en Knox Blokken) aangeboden volgens de standaardinstructies van de betreffende tests.

### 3.3 Resultaten

De accuratesse en responstijden werden afzonderlijk onderworpen aan een variantieanalyse met herhaalde metingen en de factoren groep (goed of zwak) als tussensubjectenfactor en type opgave (plaatje, tekst, combinatie van plaatje en tekst) als binnen-subjectenfactor.

#### *Werkgeheugen*

Een multivariate variantieanalyse van de scores op de werkgeheugentoetsen met groep als tussensubjectenfactor liet een multivariaat effect zien ( $F(3, 95) = 5,80, p = 0,001, \eta_p^2 = 0,16$ ). Op alle drie de onderdelen scoorden de goede rekenaars hoger. De univariate analyses lieten zien dat er geen significant effect was op Voorwaartse Cijferreeksen ( $F(1, 97) = 1,87, p = 0,18, \eta_p^2 = 0,02; M = 9,64, SD = 1,95$  resp.  $9,02, SD = 2,53$ ), maar wel op Knox Blokken ( $F(1, 97) = 5,83, p = 0,018, \eta_p^2 = 0,06; M = 8,70, SD = 1,71$  resp.  $M = 7,80, SD = 2,01$ ), en op Achterwaartse Cijferreeksen ( $F(1, 97) = 10,53, p = 0,002, \eta_p^2 = 0,10; M = 4,78, SD = 1,23$  resp.  $M = 4,02, SD = 1,07$ ).

#### *Accuratesse*

Het hoofdeffect van opgaventype was significant ( $F(2, 96) = 8,91, p < 0,001, \eta_p^2 = 0,16$ ), evenals het hoofdeffect van groep ( $F(1, 97) = 10,73, p = 0,001, \eta_p^2 = 0,10$ ), in het voordeel van de goede rekenaars. Paarsgewijze vergelijking van de drie opgaventypen met Bonferronicorrectie liet zien dat het effect van de picturale en tekstuele opgaven niet verschilden. Het effect van beide opgaventypen verschilde wel van het combinatietype ( $p < 0,05$ ), in het voordeel van het combinatietype. De richting van het effect is daarmee in strijd met de hypothese. Er was ook een marginaal significant interactie van groep en opgaventype ( $F(2, 96) = 2,86, p = 0,062, \eta_p^2 = 0,06$ ; zie Figuur 4). Simpele effectanalyse liet zien dat de groepen significant verschilden op de tekstuele opgaven ( $p = 0,002$ ), marginaal verschilden op de picturale opgaven ( $p = 0,056$ ), en niet verschilden op de combinatieopgaven ( $p > 0,05$ ). Verder bleek dat het opgaveneffect zich alleen voordeed bij de zwakke rekenaars ( $F(2, 96) = 10,77, p < 0,001, \eta_p^2 = 0,18$ ), en niet bij de goede ( $F(2, 96) < 1, n.s.$ ).

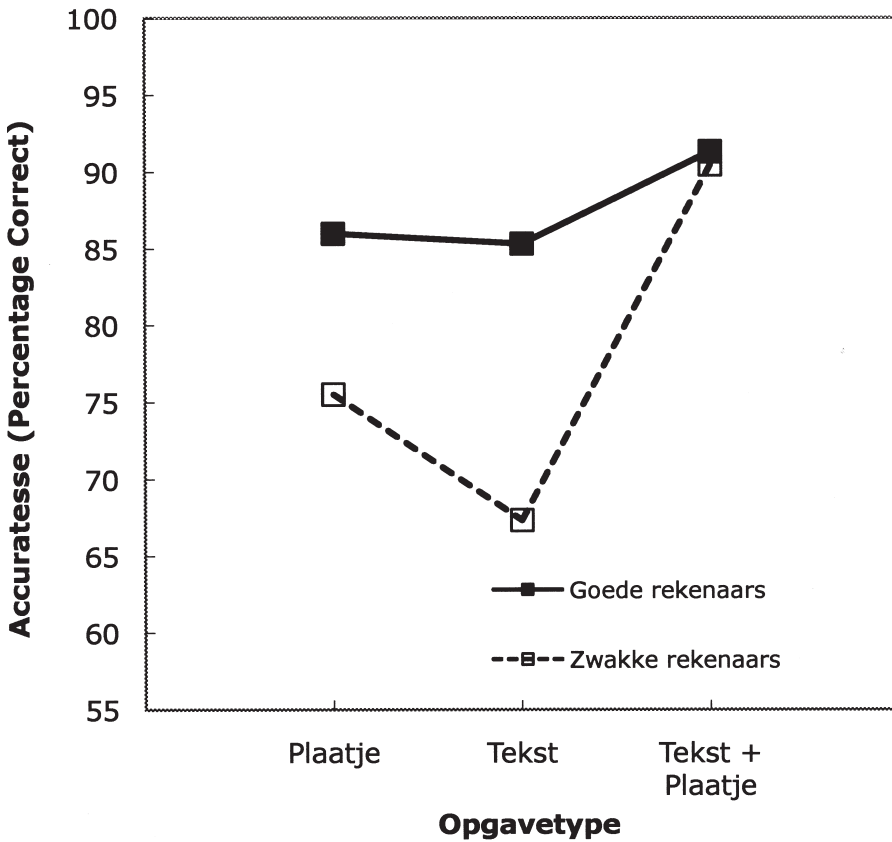
Omdat de groepen in hun prestaties verschilden op de Achterwaartse Cijferreeksen en Knox Blokken, werden aparte covariantieanalyses met telkens een van beide werkgeheugenmaten uitgevoerd. In beide analyses bleef de interactie tussen groep en opgaventype, hoewel marginaal significant, nog steeds aanwezig. De variatie in score op beide werkgeheugenmaten kon de interactie daarmee niet volledig verklaren.

Op dezelfde wijze als beschreven in paragraaf 2 werden alle analyses ook uitgevoerd op scores die gecorrigeerd waren voor hun proportionele structuur. Het patroon van significanties en marginale significanties veranderde hierdoor niet.

#### *Responstijd*

Zowel het hoofdeffect van opgaventype als dat van groep waren significant (resp.  $F(2, 96) = 328,64, p < 0,001, \eta_p^2 = 0,87$  en  $F(1, 97) = 14,62, p < 0,001, \eta_p^2 = 0,13$ ). Ook bij de responstijden was het voordeel van de combinatieopgaven boven de andere opgaven onverwacht. Er bleek ook een significante interactie van groep en opgaventype te zijn ( $F(2, 96) = 5,06, p = 0,008, \eta_p^2 = 0,10$ ; zie Figuur 5). Simpele effectanalyse liet zien dat deze interactie veroorzaakt werd door significante verschillen tussen de groepen op de picturale en tekstuele opgaven ( $p < 0,01$ ), waarbij de zwakke rekenaars het traagst waren.

De covariantieanalyse met Knox Blokken als covariaat kon de verschillen tussen de groepen niet verklaren. Dezelfde analyse met Achterwaartse Cijferreeksen leverde een genuanceerd beeld op. De interactie tussen groep en covariaat bleek marginaal significant ( $F(1, 95) = 3,85, p = 0,053, \eta_p^2 = 0,04$ ), hetgeen er op wees dat de binnengroepscoëfficiënten niet gelijk waren en een reguliere covariantieanalyse niet verantwoord was. Daarom werden om exploratieve redenen binnen de groepen en de opgaventypen de zes correlaties tussen Achterwaartse Cijferreeksen en responstijd berekend. Van deze correlaties was alleen de correlatie binnen de groep zwakke rekenaars bij de picturale opgaven significant ( $r(50) = -0,29, p = 0,04$ ). Deze correlatie werd vervolgens vergeleken met de dezelfde correlatie binnen dezelfde



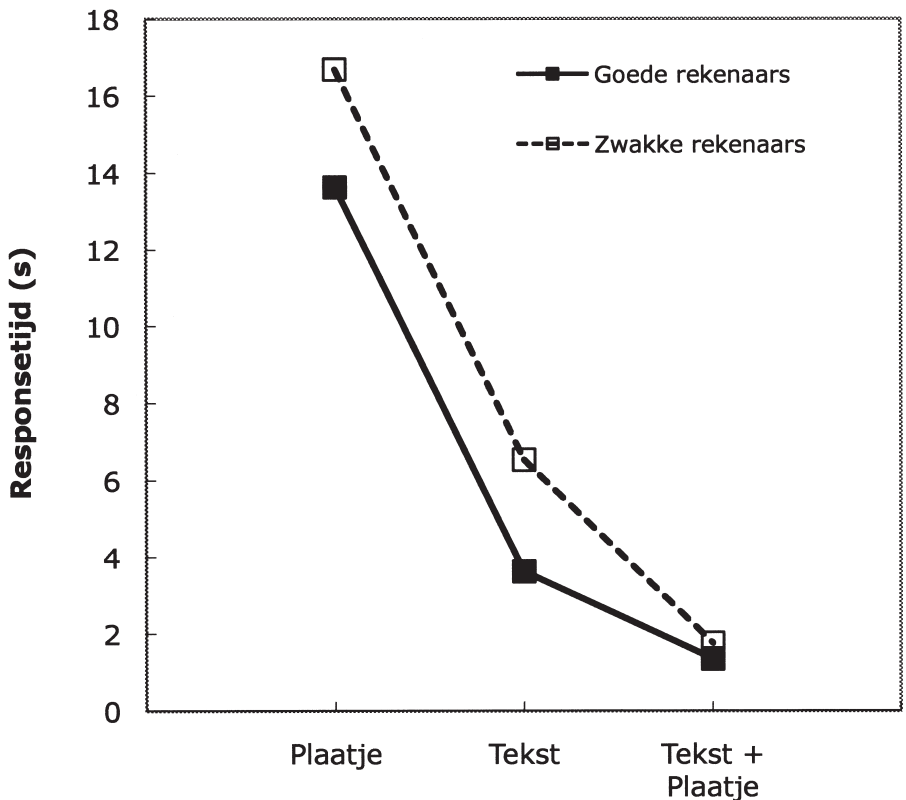
Figuur 4. Accuratesse (percentage goed) van de twee rekengroepen (goed versus zwak) bij de drie opgaventypen.

groep goede rekenaars, waar deze correlatie niet significant was ( $r(49) = 0,13, p = 0,25$ ). Het verschil was significant ( $z = 2,06, p = 0,04$ ). Alleen de zwakke rekenaars waren sneller met hun antwoord bij de picturale opgaven als zij een hogere score op Achterwaartse Cijferreeksen hadden.

### 3.4 Discussie

De verwachting (hypothese 1) was dat de combinatie van tekst en plaatje schadelijk zou zijn voor de prestaties in vergelijking met alleen een plaatje of een tekst. Het argument hiervoor was dat de combinatie van plaatje en tekst door het optreden van een redundantie-effect tot een verzwaarde cognitieve belasting zou leiden. Geheel onverwacht en ook in strijd met het voorspelde interactiepatroon (hypothese 2) bleek deze combinatie juist tot

de hoogste accuratesse en de snelste responsen te leiden bij de zwakke rekenaars en tot de snelste responsen bij de goede rekenaars. Wat kan hiervoor de reden zijn? De sleutel voor het antwoord op deze vraag is mogelijk het modaliteitsprincipe. Voordat dit duidelijk het modaliteitsprincipe uit te leggen. Dit principe houdt in dat opgaven, die zowel in picturale vorm als gesproken tekst aangeboden worden, tot een hoger leereffect leiden dan wanneer de tekst niet gesproken, maar gedrukt aangeboden wordt (Moreno & Durán, 2004; Moreno & Mayer, 1999). Het modaliteitsprincipe is een onderdeel van de cognitieve theorie van multimedialeren van Mayer (2003), een theorie die verwant is aan de cognitieve belastingstheorie. Mayer gaat, net zoals Baddeley (2002) dat in zijn werkgehe-



Figuur 5. Responsetijd (s) van de twee rekengroepen (goed versus zwak) bij de drie opgaventypen.

genmodel doet, uit van onderscheiden visuele en auditieve subsystemen, die ieder hun eigen beperkte verwerkingscapaciteit hebben. De combinatie van een plaatje en gedrukte tekst moet via hetzelfde, bijv. visuele, subsysteem worden verwerkt en vormt daardoor een grotere belasting dan de combinatie van een plaatje en gesproken tekst die via twee verschillende subsystemen (het visuele en het auditieve) worden verwerkt.

De onverwachte superioriteit van de combinatie van plaatje en tekst in dit tweede onderzoek zou met het modaliteitsprincipe als volgt kunnen worden verklaard. Om er zeker van te zijn dat de beginnende lezers in dit onderzoek een tekstbasis konden construeren vanuit de opgavetekst, lazen de proefleiders de tekst voor. Volgens de theorie van multimedia leren (Mayer, 2003) zou dit zelfs meer schadelijke redundantie hebben moeten toevoegen dan te verwachten was op basis van de combinatie van plaatje en schriftelijke tekst, omdat

de gesproken en gedrukte tekst identiek waren. Observaties tijdens het experiment suggereerden echter dat de kinderen bij het voorlezen van de tekst bij de combinatieopgaven vaak niet naar de opgavetekst keken. Helaas is dit niet systematisch geregistreerd, zodat geen kwantificering mogelijk is. Maar als het waar is dat de kinderen bij dit opgaventype de opgavetekst vaak niet bekeken, dan lijkt het erop dat de kinderen zelf een modaliteitseffect zonder redundantie-effect creëerden: zij hoorden de tekst en keken alleen naar het plaatje. De gesproken opgavetekst kan dan als een gids gewerkt hebben om de aandacht op de relevante delen van het plaatje te richten. In feite zou dit ertoe geleid kunnen hebben dat ze alleen nog maar de eindhoeveelheid hoefden te tellen. De irrelevante cognitieve belasting zou op die manier minimaal zijn geweest.

De verworpen hypothese 2, dat de prestaties van de zwakke rekenaars meer zouden lijden onder een combinatie van tekst en

plaatje, was gebaseerd op de bevestigde aanname dat ze over minder geheugencapaciteit zouden beschikken. Dit zou sneller tot een overbelasting door de redundantie leiden. Er bleek zich echter een andere interactie tussen groep en opgavetype voor te doen die inhield dat de zwakke rekenaars slechter op de tekstuele en picturale opgaven presteerden dan de goede rekenaars, terwijl er geen verschil was bij de combinatieopgaven. Als de voorgaande conclusie terecht is dat er sprake is geweest van een modaliteitseffect, dan is het niet verrassend dat het verschil tussen goede en zwakke rekenaars afwezig was bij de combinatieopgaven. Door de lage irrelevante cognitieve belasting, was er minder kans dat verschillen in werkgeheugencapaciteit een rol zouden gaan spelen.

Niettemin leken verschillen in werkgeheugencapaciteit een rol te spelen bij de snelheid waarmee de zwakke leerlingen de picturale opgaven oplosten. Omdat bij deze opgave getallen moesten worden genoemd die in het schema onder het plaatje moesten passen, trad mogelijk een gesplitstattention-effect op. Dit effect leidt tot hogere irrelevante cognitieve belasting en kan daardoor bij leerlingen met een beperktere werkgeheugencapaciteit tot een verlenging van de responstijd hebben geleid. Het feit dat de goede rekenaars niet zo'n effect lieten zien, is mogelijk het gevolg van beter ontwikkelde, geautomatiseerde en abstracte cognitieve schema's, waardoor hun irrelevante cognitieve belasting zodanig laag was dat beperkingen in het werkgeheugen geen rol speelden. Deze verklaring sluit aan op de conclusies van Seufert, Schütze en Brünken (2009). In een onderzoek met universiteitsstudenten aan wie conceptuele biochemische kennis in een multimedia-experiment werd geïnstrueerd, vonden Seufert e.a. dat de prestaties in een puur visuele conditie (illustraties met gedrukte hulpstukken) lager waren dan in een conditie met visuele en auditieve informatie (illustratie met gesproken hulpstukken). Het werkgeheugen bleek, zoals in het huidige onderzoek, geen rol te spelen bij de multimodale, maar wel bij de puur visuele opgavenpresentatie.

Een alternatieve verklaring<sup>2</sup> voor het hierboven geschetste modaliteitseffect is de volgende. Het zou kunnen dat de onverwacht be-

tere prestaties bij de combinatie van plaatjes en afgebeelde tekst het gevolg was van een groter beroep op het zelf ontdekken van de mathematische relaties in de opgaven die ofwel alleen uit een plaatje ofwel alleen uit een afgebeelde tekst bestonden. De laatste twee opgaven zouden in deze verklaring door hun vorm meer irrelevante informatie bevatten dan de combinatie-opgaven. Daarin worden de mathematische relaties mogelijk sneller duidelijk. Onderzoek in het kader van de cognitieve belastingstheorie heeft laten zien dat exploratief of ontdekkend leren minder effectief lijkt te zijn dan het leren op basis van uitgewerkte voorbeelden die door hun aard minder irrelevante informatie bevatten (Sweller & Chandler, 1991). De hier gegeven alternatieve verklaring is echter wel in strijd met de bevindingen in onderzoek 1. Verder onderzoek zal de houdbaarheid van de verschillende verklaringen moeten toetsen.

Een beperking van deze studie was dat voor de vergelijkbaarheid met het eerste onderzoek geen invulschema's bij de tekstuele opgaven en de combinatieopgaven waren opgenomen. Bij de picturale opgaven is dit wel gedaan, ten eerste, omdat zonder schema de interpretatie van het plaatje wordt bemoeilijkt en ten tweede om dicht bij de vorm van de opgaven in de rekenmethoden te blijven. In toekomstig onderzoek zou bijvoorbeeld toch bij ieder type opgave door de leerling een rekenzin geproduceerd moeten worden. Dat is bovendien van belang gezien de oorspronkelijke bedoeling van ontwikkelaars bij de constructie van picturale opgaven. Een tweede beperking is dat doordat de proefleiders de opgaven voorlezen, onbedoeld mogelijk een modaliteitseffect optrad waardoor geen toetsing van het veronderstelde redundantie-effect meer mogelijk was. Een interessante mogelijkheid om hier verder duidelijkheid te brengen is het observeren van oogfixaties (Van der Schoot, Bakker, Horsley, & Van Lieshout, 2009) om na te gaan in welke mate de opgavetekst wordt geraadpleegd.

#### 4 Algemene discussie

Naar aanleiding van de resultaten van het tweede onderzoek werd geopperd dat een



modaliteitseffect verantwoordelijk is geweest voor het vinden van andere dan de voorspelde resultaten. Het ging om het niet optreden van een prestatieverlagend redundantie-effect bij de combinatie van plaatje en tekst en om het niet lager presteren op dit type opgave door de zwakke rekenaars in vergelijking met de goede. De vraag is waarom dit effect zich niet heeft voorgedaan in de eerste studie waar de proefleiders immers ook de opgavetekst voorlezen. Een mogelijke verklaring ligt in het verschil in leesvaardigheid van de groep 7 leerlingen in onderzoek 1 en de groep 3 leerlingen van onderzoek 2. Terwijl in de laatste groep de leesinstructie net aangevangen was, zullen de meeste leerlingen van groep 7 een redelijk niveau van automaticiteit in woordherkenning hebben bereikt. Daardoor zullen zij mogelijk eerder last, in de vorm van een redundantie-effect, hebben gehad van de zichtbare opgavetekst. Ook hierin zou oogfixatieonderzoek uitsluitsel kunnen geven. Ook is het mogelijk dat de gidsfunctie van de voorgelezen tekst in onderzoek 2 de aandacht naar de telbare eindverzameling bracht, terwijl in onderzoek 1 dan nog steeds iets uitgerekend moest worden.

Verder geldt voor beide onderzoeken dat het geen instructieonderzoeken zijn. Welke opgaventypen zich beter lenen voor het aanleren van de beoogde vaardigheden staat nog te bezien. Het zou zo kunnen zijn dat opgaven, die aanvankelijk moeilijk zijn, juist mogelijkheden tot een leerproces geven waar gemakkelijker opgaven niet in voorzien. Maar volgens de cognitieve belastingstheorie moet het leren niet gebeuren met opgaven die het de leerling door irrelevante cognitieve belasting juist extra moeilijk maken (Sweller & Chandler, 1991).

Een voorlopige implicatie van het eerste onderzoek is dat ontwikkelaars van rekenmethoden en -toetsen zich moeten realiseren dat het gebruik van illustraties negatieve effecten op de directe prestaties kunnen hebben. In het tweede onderzoek blijft nog onduidelijk of plaatjes gecombineerd met tekst een negatief effect hebben. Maar ook als de opgaven alleen in de vorm van een plaatje worden aangeboden (zoals in de rekenmethoden gebeurt) lijkt bij de zwakke rekenaars een geringere capaciteit van de centrale executieve van het

werkgeheugen hun efficiëntie bij het uitvoeren van de rekentaak te beperken. Als de verklaring van een modaliteitseffect voor de onverwachte superioriteit van de combinatie van plaatjes en gesproken tekst in het tweede onderzoek juist is, opent dat nog verder te onderzoeken mogelijkheden voor rekeninstructie met een auditieve component.

Een belangrijke nieuwe stap in het toepassen van de cognitieve belastingstheorie op het soort eenvoudige opgaven die in rekenmethoden voorkomen, zou het doen van instructieonderzoek zijn waarbij nagegaan wordt welke opgavenkenmerken het leerproces bevorderen of verstoren. Pas in een instructiestudie kan duidelijk worden of opgavenkenmerken de relevante (*germane*) belasting bevorderen, dat wil zeggen de belasting die het leren en generaliseren bevordert. In het algemeen lijkt deze theorie vruchtbaar te zijn bij het onderzoek naar het effect van diverse componenten in de rekenopgaven van moderne rekenmethoden en rekentoetsen.

## Noten

- 1 Wij willen graag dank zeggen aan Freek Büscher, Alis van Lieshout en Judith Ludikhuizen voor de uitvoering van het eerste onderzoek en aan Lieneke Bodewes, Bettine van der Haar en Janneke Oussoren voor de uitvoering van het tweede onderzoek.
- 2 Deze verklaring werd aangereikt door een van de beoordelaars.

## Literatuur

- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7, 85–97.
- Berends, I., & Lieshout, E. C. D. M. van. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction*, 19, 345–353.
- Briars, D. J., & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. *Cognition and Instruction*, 1, 245–296.
- Campbell, J. I. D., Fuchs-Lacelle, S., & Phenix, T. L. (2006). Identical elements model of arithmetic memory: Extension to addition and subtraction. *Memory & Cognition*, 34, 333–647.

- De Corte, E., & Verschaffel, L. (1986). Navertellen van een redactie-opgave als toegang tot de probleemrepresentatie. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 2, 88-104.
- Florax, M., & Ploetzner, R. (in druk). What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity. *Learning and Instruction*.
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40, 177-190.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, 88, 121-151.
- Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: a calculus course as an example. *Educational Studies in Mathematics*, 39, 111-129.
- Janssen, J., & Kraemer, J.M. (2002). *Rekenen-Wiskunde*. Arnhem, Nederland: Cito.
- Lieshout, E. C. D. M. van (2009, augustus). *Cognitive load of pictorial and verbal arithmetic problems in grade 1 children with a high or low arithmetic level*. Paper gepresenteerd op de 13<sup>e</sup> tweejaarlijkse EARLI conferentie, Amsterdam.
- Lieshout, E. C. D. M. van, Jaspers, M. W. M., & Landewé, B. H. M. (1994). Mathematical word problem solving of normally achieving and mildly mentally retarded children. In J.E.H. van Luit (Ed.), *Research on learning and instruction of mathematics in kindergarten and primary school* (pp. 344-365). Doetinchem, Nederland: Graviant.
- Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, 13, 125-139.
- Merriënboer J. J. G. van, & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147-177.
- Moreno, R., & Durán, R. (2004). Do multiple representations need explanations? The role of verbal guidance and individual differences in multimedia mathematics learning. *Journal of Educational Psychology*, 96, 492-503.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia Learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368
- Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction*, 16, 173-199.
- Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, 32, 1-8.
- 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.
- Richardson, J. T. E. (2005). Knox's cube imitation test: A historical review and an experimental analysis. *Brain and Cognition*, 59, 183-213.
- Schoot, M. van der, Bakker, A., Horsley, T. M., & Lieshout, E. C. D. M. van. (2009). The consistency effect depends on markedness in less successful but not successful problem solvers: An eye movement in primary school children. *Contemporary Educational Psychology*, 34, 58-66.
- Seufert, T., Jänen, I., & Brünken, R. (2007). The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence formation. *Computers in Human Behavior*, 23, 1055-1071.
- Seufert, T., Schütze, M., & Brünken, R. (2009). Memory characteristics and modality in multimedia learning: An aptitude-treatment-interaction study. *Learning and Instruction*, 19, 28-42.
- Snijders, J. T., & Snijders-Oomen, N. (1970). *Snijders-Oomen Niet-Verbale Intelligentieschaal: SON-'58*. Groningen, Nederland: Wolters-Noordhoff.
- Staphorsius, G., & Krom, R. S. H. (1998). *Toetsen Begrijpend Lezen, handleiding*. Arnhem, Nederland: CITOgroep.
- Swanson, H. L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Memory & Cognition*, 32, 648-66.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive*

*Science*, 12, 257-285.

- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295-312.
- Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, 32, 9-31.
- Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, 8, 351-362.
- Sweller, J., Merriënboer, J. J. G., van & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.
- Treffers, A. (1983). Geïntegreerd cijferen volgens progressieve schematisering. *Pedagogische Studiën*, 60, 351-362.
- Treffers, A. (1986). Analyseren en ontwikkelen van reken/wiskundeonderwijs vanuit twee verschillende basisconcepties. *Pedagogische Studiën*, 63, 14-25.
- Wechsler, D. (2002). *Wechsler intelligence scale for children – III NL*. Lisse, Nederland: Harcourt.
- Winer, J. B. (1962). *Statistical principles in experimental design* (2nd ed.). Tokyo: McGraw-Hill.

Manuscript aanvaard: 28 juli 2009

## Auteurs

**Ernest C. D. M. van Lieshout** is werkzaam bij de Afdeling Orthopedagogiek, Faculteit der Psychologie en Pedagogiek, Vrije Universiteit Amsterdam en **Inez E. Berends** bij PI Research, Amsterdam.

*Correspondentie-adres:* Ernest C. D. M. van Lieshout, Afdeling Orthopedagogiek, Faculteit der Psychologie en Pedagogiek, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam. Email: ECDM.van.Lieshout@psy.vu.nl.

## Abstract

### The effect of illustrations for arithmetic problems: help or hindrance?

The cognitive load theory (Sweller, 1988) concerns working memory load that originates from different factors during problem solving. These factors can concern two simultaneously existing sources of information like a picture and text. To date, research within this theoretical framework has scarcely been engaged in primary education. However, math problems in modern Dutch math education often consist of illustrations and text. Although these illustrations are intended as support, from cognitive load theory negative effects are to be expected as a result of overload. A number of concepts of the theory that are relevant for solving illustrated math problems will be clarified on the bases of two studies with 130 older and 100 younger students in primary education. The performance of the older children indeed appeared to suffer from illustrations. For the younger children this was unclear as yet. Differences in working memory seem also to be important.