



Universiteit
Leiden
The Netherlands

Explanatory latent variable modeling of mathematical ability in primary school : crossing the border between psychometrics and psychology

Hickendorff, M.

Citation

Hickendorff, M. (2011, October 25). *Explanatory latent variable modeling of mathematical ability in primary school : crossing the border between psychometrics and psychology*. Retrieved from <https://hdl.handle.net/1887/17979>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/17979>

Note: To cite this publication please use the final published version (if applicable).

Summary in Dutch

(Samenvatting)

De rekenvaardigheid van Nederlandse leerlingen is een onderwerp van felle debatten. Eén punt van discussie is het *rekenonderwijs* dat de afgelopen decennia een hervorming van internationale reikwijdte heeft ondergaan. Deze vernieuwing kan grofweg omschreven worden als het verlaten van de traditionele aanpak waarin leerkrachten rekenkennis en -vaardigheden direct instrueren aan hun leerlingen, die op hun beurt moeten oefenen en 'stampen'. Hier is een aanpak voor in de plaats gekomen waarin de informele voorkennis die leerlingen hebben aan de basis van een leerlijn ligt. Het doel is niet alleen procedurele kennis en vaardigheden bij de leerling te bewerkstelligen, maar vooral ook inzicht, flexibiliteit en creativiteit. Een ander discussiepunt is het *prestatieniveau* van leerlingen in zowel het basisonderwijs als in het voortgezet onderwijs. Op regelmatige basis vinden grootschalige nationale en internationale peilingsonderzoeken van de rekenvaardigheid van leerlingen plaats, waarin wordt gerapporteerd over ontwikkelingen over tijd, vergelijkingen tussen landen en afwijkingen van de onderwijsdoelen die binnen een land gelden. Deze resultaten vormen vaak de aanleiding voor de discussie over het prestatieniveau.

Dit proefschrift – het resultaat van een samenwerking tussen het Instituut Psychologie van de Universiteit Leiden en Cito Instituut voor Toetsontwikkeling – richt zich op de rekenvaardigheid van Nederlandse basisschoolleerlingen. Het startpunt was de resultaten van de meest recente *Periodieke Peiling van het OnderwijsNiveau* (PPON) rekenen-wiskunde aan het einde van het basisonderwijs (groep 8 – internationaal *sixth grade*; leerlingen van 12 jaar oud). Deze peiling is in 2004 uitgevoerd door Cito (J. Janssen et al., 2005; zie ook Van der Schoot, 2008). PPON-2004 was de vierde peilingscyclus, met

eerdere peilingen in 1987, 1992 en 1997; de dataverzameling voor de vijfde cyclus vindt in 2011 plaats. De ontwikkelingen van het rekenvaardigheidsniveau over de periode 1987 tot 2004 zijn uiteenlopend: op sommige domeinen gingen de prestaties vooruit, terwijl ze op andere domeinen juist achteruit gingen. Daarnaast bleek dat het niveau op vrijwel alle onderdelen achterbleef bij de kerndoelen – op sommige onderdelen meer dan op andere. Deze resultaten haalden de kranten en andere media, waarin vervolgens verschillende mensen hun mening hierover gaven. Een terugkerend element hierin is de didactische theorie van het *realistisch rekenen* (RME; zie bijvoorbeeld Freudenthal, 1973, 1991; Treffers, 1993) die de meest invloedrijke theorie in het hedendaagse Nederlandse rekenonderwijs is geworden sinds de jaren '80 en '90 van de vorige eeuw. Het realistisch rekenen roept sterke gevoelens op en heeft uitgesproken vóór- en tegenstanders. In het maatschappelijk debat hebben overtuigingen en op anekdotes gebaseerde persoonlijke sentimenten echter de overhand op robuuste empirische onderzoeksresultaten over wat leerlingen kennen en kunnen op het gebied van rekenen, en over wat de effecten van verschillende rekeninstructiemethoden zijn. Het doel van dit proefschrift is om deze op onderzoek gebaseerde inzichten te verschaffen.

Om een overzicht te geven van wat uit eerder onderzoek bekend is – en van wat *niet* bekend is – over de resultaten van verschillende rekenprogramma's of -curricula wordt als eerste een onderzoekssynthese gepresenteerd van empirische studies die deze vraag probeerden te beantwoorden voor Nederlandse basisschoolleerlingen. Om inzicht te krijgen in de kennis en vaardigheden van leerlingen op het gebied van rekenen zijn de peilingsonderzoeken een rijke bron van gegevens. Deze peilingen zijn echter *surveys* en zijn daardoor beperkt tot beschrijvende analyses. Verklaringen voor gevonden verschillen of veranderingen over tijd vereisen nader onderzoek, en dat is precies wat in het huidige promotieproject is gedaan. Dit proefschrift bevat zes empirische studies, waarin gegevens van in totaal bijna 5000 leerlingen uit groep 3, 4, 5 en 8 zijn geanalyseerd.

In deze empirische studies naar determinanten van rekenvaardigheid op het gebied van optellen, aftrekken, vermenigvuldigen en delen wordt de afstand tussen de wetenschappelijke disciplines van de inhoudelijke psychologie (cognitieve psychologie en onderwijspsychologie) enerzijds en de psychometrie anderzijds overbrugd. Inhoudelijk spelen *oplossingsstrategieën* een hoofdrol in vijf van de zes studies. Strategieën worden belangrijk geacht vanuit de onderwijspsychologie, omdat ze een speerpunt van het realistisch rekenen zijn, maar ook vanuit de cognitieve psychologie waarin mechanismen van strategiekeuze en concepten als strategische competentie (zie bijvoorbeeld Lemaire

en Siegler, 1995) belangrijke onderzoeksgebieden zijn. In de huidige studies werden oplossingsstrategieën als uitkomstvariabelen gebruikt in analyses van determinanten van strategiekeuze, en als verklarende variabelen in analyses van determinanten van rekenprestaties. Gerelateerde terugkerende elementen zijn individuele verschillen in strategiekeuze en in prestaties, en groepsverschillen zoals die tussen jongens en meisjes.

Vanuit een psychometrisch perspectief is het relevant dat de onderzoeksgegevens een complexe structuur hebben, waardoor geavanceerde statistische analyses nodig zijn. In de inhoudelijke disciplines van onderwijspsychologie en cognitieve psychologie zijn dit soort technieken niet erg gebruikelijk. Dat is dan ook de reden waarom dit proefschrift gezien kan worden als een poging tot het samenbrengen van psychologie en psychometrie, zoals bepleit door Borsboom (2006). Een opvallend kenmerk in de huidige studies is dat de data bestaan uit herhaalde observaties: elke leerling beantwoordt meerdere rekenopgaven. Dit leidt tot een gecorreleerde of afhankelijke gegevensstructuur. We betogen dat *latente variabele modellen* geschikt zijn om rekening te houden met deze afhankelijkheden. Eén of meerdere latente variabelen geven de individuele verschillen tussen leerlingen weer, en de afhankelijke responsen binnen elke leerling worden weerspiegeld in een positie op deze variabele(n). Latente variabelen kunnen ofwel categorisch zijn, wanneer ze kwalitatieve individuele verschillen tussen leerlingen modelleren, ofwel continu, wanneer ze kwantitatieve individuele verschillen modelleren. Ook kan de invloed van *verklarende variabelen* (zoals peilingsjaar of het geslacht van de leerling) vastgesteld worden door het effect op de latente variabele te analyseren.

In de onderzoeken zijn de responsen op elke *trial* (wanneer een leerling geconfronteerd wordt met een opgave) van categorisch meetniveau. Meer specifiek zijn er twee typen responsen: de strategie die gehanteerd is om de opgave op te lossen (meerdere ongeordende categorieën) en de goed/fout score van het gegeven antwoord (dichotoom). Om individuele verschillen in strategiekeuze te analyseren gebruiken we latente klasse analyse (LCA; bijvoorbeeld Goodman, 1974; Lazarsfeld & Henry, 1968). In latente klasse modellen wordt aangenomen dat er ongeobserveerde subgroepen (klassen of clusters) van leerlingen zijn die gekarakteriseerd worden door een specifiek patroon van responsen, in dit geval een specifiek patroon van strategiekeuzes. Latente klasse modellen met covariaten (Vermunt & Magidson, 2002) zijn gebruikt om de invloed van leerlingkenmerken op klasselidmaatschap vast te stellen.

Om individuele verschillen in prestatie te analyseren gebruiken we modellen uit

de item-responstheorie (IRT; bijvoorbeeld Embretson & Reise, 2000; Van der Linden & Hambleton, 1997). Hierin wordt de kans op een goed antwoord op een specifieke opgave bepaald door één of meerdere continue latente (vaardigheids)dimensies. Deze IRT-meetmodellen zijn uitgebreid met een verklarend deel, waarin predictoren op persoonsniveau, opgaveniveau of persoons-bij-opgaveniveau worden opgenomen (De Boeck & Wilson, 2004; Rijmen et al., 2003). Een innovatieve toepassing van dergelijke verklarende IRT-modellen wordt gepresenteerd, waarin de strategie die op een opgave is gehanteerd als persoons-bij-opgave predictor wordt opgenomen. Daarmee wordt de strategie-accuratesse gemodelleerd (de kans op een goed antwoord met een bepaalde oplossingsstrategie), statistisch gecorrigeerd voor individuele verschillen in de vaardigheid van de leerlingen die de strategie gebruiken en voor verschillen in moeilijkheidsgraad tussen de opgaven waarop deze strategie gebruikt wordt. Dit was nog niet eerder bewerkstelligd in psychologisch onderzoek naar oplossingsstrategieën.

Hierna volgt een samenvatting per hoofdstuk. Hoofdstuk 1 van dit proefschrift bevat een onderzoekssynthese van resultaten van Nederlandse empirische studies naar de relatie tussen rekendidactiek en rekenvaardigheid. Dit hoofdstuk is gebaseerd op literatuuronderzoek dat is uitgevoerd voor de adviescommissie *Rekenonderwijs op de basisschool*¹ ingesteld door de *Koninklijke Nederlandse Akademie van Wetenschappen* (KNAW), wier rapport in 2009 is uitgekomen. Deze systematische kwantitatieve onderzoekssynthese laat geen eenduidige conclusies over het effect van verschillende rekeninstructiemethoden of rekencurricula toe. Enerzijds zijn er weinig methodologisch degelijk opgezette *interventiestudies* waarin de effecten van verschillende instructie-aanpakken vergeleken worden. De wel beschikbare studies zijn bovendien beperkt in verschillende aspecten, zoals steekproefgrootte of inhoudsdomein. Ook zijn didactische kenmerken en instructiekenmerken vaak met elkaar verweven in de programma's die vergeleken zijn, zodat het onmogelijk is de unieke effecten van verschillende kenmerken vast te stellen. Anderzijds zijn de *curriculumstudies* waarin de uitkomsten van leerlingen die verschillende rekencurricula (rekenmethodes) gevolgd hebben worden vergeleken, beperkt in de mate van controle over de implementatie van het curriculum en in de mogelijkheid tot het corrigeren voor verstorende variabelen.

Hoewel er dus geen algemene hoofdconclusie getrokken kan worden, zijn er wel wat

¹ Ik heb als toegevoegd onderzoeker voor deze commissie gewerkt, en heb in opdracht van de commissie het systematische literatuuronderzoek dat ten grondslag ligt aan hoofdstuk 4 van het rapport (KNAW, 2009) uitgevoerd. Dat vormt de basis voor hoofdstuk 1 van dit proefschrift.

specifieke patronen die uit de bestudeerde onderzoeksresultaten naar voren komen. Ten eerste is het opvallend dat de prestatieverschillen *binnen* een bepaald type instructie-aanpak groter zijn dan *tussen* verschillende aanpakken. Blijkbaar spelen didactische principes een kleinere rol dan de praktische implementatie door de leerkracht en de interactie tussen de leerkracht en de leerling, bevindingen die in overeenstemming zijn met die van bijvoorbeeld Slavin en Lake (2008) in hun grootschalige internationale onderzoekssynthese. Een tweede patroon is dat meer onderwijstijd voor rekenen betere rekenprestaties tot gevolg heeft, wat in overeenstemming is met de stelling van Hiebert en Grouws (2007) dat *opportunity to learn* the belangrijkste voorspeller van prestatie-uitkomsten is. Ten derde is gebleken dat, als de onderwijstijd gelijk is, experimentele rekeninstructieprogramma's die in kleine groepjes leerlingen buiten de klas worden uitgevoerd positieve effecten op de prestaties hebben ten opzichte van de reguliere klassepraktijk, vergelijkbaar met het gerapporteerde positieve effect van *small-group tutoring* door Slavin en Lake (2008). Ten slotte zijn veel van de bestudeerde onderzoeken gericht op zwakke rekenaars. Deze leerlingen lijken minder gebaat bij een vrije vorm van instructie en meer behoefte te hebben aan een sturende rol van de leerkracht in hun leerproces, in overeenstemming met conclusies van internationale reviews (Gersten et al., 2009; Kroesbergen & Van Luit, 2003; Swanson & Carson, 1996; Swanson & Hoskyn, 1998). Veel minder onderzoeks aandacht is uitgegaan naar instructie voor middelmatige en sterke rekenaars. Ook mogelijke differentiële instructie-effecten voor jongens en meisjes zijn weinig onderzocht.

In de overige zes empirische hoofdstukken is de aandacht verlegd van instructie-effecten naar andere aspecten van de rekenvaardigheid van basisschoolleerlingen in het hedendaagse reken-wiskundeonderwijs. In hoofdstuk 2 en 3 zijn secundaire analyses uitgevoerd op het ruwe leerlingmateriaal (de ingevulde opgaveboekjes) van ongeveer 1600 leerlingen die deelnamen aan één van de twee meest recente peilingen in groep acht. Deze analyses hadden als doel meer inzicht te krijgen in de vaardigheid op het gebied van het vermenigvuldigen en delen met grote getallen en kommagetallen door oplossingsstrategieën in beschouwing te nemen. Het prestatieniveau van de groep 8-leerlingen op deze bewerkingen bleek tussen PPON 1997 en 2004 sterk gedaald, en bovendien ver achter te blijven bij de standaarden. Voor beide bewerkingen zijn de strategieën die leerlingen gebruikten om de opgaven op te lossen gecodeerd in vier hoofdcategoryën: het traditionele algoritme, ook wel cijferen genoemd (de staartdeling en het onder elkaar vermenigvuldigen), niet-traditionele strategieën die met gehele

getallen werken, het geven van een antwoord zonder een schriftelijke uitwerking (complete berekening of tussenstappen) te noteren en overige strategieën (onduidelijke of verkeerde aanpakken en overgeslagen opgaven). Een subcategorie van de niet-traditionele strategieën is het zogenaamde *kolomsgewijs rekenen*, een onderdeel van de didactiek van het realistisch rekenen (Van den Heuvel-Panhuizen, 2008; Van den Heuvel-Panhuizen et al., 2001). Het kolomsgewijs rekenen kan als overgang tussen informele aanpakken en het traditionele algoritme gezien worden: er wordt gewerkt met gehele getallen in plaats van individuele cijfers (zoals in informele strategieën) maar de procedure is min of meer gestandaardiseerd (zoals in de traditionele algoritmen).

In hoofdstuk 2 worden ook de latente variabele modellen geïntroduceerd. Latente klasse analyses zijn gebruikt om individuele verschillen in keuze van strategieën in kaart te brengen en om verschuivingen over tijd zichtbaar te maken. Verklarende item-responsmodellen zijn gebruikt om strategie-accuratesse en verschuivingen daarin zo zuiver mogelijk te analyseren. Dit hoofdstuk richt zich inhoudelijk op de oplossingsstrategieën op de deelsommen van PPON 1997 en 2004. Het blijkt dat twee veranderingen hebben bijgedragen aan de prestatiedaling tussen deze twee peilingen. Ten eerste heeft een verschuiving in strategiegebruik plaatsgevonden: in 2004 waren minder leerlingen geneigd de staartdeling te gebruiken dan in 1997, terwijl meer leerlingen geneigd waren deze complexe deelopgaven zonder uitwerking (vermoedelijk uit het hoofd) uit te rekenen. Verrassend genoeg is het percentage leerlingen dat kolomsgewijs deelde – wat het eindpunt is van hedendaagse leerlijnen en daarmee de vervanging van de staartdeling in het rekenonderwijs – nagenoeg gelijk gebleven. Uit de analyses van de accuratesse van elke strategie blijkt dat het antwoorden zonder uitwerking een stuk minder succesvol was dan de beide schriftelijke strategieën (traditionele en niet-traditionele aanpakken). Deze strategieverschuiving, die overigens vooral op het conto van jongens geschreven moet worden, heeft dus een ongunstig effect op de prestaties gehad. Ten tweede blijkt dat er nog een deel van de prestatiedaling onverklaard blijft na statistische correctie voor deze ongunstige verschuivingen in strategiegebruik. Dit betekent dat de kans op een goed antwoord met een bepaalde strategie op een bepaalde opgave in 2004 lager lag dan in 1997. Deze algemene daling van de prestaties binnen elke strategie is zorgelijk.

De resultaten op het gebied van vermenigvuldigen zijn tot op zekere hoogte vergelijkbaar met die van delen, zoals blijkt uit hoofdstuk 3. Ook bij vermenigvuldigen is een verschuiving in strategiegebruik gevonden, waarbij er in 2004 minder leerlingen waren die vooral traditioneel cijferend vermenigvuldigden en meer leerlingen die zonder

schriftelijke uitwerking rekenden (opnieuw vooral jongens). Deze toename in het uit het hoofd rekenen is bij vermenigvuldigen echter van kleinere omvang dan bij delen, en wordt bovendien aangevuld met een toename in het gebruik van niet-traditionele strategieën. Ook was het traditionele vermenigvuldigingsalgoritme in beide peilingsjaren de meest voorkomende strategie, wat in overeenstemming is met het feit dat het algoritme nog wel het eindpunt in de leerlijn voor vermenigvuldigen is (dit is dan ook in tegenstelling tot de leerlijn voor delen). Het is daarmee dan ook opvallend te noemen dat juist bij vermenigvuldigen het gebruik van niet-traditionele strategieën is toegenomen over tijd terwijl het bij delen juist constant is gebleven; het tegengestelde patroon had meer in de lijn der verwachting gelegen.

Net zoals bij delen zijn ook bij vermenigvuldigen de schriftelijke strategieën succesvoller dan het antwoorden zonder uitwerking. Daarentegen blijken bij vermenigvuldigen de niet-traditionele strategieën minder succesvol dan het traditionele algoritme, terwijl bij delen deze strategieën ongeveer even succesvol waren (behalve voor de gemiddelde rekenaars). De afname van het traditioneel cijferend vermenigvuldigen ten koste van de toename in niet-traditionele strategieën en het hoofdrekenen hebben dus ook bij vermenigvuldigen een ongunstig effect op de prestaties gehad. In hoofdstuk 3 is ook de invloed van de strategie-instructie door de leerkracht op het strategiegebruik van leerlingen op vermenigvuldigen en delen in PPON-2004 onderzocht. Vooral het gebruik van de staartdeling wordt beïnvloed door de instructie van de leerkracht: vrijwel uitsluitend leerlingen van wie de leerkracht de staartdeling onderwees (en daarmee dus afweek van de rekenmethode) gebruikten de staartdeling. Bij vermenigvuldigen zijn de trends vergelijkbaar, maar wel een stuk minder uitgesproken.

Uit de resultaten van hoofdstuk 2 en 3 is gebleken dat – vooral bij delen – het meest relevante onderscheid in oplossingsstrategieën dat tussen schriftelijke strategieën en hoofdrekenen is. Daarom zijn in de volgende twee empirische studies nieuwe gegevens verzameld om de kenmerken van deze strategieën (zoals de verdeling over de opgaven, accuratesse, snelheid, en adaptiviteit van strategiekeuzes) voor de bewerking delen op een zuivere manier vast te stellen. Daartoe is een gedeeltelijke (hoofdstuk 4) en een complete (hoofdstuk 5) *choice/no-choice* studie (zie Siegler & Lemaire, 1997) opgezet. In het onderzoek in hoofdstuk 4 maakten 362 leerlingen uit groep 8 twee parallelle sets van elk vier deelopgaven (bijvoorbeeld $782 \div 32$) onder twee condities: eerst de ene set in een conditie waarbij ze zelf mochten kiezen of ze een schriftelijke strategie gebruikten of uit het hoofd rekenden (de *choice* of vrije keuze conditie), en vervolgens

de andere set in een conditie waarbij ze verplicht waren schriftelijk te rekenen (de *no-choice* of verplicht schriftelijke conditie). In de vrije keuze conditie waren ook nog vijf andere deelopgaven opgenomen die makkelijker getalskenmerken hadden (bijvoorbeeld $3240 \div 4$), en waarvan daarom verwacht werd dat ze uit het hoofd uit te rekenen waren.

Latente klasse analyse op de strategiekeuzes in de vrije keuze conditie liet zien dat er drie subgroepen van leerlingen onderscheiden kunnen worden: zij die voornamelijk uit het hoofd rekenden (meer jongens dan meisjes), zij die voornamelijk schriftelijk rekenden (meer meisjes dan jongens), en zij die van strategie wisselden door op de moeilijkere opgaven voornamelijk schriftelijk te rekenen en op de makkelijker opgaven voornamelijk uit het hoofd te rekenen. Deze laatste groep leerlingen vertoonde dus een adaptief en flexibel patroon van strategiekeuzes. Opvallend is dat in deze groep vrijwel geen zwakke rekenaars zaten. Opnieuw blijken spontaan gekozen schriftelijke strategieën succesvoller dan spontaan gekozen hoofdrekenden, net als in hoofdstuk 2. Daarnaast is vastgesteld dat de prestaties van leerlingen die als ze zelf mochten kiezen hoofdrekenden op een bepaalde opgave, verbeterden als ze gedwongen werden schriftelijk te rekenen. Voor leerlingen die in de vrije keuze conditie al schriftelijk rekenden had het gedwongen schriftelijk rekenen geen effect op de prestaties. Het blijkt dus dat schriftelijke strategieën succesvoller zijn dan hoofdrekenden, niet alleen in vergelijkingen tussen opgaven en leerlingen, maar ook binnen dezelfde opgave en dezelfde leerling. Een hieruit volgende aanbeveling is dan ook dat leerlingen meer aangespoord moeten worden om tussenuitkomsten of complete berekeningen te noteren bij het oplossen van deelopgaven met grote getallen en kommagetallen. Dit lijkt vooral van belang voor jongens, die meer geneigd zijn tot hoofdrekenden dan meisjes, en voor zwakke rekenaars, voor wie een groter verschil in accuratesse tussen de twee strategieën is vastgesteld dan voor gemiddelde en betere rekenaars.

Waarom kiezen leerlingen voor het riskante hoofdrekenden? Deze vraag komt aan bod in hoofdstuk 5, waarin het onderzoek van hoofdstuk 4 op twee manieren wordt aangevuld. Ten eerste is er een derde onderzoeksconditie opgenomen waarin leerlingen verplicht moesten hoofdrekenden, zodat zuiver vastgesteld kan worden hoe leerlingen met deze strategie presteren. Ten tweede is er een tweede prestatie-uitkomst opgenomen: naast het vaststellen of het goede of foute antwoord is gegeven, is ook de tijd opgenomen die leerlingen nodig hadden om tot hun antwoord te komen. Daardoor is het mogelijk niet alleen de accuratesse van beide strategieën te analyseren, maar ook de snelheid in de analyses mee te nemen. Zesentachtig leerlingen uit groep acht maakten dan ook drie

sets van vier opgaven in drie verschillende condities: de vrije keuze conditie, de verplicht hoofdrekenen conditie en de verplicht schriftelijk rekenen conditie.

De resultaten laten zien dat hoofdrekenen vooral gekozen wordt omdat het sneller is dan schriftelijk rekenen, terwijl schriftelijk rekenen vooral werd gekozen omdat het accurater is. In dit onderzoek kan ook de adaptiviteit van strategiekeuzes worden vastgesteld: in hoeverre kiest een leerling op een opgave de strategie die voor hem of haar het beste is? Dit is gerelateerd aan het concept van *adaptive expertise* (Baroody & Dowker, 2003; Torbeyns, De Smedt, et al., 2009b). De resultaten laten zien dat een substantieel deel van de leerlingen niet hun beste strategie – gedefinieerd als de strategie die het snelst tot het goede antwoord leidt – koos. Deze suboptimale strategiekeuzes spreken aannames in cognitieve modellen van strategiekeuze tegen (zoals Shrager & Siegler, 1998; Siegler & Shipley, 1995). Een mogelijke verklaring hiervoor is dat deze modellen niet expliciet zijn over de invloed van individuele verschillen in de voorkeur voor snelheid en accuratesse (Phillips & Rabbitt, 1995), die ervoor kunnen zorgen dat sommige leerlingen het snelle maar foutgevoeliger hoofdrekenen kiezen. Ook de invloed van de socioculturele context wordt onderbelicht in de cognitieve modellen (Ellis, 1997; Luwel et al., 2009; Verschaffel et al., 2009). In de klas kan het bijvoorbeeld zo zijn dat hoofdrekenen meer gewaardeerd wordt dan schriftelijk rekenen.

Verder zijn er interessante verschillen tussen jongens en meisjes gevonden. Opnieuw blijken jongens meer geneigd tot hoofdrekenen dan meisjes, maar er is nu gevonden dat dat gerelateerd lijkt te zijn aan verschillen in de relatieve voorkeur voor accuratesse en snelheid. Meisjes kiezen hun strategie op basis van de accuratesse, daarmee snelheid negerend, terwijl de strategiekeuze van jongens meer gebaseerd is op snelheidsoverwegingen dan op accuratesse. Ook het algemene rekenniveau van de leerling is van invloed. Naast het weinig verrassende patroon dat bovengemiddelde rekenaars beter presteerden, een hogere strategie-accuratesse hadden en met elke strategie sneller rekenden dan ondergemiddelde rekenaars, zijn ook verschillen in de adaptiviteit van strategiekeuzes gevonden. Ondergemiddelde rekenaars pasten hun strategiekeuze zowel niet aan snelheid als aan accuratesse aan, terwijl bovengemiddelde rekenaars hun keuze aan beide componenten aanpasten.

De onderzoeken beschreven in de voorgaande hoofdstukken zijn merendeels gebaseerd op rekenopgaven in een realistische context: een meestal verbale beschrijving van een mathematisch probleem, vaak aangevuld met een illustratie. Dit type opgaven wordt veel gebruikt in het hedendaagse reken-wiskundeonderwijs (zie bijvoorbeeld

Gravemeijer & Doorman, 1999; Verschaffel et al., 2000) en in de rekentoetsen, zoals Cito's Leerling- en Onderwijs Volg Systeem (LOVS) en PPON. De vraag die gesteld kan worden is in hoeverre de dominantie van contextopgaven effect heeft op het toetsen van rekenvaardigheden. Dit is in twee studies onderzocht: in hoofdstuk 6 voor de prestaties van leerlingen in groep drie, vier en vijf, en in hoofdstuk 7 voor de prestaties en strategiegebruik van leerlingen in groep acht. De bevindingen zijn gemengd.

In hoofdstuk 6 wordt een onderzoek beschreven waarin ruim 2000 leerlingen in de onderbouw van het basisonderwijs naast de reguliere LOVS-taken voor hun jaargroep – die voornamelijk uit contextopgaven bestaan – een extra taak met alleen 'kale' getalsopgaven maakten. Met meerdimensionale IRT-modellen is vastgesteld dat twee samenhangende maar wel verschillende vaardigheden ten grondslag liggen aan het oplossen van deze twee typen opgaven. Dit betekent dat leerlingen die goed zijn in het oplossen van opgaven van het ene type niet direct ook goed zijn in het oplossen van het andere type, en dat een profiel van relatieve prestaties nuttig zou kunnen zijn voor de leerkracht. De samenhang lijkt wel toe te nemen met jaargroep, met een latente correlatie van .81 in groep drie oplopend naar .87 in groep vijf. Verder hebben indicatoren van het taalniveau van de leerling (thuis taal en niveau van begrijpend lezen) een verschillend effect op beide vaardigheden. In zijn algemeenheid zijn de effecten van taalniveau groter op de vaardigheid van het oplossen van contextopgaven dan op het oplossen van kale getalsopgaven. De prestatieachterstand van leerlingen die thuis geen Nederlands spreken ten opzichte van zij die dat wel doen is op de contextopgaven een stuk groter dan op de kale getalsopgaven. Dit betekent dat meer balans in de toetsen tussen contextopgaven en kale getalsopgaven het gat tussen leerlingen die thuis geen Nederlands spreken en zij die dat wel doen waarschijnlijk zal verkleinen. Een beperking van deze studie is echter dat het effect van een context in een rekenopgave niet rechtstreeks kan worden vastgesteld, omdat de twee typen opgaven niet vergelijkbaar waren qua getalskenmerken.

Daarom is een ander onderzoek verricht waarin dit wel mogelijk was, ditmaal in groep acht (hoofdstuk 7). In totaal 685 leerlingen maakten een set opgaven op het gebied van optellen, aftrekken, vermenigvuldigen en delen met grote getallen en kommagetallen. Er waren acht paren van opgaven: elk paar bestond steeds uit een contextopgave (ontleend aan PPON) en de kale getalsversie van die opgave: de rekenkundige bewerking, ontdaan van de context. Het effect van de contexten is vastgesteld op twee aspecten van probleemoplossen: de prestaties en het strategiegebruik. In tegenstelling tot in de onderbouw (hoofdstuk 6) blijkt er in groep acht slechts één latente vaardigheidsdimensie

ten grondslag te liggen aan het oplossen van de twee typen opgaven. Verder blijkt dat de contexten de moeilijkheidsgraad van een opgave nauwelijks beïnvloeden: alleen bij de deelopgaven was de opgavevariant met context iets moeilijker dan die zonder context. Ook de strategiekeuze en de strategie-accuratesse worden niet beïnvloed door de aanwezigheid van een context. Een verdere belangrijke bevinding is dat dit alles geldt voor jongens en voor meisjes, voor leerlingen die thuis Nederlands spreken en zij die een andere taal spreken, en voor leerlingen met een verschillend niveau van begrijpend lezen.

Samenvattend zijn dus in groep acht nauwelijks tot geen effecten gevonden van de aanwezigheid van contexten in rekenopgaven op verschillende aspecten van het probleemoplossen. Dit is tegenstelling tot de verwachtingen op basis van hoofdstuk 6 en ook tot vaak gehoorde overtuigingen. In combinatie met de onderzoeksresultaten bij jongere kinderen lijken de bevindingen van hoofdstuk 7 erop te wijzen dat de invloed van contexten in rekentoetsen in de loop van de basisschooltijd afneemt en zelfs verdwijnt.

Ter afsluiting kan geconcludeerd worden dat de in dit proefschrift gerapporteerde onderzoeken elk meer inzicht hebben verschaft in hoe basisschoolleerlingen rekenen in het hedendaagse reken-wiskundeonderwijs. Zonder geavanceerde psychometrische technieken, in het bijzonder latente variabele modellen waarin verklarende variabelen zijn opgenomen, was het niet mogelijk geweest deze resultaten en inzichten te verkrijgen. Het samenbrengen van de psychometrie en de psychologie heeft dan ook zijn nut en noodzaak bewezen.

