



KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN

REKENONDERWIJS OP DE BASISCHOOL

ANALYSE EN SLEUTELS
TOT VERBETERING



ADVIES


REKENONDERWIJS OP DE BASISCHOOL

Koninklijke Nederlandse Akademie van Wetenschappen
Postbus 19121, 1000 GC Amsterdam
Telefoon + 31 20 551 0700
Fax + 31 20 620 4941
knaw@bureau.knaw.nl
www.knaw.nl

Basisvormgeving: edenspiekermann, Amsterdam
Opmaak: Ellen Bouma, Alkmaar
Druk: Bejo druk & print, Alkmaar

Foto cover: Nationale beeldbank/Naomi den Hertog

ISBN: 978-90-6984-600-2

Het papier van deze uitgave voldoet aan  iso-norm 9706 (1994) voor permanent houdbaar papier.

Dit advies is gemaakt van FSC-papier en gecertificeerd onder nummer CU-COC-804134-N.



© 2009 Koninklijke Nederlandse Akademie van Wetenschappen (KNAW)
Niets uit deze uitgave mag worden verveelvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, via internet of op welke wijze dan ook, zonder voorafgaande schriftelijke toestemming van de rechthebbende, behoudens de uitzonderingen bij de wet gesteld.

REKENONDERWIJS OP DE BASISCHOOL

ANALYSE EN SLEUTELS TOT VERBETERING

Koninklijke Nederlandse Akademie van Wetenschappen
Advies KNAW-Commissie rekenonderwijs basisschool

TEN GELEIDE

De staartdeling is een regelmatige gast op de opiniepagina. Hoe kan een technisch aspect van de rekendidactiek op de basisschool zulke heftige gevoelens oproepen? Een deel van het antwoord op deze vraag is dat het scherpe debat van de afgelopen jaren over het rekenonderwijs aan een breder thema raakt. Op verschillende plaatsen in het onderwijs zijn basisvaardigheden van leerlingen en docenten onder druk komen te staan en wordt er een gebrek gevoeld aan belangstelling voor de inhoud van het vak, zowel in het leslokaal als op de lerarenopleidingen. Dit alles is een grote maatschappelijke zorg.

Wat heeft een geleerd gezelschap als de Koninklijke Nederlandse Akademie van Wetenschappen met primaire rekendidactiek van doen? Er zijn minstens twee aspecten van deze problematiek die de kern van de KNAW als stem, geweten en forum van de wetenschap raken.

Allereerst zijn er de lange lijnen van het onderwijs. Het is maar al te gemakkelijk om vanuit een hiërarchisch perspectief naar het onderwijs te kijken. De verschillende onderwijsniveaus zijn echter nauw met elkaar verbonden. Een zwak fundament doet het gehele gebouw wankelen. De afgelopen jaren heeft de KNAW zich daarom nadrukkelijk met de volle breedte van het onderwijs verbonden. Initiatieven als het rapport *Robuuste profielen in het voortgezet onderwijs* (2003) en sinds kort de KNAW Onderwijsprijs bewijzen dit. Op dit moment speelt de Akademie een voorhoederol bij het behartigen van het belang van de wetenschap op de basisschool, in lijn met de acties van een aantal buitenlandse zusteracademies. Oud-presidenten Pim Levelt en Frits van Oostrom hebben meermaals de fundamentele rol van de goed toegeruste en gemotiveerde docent benadrukt.

Het tweede aspect is de bijdrage die de wetenschap zelf kan leveren aan discussies over onderwijsmethoden. Zijn er wetenschappelijke gegevens waarop scholen en

docenten een weloverwogen keuze voor een bepaalde didactiek kunnen baseren? Of, in dit geval: zijn er cijfers over het rekenen? Zo ja, wat kunnen we daaruit concluderen? Zo nee, zouden we die data dan niet eerst moeten verzamelen?

Tegen deze achtergrond en daartoe aangemoedigd door het ministerie van Onderwijs, Cultuur en Wetenschap heeft de Akademie in 2008 een commissie geïnstalleerd onder leiding van Jan Karel Lenstra, directeur van het Centrum Wiskunde & Informatica in Amsterdam. De KNAW dankt de commissie, en in het bijzonder haar voorzitter, voor de grote inspanningen die zij heeft geleverd door een studie naar dit controversiële onderwerp te verrichten op basis van een grondige analyse van inhoudelijke inzichten en empirisch feitenmateriaal.

De belangrijkste conclusie die de commissie trekt, is dat de verkeerde discussie is gevoerd en dat dit ons ervan heeft weerhouden tot de kern van de zaak te komen. Zij constateert dat de wetenschappelijke kennis van de effectiviteit van rekendidactieken beperkt is en dat relevante data grotendeels ontbreken. De commissie vraagt aandacht voor het scherper monitoren en voor meer vergelijkend onderzoek. Zij signaleert dat de diversiteit aan lesmethoden de afgelopen jaren is verminderd. Zorgwekkend, want monoculturen zijn een verarming van het ecosysteem, zo waarschuwen biologen ons.

Maar de scherpste conclusie van deze studie is dat de kwaliteit van het rekenonderwijs op de lerarenopleiding ernstig onder druk is komen te staan. Dat is een zorgelijke ontwikkeling die onmiddellijke aandacht en actie vraagt. Het niveau van de gehele onderwijsketen wordt gedragen door de kwaliteit van de docent. De KNAW hoopt dan ook dat dit rapport ertoe leidt dat de opleiding en de nascholing van leerkrachten op het terrein van het rekenonderwijs grondig tegen het licht wordt gehouden. De docent is de spil in het onderwijs en moet kunnen rekenen op het allerbeste gereedschap.

Robbert Dijkgraaf
President KNAW

INHOUD

SAMENVATTING 9

SUMMARY 10

1. INLEIDING 17

1.1 Aanleiding voor dit advies 17

1.2 Opdracht en samenstelling van de commissie 18

1.3 Werkwijze van de commissie 18

1.4 Opzet van het rapport 19

2. ONTWIKKELINGEN IN HET REKENONDERWIJS 21

2.1 Een halve eeuw geschiedenis 21

2.2 Wat is traditioneel rekenen? 23

2.3 Wat is realistisch rekenen? 24

2.4 Traditioneel versus realistisch rekenen: discussie 25

2.5 Samenvatting en conclusies 27

3. REKENVAARDIGHEID 29

3.1 Nationaal en internationaal empirisch onderzoek 29

3.2 PPO: Periodieke peiling van het onderwijsniveau 30

3.3 TIMSS: Trends in international mathematics and science study 34

3.4 Commissie Meijerink: Over de drempels met rekenen 39

3.5 Samenvatting en conclusies 41

4. RELATIE TUSSEN REKENDIDACTIEK EN REKENVAARDIGHEID 43

4.1 Gevolgde werkwijze 43

4.2 Internationale literatuur 46

4.3 Interventiestudies 56

4.4 Curriculumstudies 68

4.5 Samenvatting, conclusies en aanbevelingen 74

5.	ONTWIKKELING IN HET AANBOD VAN DE REKENMETHODEN	79
5.1	Een markt in beweging	79
5.2	Conclusie en aanbevelingen	80
6.	ROL, OPLEIDING EN NASCHOLING VAN DE LERAAR	83
6.1	Rol van de leraar	83
6.2	Opleiding: de situatie op de pabo	85
6.3	Nascholing en begeleiding	86
6.4	Samenvatting, conclusies en aanbevelingen	88

LITERATUUR 91

BIJLAGEN

1	Geïnterviewde deskundigen	95
2	Berekening effectgrootte	97
3	Tabellen met effectgroottes	98

SAMENVATTING

Bezorgdheid over de rekenvaardigheid van kinderen heeft de laatste jaren geleid tot een publieke discussie over het rekenonderwijs in ons land. Daarin staan de aanhangers van de ‘traditionele’ en de ‘realistische’ rekendidactiek tegenover elkaar. Het debat werkt polariserend en lijkt slechts in geringe mate gevoerd te worden op basis van wetenschappelijk gefundeerde kennis over de zaken die ter discussie staan. De KNAW besloot daarom een commissie ‘Rekenonderwijs op de basisschool’ in te stellen. Toen de staatssecretaris van OCW, mevrouw Dijkma, een studie naar rekenmethodieken aankondigde, werd besloten beide initiatieven te combineren.

De commissie had de volgende opdracht: *Breng in kaart wat er bekend is over de relatie tussen rekendidactiek en rekenvaardigheid op grond van bestaande inhoudelijke inzichten en empirisch feitenmateriaal. Geef daarbij aan hoe ruimte kan worden geschaffen voor leraren en ouders om te kiezen op basis van informatie over deze relatie tussen rekendidactiek en effect.*

Naast de vraag wat er feitelijk bekend is over de relatie tussen rekendidactiek en rekenvaardigheid, heeft de commissie zich ook de vraag gesteld hoe het staat met de rekenvaardigheid van kinderen. De antwoorden op deze vragen zijn gebaseerd op wetenschappelijke publicaties. De verkenning door de commissie van de ruimte die er is om te kiezen tussen rekendidactieken en, in bredere zin, om het rekenonderwijs te verbeteren is gebaseerd op interviews die de commissie heeft gehouden en op feitenmateriaal, voor zover dat binnen het gegeven tijdsbestek kon worden verzameld.

HOOFDCONCLUSIES

1. De bezorgdheid over de rekenvaardigheid van basisschoolleerlingen is op zijn plaats. Nederland dreigt zijn sterke internationale positie te verliezen. Achteruitgang bij bewerkingen met grotere getallen en kommagetallen wordt niet gerechtvaardigd door vooruitgang bij onderdelen als getalbegrip en schattend rekenen. Het rekenpeil kan en moet over de gehele linie omhoog.
2. Het publieke debat overdrijft de tegenstelling tussen de traditionele en de realistische rekendidactiek en gaat bovendien over het verkeerde onderwerp, namelijk een vermeend verschil in het effect van beide didactieken. Er is geen overtuigend verschil aangetoond.
3. De sleutel tot verbetering van de rekenvaardigheid ligt in het niveau van de leraar. De opleiding en nascholing van de leraar zijn in ernstige mate geërodeerd. Het Ministerie van OCW dient de pabo-opleiding aan een grondig onderzoek te onderwerpen en nascholing in rekenvaardigheid en rekendidactiek krachtig te stimuleren.

SUMMARY

Growing concern about Dutch children's mathematical proficiency has led in recent years to a public debate about the way mathematics is taught in the Netherlands. There are two opposing camps: those who advocate teaching mathematics in the "traditional" manner, and those who support "realistic" mathematics education. The debate has had a polarizing effect and appears to have little basis in scholarly research. The Royal Netherlands Academy of Arts and Sciences therefore decided to set up a Committee on Primary School Mathematics Teaching. When the State Secretary for Education, Culture and Science, Ms Sharon Dijksma, announced a study on mathematics education, these two initiatives were combined.

The Committee's mission was the following: *To survey what is known about the relationship between mathematics education and mathematical proficiency based on existing insights and empirical facts. Indicate how to give teachers and parents leeway to make informed choices, based on our knowledge of the relationship between approaches to mathematics teaching and mathematical achievement.*

In addition to surveying what we know about the relationship between teaching approaches and mathematical achievement, the Committee also looked into the state of children's mathematical proficiency. The answers to these questions are based on scholarly publications. The Committee's investigation of the leeway available to choose between various teaching approaches and, in the broader sense, to improve mathematics education is based on interviews that it conducted and on factual material, in so far as it was possible to collect it within the given time frame.

MAIN CONCLUSIONS

1. There is good reason to be worried about the mathematical proficiency of Dutch primary school pupils. The Netherlands is at risk of losing its strong international position. The progress made in such areas as number sense and estimation does not justify the decline in children's ability to perform mathematical operations with large numbers and decimal numbers. Children's mathematical proficiency needs improvement.
2. The public debate exaggerates the differences between the traditional and realistic approaches to mathematics teaching. It also focuses erroneously on a supposed difference in the effect of the two instructional approaches whereas in fact, no convincing difference has been shown to exist.
3. The key to improving mathematical proficiency lies in the teacher's competences. Teacher training and post-graduate courses have been seriously undermined. The Ministry of Education, Culture and Science should subject teaching training programmes to a thorough investigation and encourage post-graduate training in mathematics and mathematics teaching.

ONTWIKKELINGEN IN HET REKENONDERWIJS

Het primaire onderwijs in Nederland heeft de laatste vijftig jaar ingrijpende veranderingen ondergaan. De oprichting van basisschool en pabo, de formulering van kerndoelen, de invoering van het zelfstandig werken en de goedkope rekenmachine hebben een grote invloed uitgeoefend.

In het rekenonderwijs was de belangrijkste wijziging de invoering van de realistische rekendidactiek. Deze geeft een prominente rol aan contexten die eigen oplossingsstrategieën van de leerlingen uitlokken; interactie hierover en reflectie hierop onder leiding van de leraar leiden dan tot de opbouw van inzicht, kennis en vaardigheid. In de traditionele rekendidactiek staat het systematisch aanleren van één standaardalgoritme per bewerking centraal, via een directere sturing van de leerling en voornamelijk aan de hand van contextloze opgaven.

Traditioneel en realistisch rekenonderwijs zijn echter geen eenduidige begrippen. De gedaanten van een didactiek variëren van ideëel concept via methode in het boekje en perceptie van de leraar tot de praktijk in de klas. Anno 2009 zijn sommige aspecten van realistisch rekenen minder goed ingevoerd dan de ontwikkelaars voor ogen stond; traditioneel rekenen zal kerndoelen die verder reiken dan de beheersing van één standaardalgoritme per bewerking niet kunnen negeren. Een realistische aanpak stelt wellicht hogere eisen aan de leraar dan een traditionele.

REKENVAARDIGHEID

In nationale peilingen zijn de rekenprestaties de afgelopen twintig jaar op veel onderdelen vrij stabiel gebleven. Tegenover een achteruitgang op sommige onderdelen staat een vooruitgang op andere. Internationaal neemt Nederland nog steeds een sterke positie in, maar deze kalf af.

De rekenvaardigheid van de Nederlandse leerlingen stemt de commissie niet tot tevredenheid. Positieve en negatieve ontwikkelingen mag men niet tegen elkaar wegstrepen. Er zijn veel aspecten waarbij het prestatiepeil en de ontwikkeling daarvan onvoldoende zijn. Ook bij onderdelen waar vooruitgang is geboekt, zoals getalbegrip en schattend rekenen, blijft het peil ver achter bij de gestelde standaard. Bij bewerkingen met grotere getallen en kommagetallen is het peil sterk gedaald, deels door een keuze voor uit het hoofd rekenen waar schriftelijk rekenen geboden is. Dit mag niet worden gerelativeerd door te wijzen op de rekenmachine. Te weinig leerlingen bereiken een geavanceerd niveau. Nederland blijft achter bij de Aziatische landen en verliest terrein aan andere West-Europese landen.

RELATIE TUSSEN DIDACTIEK EN REKENVAARDIGHEID

De studie door de commissie naar de relatie tussen rekendidactiek en rekenvaardigheid heeft zich geconcentreerd op empirisch onderzoek dat de laatste twintig jaar in Nederland is verricht, maar omvat ook een beknopte inventarisatie van buitenlands onderzoek.

De conclusie is dat dit materiaal geen eenduidig beeld oplevert en geen algemene wetenschappelijke uitspraken rechtvaardigt over de relatie tussen rekendidactiek en

TRENDS IN MATHEMATICS EDUCATION

Primary education in the Netherlands has changed seriously in the past fifty years. Various trends and developments have had a major impact: the establishment of new-style primary schools and teacher training colleges; the decision to define educational standards; the introduction of self-directed learning; and the widespread availability of inexpensive calculators. The most important change in mathematics teaching was the introduction of realistic mathematics education (RME). In RME, the emphasis is on providing contexts that elicit pupils' own solution strategies; teacher-guided interaction and reflection then enable the pupils to build mathematical knowledge, understanding and skills. Traditional mathematics education, on the other hand, has pupils systematically acquiring one standard algorithm per operation. The teacher provides more direct instruction and pupils learn mainly by solving problems presented without any context.

Traditional and realistic mathematics education are by no means simple concepts, however. Teaching approaches take on many different forms, from an ideal curriculum, through an instructional approach as described in a teaching guide and teachers' own perceptions of this approach, to the actual classroom practices. On the one hand, some aspects of RME have not taken hold as its developers had envisaged; on the other, traditional mathematics education cannot ignore educational standards that go beyond mastering a single standard algorithm per operation. RME may well put heavier demands on teachers than the traditional approach.

MATHEMATICAL PROFICIENCY

National assessments have shown that in many respects, children's mathematics achievement has remained fairly stable in the past twenty years. While achievement has declined on some aspects, it has improved in others. There are signs that the Netherlands' strong international position is foundering.

The Committee is not satisfied with the current level of mathematical proficiency of Dutch pupils. Positive and negative trends should not cancel each other out. There are several areas in which the level of performance and the overall trend in that level are insufficient. Even in those areas where progress has been made – for example number sense and estimation – pupil's achievements are far below the required standard. Mathematical proficiency has declined notably with respect to performing operations with large numbers and decimal numbers, in part because pupils choose a mental solution strategy, while they would have been better of working out their calculations on paper. The fact that pupils have access to calculators is no excuse. Too few pupils work their way up to an advanced achievement level. The Netherlands lags behind Asia and is losing ground to other western European countries.

RELATIONSHIP BETWEEN INSTRUCTIONAL APPROACHES AND MATHEMATICAL PROFICIENCY

The Committee's investigation of the relationship between instructional approaches and mathematical proficiency was based on empirical research conducted in the

rekenvaardigheid. Het onderzoek is beperkt en biedt geen overtuigende empirische ondersteuning voor de claims van enige partij in de discussie over traditioneel versus realistisch rekenen.

Het bereik van het wetenschappelijk onderzoek op het gebied van de effectiviteit van het rekenonderwijs is smal. *Interventiestudies* en *curriculumstudies* kennen, ondanks een vaak zorgvuldige opzet, serieuze beperkingen qua omvang of qua gebrek aan controle over externe variabelen. *Design experiments* zijn veelal gericht op de ontwikkeling en evaluatie van leergangen, niet op vergelijkend onderzoek.

Toch bestaat er op een aantal punten wel duidelijkheid:

- *Binnen* een bepaalde rekendidactiek bestaan vaak grotere verschillen in de leerlingprestaties dan *tussen* rekendidactieken. De specifieke uitwerking van de didactiek en de interactie tussen leraar en leerling spelen kennelijk een grotere rol dan de algemene rekendidactische principes.
- Meer onderwijstijd en aandacht voor rekenen leidt tot betere resultaten.
- Rekenzwakke kinderen lijken minder gebaat bij een vrije vorm van instructie en hebben meer behoefte aan een sturende rol van de leraar.

Het Ministerie van OCW moet maatregelen nemen om het wetenschappelijk onderzoek naar het rekenonderwijs in omvang en variatie te doen toenemen. Naast een reflectie op de kerndoelen en ontwikkelingsexperimenten vraagt de commissie om aandacht voor methodologisch verantwoord vergelijkend onderzoek naar het effect van rekendidactieken en realisaties daarvan, nadere analyses van gegevens van grootschalige nationale en internationale peilingen, en studies naar rol en effect van informatietechnologie. Universiteiten moeten de inrichting van hun opleidingen voor leraar primair onderwijs aangrijpen als een gelegenheid het palet van het onderzoek te vergroten. NWO moet de multidisciplinaire samenwerking tussen wiskundigen, vakdidactici, pedagogen en psychologen bij het uitvoeren van onderzoek naar het rekenonderwijs stimuleren.

ONTWIKKELING IN HET AANBOD VAN REKENMETHODEN

Het marktaandeel van realistische rekenmethoden is toegenomen van 15 procent in 1987 tot 100 procent in 2004. Rond 2010 komen er nieuwe uitgaven op de markt. De ervaringen in de klas en ook het huidige debat stimuleren een aanpassing, met meer rust in de presentatie en met meer aandacht voor oefenen en het onderhouden van basale vaardigheden en cijferen. Ook komen er nieuwe rekenmethoden beschikbaar met minder ruimte voor contexten en de eigen oplossingswijzen van leerlingen.

De toenemende diversiteit in het aanbod geeft scholen meer ruimte om te kiezen. Scholen zijn echter nauwelijks in staat om rekenboeken goed met elkaar te vergelijken. Het Ministerie van OCW, in samenwerking met partijen in het veld, moet er daarom voor zorgen dat rekenmethoden objectief worden geanalyseerd, opdat scholen een verantwoorde keuze kunnen maken.

De onderwijsinspectie moet zich blijven beperken tot de controle *of* de doelen en referentieniveaus worden bereikt en niet sturen op *hoe* ze worden bereikt. Cito moet

Netherlands in the past twenty years, and on a brief and general survey of studies carried out abroad.

Its conclusion is that this empirical material is not unequivocal and does not permit any general, scientifically-grounded statements about the relationship between mathematics instructional approaches and mathematical proficiency. The research is limited and does not provide convincing empirical evidence for the claims made by either side of the debate about the effectiveness of traditional methods versus RME.

Research on the effectiveness of mathematics teaching covers a narrow bandwidth. Although they are often meticulously designed, *intervention studies* and *curriculum studies* have serious limitations in terms of scale or lack of control over external variables. *Design experiments* often focus on developing and evaluating an innovative instructional approach, and not on comparative research.

Nevertheless, a number of conclusions can be made:

- There are often larger differences in pupil performance *within* a particular mathematics instructional approach than *between* the two different approaches. The particular way the approach is implemented in the classroom and the interaction between teacher and pupil play a larger role than general instructional principles.
- Setting aside more time and effort for mathematics leads to better results.
- Children with weak mathematical proficiency seem to derive less benefit from an open style of instruction and need more directed instruction from teachers.

The Dutch Ministry of Education, Culture and Science must take steps to foster more extensive and more varied research into mathematics education. In addition to reflections on educational standards and design experiments, the Committee would like to see methodologically sound comparative research on the effect of mathematics instructional approaches and how they are put into practice; in-depth analyses of data collected in large-scale national and international assessments and studies on the role and impact of information technology. When setting up primary school teacher training programmes, universities should take advantage of the occasion to extend the scope of research. The Netherlands Organisation for Scientific Research should encourage mathematicians, maths instructors, educationalists and psychologists to research mathematics teaching in multidisciplinary teams.

DEVELOPMENT IN THE AVAILABILITY OF MATHEMATICS TEXTBOOKS

Textbooks based on realistic mathematics education increased their market share from 15% in 1987 to 100% in 2004. New editions will be published in around 2010. Lessons learned from teaching practice and the current debate are both causing adjustments to be made: the presentation of the learning material is more muted, and there is more emphasis on exercises, maintaining basic skills and practicing calculations. New textbooks are also becoming available that place less emphasis on contexts and on pupils' own solution strategies.

The growing diversity of textbooks available gives schools more choice, but schools are not really capable of making critical comparisons between mathematics textbooks

in zijn toetsen een balans aanbrengen tussen opgaven met en zonder context, tussen inzicht en vaardigheid.

ROL, OPLEIDING EN NASCHOLING VAN DE LERAAR

De leraar is de spil in het onderwijsleerproces. De kwaliteit van de leraar heeft direct effect op de leerprestaties. De rol van de leraar staat echter onder druk door het zelfstandig werken, door een tekortschietende pabo-opleiding, en door beperkte nascholing en begeleiding.

De commissie plaatst vraagtekens bij de effectiviteit van niet-begeleid zelfstandig werken tijdens de rekenles en pleit voor een grotere inhoudelijke rol voor de leraar. Sturing door en interactie met de leraar en instructie, oefening en nabespreking zijn noodzakelijk.

Er bestaat grote zorg over de situatie op de pabo. Het niveau van de instroom neemt af. Het aantal contacturen voor rekenen is laag. Vakdocenten zijn niet inhoudelijk betrokken bij de stages van hun studenten. Er wordt vaak integraal getoetst, wat slechte rekenresultaten kan camoufleren. Er is nog geen landelijke afstemming van de te stellen eisen.

Het Ministerie van OCW dient de situatie op de pabo's aan een grondig onderzoek te onderwerpen, met als doel het vakinhoudelijke en vakdidactische niveau van de opleiding te verhogen. Het gaat daarbij om de volgende vragen: Is het mogelijk hogere eisen te stellen aan de instroom? Is er een goede balans tussen algemene pedagogische en vakspecifieke leergebieden en tussen rekenvaardigheid en rekendidactiek? Kan de tijd voor rekenen worden uitgebreid, ter verhoging van de professionele gecijferdheid en vakdidactische kennis? Biedt de kennisbasis die de HBO-raad heeft laten ontwikkelen kans op een landelijke normering? Is het zinvol specialisaties voor onderbouw en bovenbouw op de pabo in te voeren? Kunnen universitaire pabo's een rol spelen bij de verbetering van het rekenonderwijs?

Nascholing in rekenen is van belang, maar de vraag ernaar is de afgelopen jaren sterk afgenomen. Internationaal gezien is Nederland op dit punt hekkensluiter. Het Ministerie van OCW dient nascholing in rekenen, in combinatie met begeleiding op de werkvloer, krachtig te stimuleren. Scholen moeten overwegen rekencoördinatoren aan te stellen, ter ondersteuning van de leraar.

on their own. The Ministry of Education, Culture and Science must therefore work with other parties in the field to ensure that mathematics textbooks are analysed objectively so that schools can make well-informed choices.

The education inspectorate would continue to restrict itself to examining *whether* the educational standards and reference levels are being achieved, and not *how* they are being achieved. Test developers such as CITO should design tests that strike a good balance between problems with and without context, between understanding and skills.

ROLE OF TEACHER, TEACHER TRAINING AND POST-GRADUATE COURSES

Teachers are at the heart of the teaching-learning process. The teacher's competence has a direct impact on pupil outcomes. The role of the teacher is under pressure, however: pupils today work too independently, teacher training programmes are inadequate, and post-graduate courses and coaching are limited.

The Committee questions the effectiveness of letting pupils work independently and without guidance during mathematics lessons; it would prefer teachers to take a more active role. Pupils need to be guided by and interact with their teachers; they require instruction, practice and follow-up.

The situation at teacher training colleges is very worrisome. The standard of students entering such programmes is declining. Mathematics merits only a small number of teaching hours. Specialist instructors are not involved in their students' traineeships. Many of the tests cover a wide range of subjects, camouflaging poor maths results. The requirements have not yet been coordinated at national level.

The Ministry of Education, Culture and Science should conduct a thorough investigation of the situation at teacher training colleges, the aim being to improve the programme on subject matter and teaching methods. The following questions should be considered: Would it be possible to require a higher entrance level of students? Has a good balance been found between general pedagogical and subject-specific knowledge, and between students mathematical proficiency and their knowledge of mathematics instructional approaches? Can the time spent on mathematics be increased in order to improve teachers' numeracy and their ability to teach the subject? Does the knowledge base developed at the request of the Netherlands Association of Universities of Applied Sciences make it more likely that national standards will be set? Would it be sensible to introduce specialist teacher training programmes for the lower and upper years of primary school? Can university teacher training programmes play a role in improving mathematics education?

Post-graduate courses in mathematics are important, but the demand for such courses has declined sharply in the past few years. The Netherlands is at the bottom of the international rankings in this respect. The Ministry of Education, Culture and Science must do its utmost to encourage post-graduate training in mathematics, combined with classroom coaching. Schools should consider appointing mathematics coordinators to support teachers.

1. INLEIDING

1.1 Aanleiding voor dit advies

De laatste jaren is er sprake van een aanhoudende kritiek op het huidige rekenonderwijs in ons land. Deze kritiek wordt gevoed door bezorgdheid over de rekenvaardigheid van kinderen. De discussie over rekenonderwijs en rekenvaardigheid heeft niet alleen plaats in de vakliteratuur maar ook in de media. Het publieke debat kent een felle toon en wordt gevoerd met de kracht van overtuiging, vaak met meer ruimte voor anekdotiek en karikatuur dan voor feitelijke onderbouwing. Het debat spitst zich vooral toe op de ‘realistische’ rekendidactiek. Men bedenke hierbij dat het rekenonderwijs op de basisschool de laatste decennia belangrijke veranderingen heeft ondergaan, waarbij de overgang van ‘traditioneel’ naar ‘realistisch’ rekenen het meest in het oog springt.

De Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) is van mening dat dit publieke debat polariserend werkt en niet alleen schadelijk is voor het basisonderwijs maar ook doorwerkt in het voorgezet onderwijs en in het wetenschappelijke veld. Het debat zou zoveel mogelijk gevoerd moeten worden op basis van wetenschappelijk gefundeerde kennis over de zaken die hier ter discussie staan. De KNAW besloot daarom een commissie ‘Rekenonderwijs basisschool’ in te stellen. Toen de staatssecretaris van OCW, mevrouw Dijksema, een overzichtsstudie naar effectieve rekenmethodieken aankondigde (*kamerbrief PO/KU/56359*, 29 september 2008), besloten de KNAW en het Ministerie van OCW beide initiatieven te combineren. Het nu voorliggende advies is daarvan het resultaat.

1.2 Opdracht en samenstelling van de commissie

De KNAW heeft de volgende opdracht geformuleerd:

Breng in kaart wat er bekend is over de relatie tussen rekendidactiek en rekenvaardigheid op grond van bestaande inhoudelijke inzichten en empirisch feitenmateriaal. Geef daarbij aan hoe ruimte kan worden geschapen voor leraren en ouders om te kiezen op basis van informatie over deze relatie tussen rekendidactiek en effect. Het initiëren of uitvoeren van nieuw onderzoek behoorde niet tot de taak van de commissie.

De samenstelling van de commissie was als volgt:

1. Prof. dr. ir. Hester Bijl, Faculteit Luchtvaart- en Ruimtevaarttechniek, Technische Universiteit Delft
2. Dr. Marjolein Kool, PABO Hogeschool Domstad en Freudenthal Instituut, Utrecht
3. Prof. dr. Jan Karel Lenstra (voorzitter), Centrum Wiskunde & Informatica, Amsterdam, en Faculteit Wiskunde & Informatica, Technische Universiteit Eindhoven
4. Drs. Anneke Noteboom, SLO-nationaal expertisecentrum leerplanontwikkeling, Enschede
5. Dr. Cornelis van Putten, Instituut Psychologie, Universiteit Leiden
6. Prof. dr. Rob Tijdeman, Mathematisch Instituut, Universiteit Leiden
7. Prof. dr. Lieven Verschaffel, Center for Instructional Psychology and Technology, Katholieke Universiteit Leuven

Drs. Marian Hickendorff (Instituut Psychologie, Universiteit Leiden) heeft als toegevoegd onderzoeker de commissie inhoudelijk ondersteund. Ir. Arie Korbijn (KNAW) was secretaris.

1.3 Werkwijze van de commissie

We hebben als commissie onze taak breed opgevat. Naast de vraag wat er feitelijk bekend is over de relatie tussen rekendidactiek en rekenvaardigheid, hebben we ons de vraag gesteld wat de huidige stand van zaken is met betrekking tot de rekenvaardigheid van kinderen. Wat betreft het tweede gedeelte van onze opdracht proberen we aan te geven hoe er ruimte is te scheppen voor het maken van didactische keuzes, maar ook hoe het rekenonderwijs verbeterd kan worden.

Onze analyses van de rekenvaardigheid en van de relatie daarvan met de rekendidactiek zijn gebaseerd op in de literatuur gerapporteerde feiten en onderzoeksresultaten. Onze verkenning van de ruimte om te kiezen tussen rekendidactieken en om het rekenonderwijs te verbeteren is primair gebaseerd op de interviews die wij hebben gehouden. Een uitputtende analyse van alle factoren die hier een rol spelen behoorde niet tot onze opdracht en zou ook niet binnen het ons gegeven tijdbestek hebben gepast. Wel hebben we feitenmateriaal verzameld over enkele deelaspecten van deze problematiek, in het bijzonder betreffende de opleiding en nascholing van leraren.

1.4 Opzet van het rapport

Hoofdstuk 2 is gewijd aan de ontwikkelingen in het rekenonderwijs. We bespreken de geschiedenis daarvan in de laatste halve eeuw en beschrijven de rekendidactieken die elkaars tegenpolen heten te zijn, de traditionele en de realistische.

Hoofdstuk 3 bespreekt de stand van zaken met betrekking tot de rekenvaardigheid. We gaan daarbij uit van periodiek uitgevoerde nationale en internationale peilingen van de prestaties van leerlingen, PPON en TIMSS.

Hoofdstuk 4 brengt in kaart wat er bekend is over de relatie tussen rekendidactiek en rekenvaardigheid. We volgen hier een uitgebreide overzichtsstudie, die op ons verzoek werd uitgevoerd door drs. Marian Hickendorff. Tevens geven we een beknopte inventarisatie van de meest toonaangevende internationale literatuur.

Hoofdstuk 5 beschrijft de ontwikkeling in de beschikbare rekenmethoden.

Hoofdstuk 6 heeft de rol, opleiding en nascholing van de leraar tot onderwerp. Dit lijkt de algemene noemer waaronder vele manieren om het rekenonderwijs te verbeteren vallen.

Elk van deze hoofdstukken eindigt met een korte samenvatting, waarin onze conclusies en aanbevelingen zijn opgenomen.

2. ONTWIKKELINGEN IN HET REKENONDERWIJS

2.1 Een halve eeuw geschiedenis¹

In de jaren vijftig en zestig werd het vak rekenen op de lagere school op traditionele wijze gegeven. Er bestond echter nationaal en internationaal ontevredenheid over het rekenonderwijs. Er was zowel inhoudelijk als didactisch sprake van verstarring en de resultaten waren niet goed. De lancering van de kunstmaan Spoetnik door de Sovjet-Unie in 1957 wordt veelal gezien als de aanleiding tot grootse internationale hervormingen in het reken- en wiskundeonderwijs. Deze lancering veroorzaakte een schok in het Westen en leidde tot actie. In 1959 vond in Royaumont bij Parijs op initiatief van de OEEC een conferentie plaats die een drastische hervorming van het wiskundeonderwijs beoogde. De nadruk moest gelegd worden op wiskunde als structuur. Euclidische meetkunde moest verdwijnen en vervangen worden door vectormeetkunde en lineaire algebra. Verzamelingenleer en logica moesten de kern van het curriculum worden. Algoritmen zouden een geringere betekenis krijgen in verband met de opkomst van computers en rekenmachines. De conferentie stond in het teken van economische en technische vooruitgang als maatschappelijk belang. De initiatiefnemers waren onderzoekers uit de wiskunde, psychologie, onderwijskunde en techniek. Deze conferentie heeft grote invloed gehad op het wiskundeonderwijs in Europa en Amerika. Beleidsmakers, politici en lerarenorganisaties schaarden zich achter deze ontwikkelingen. In landen als België, Frankrijk en Duitsland werd niet

1 Voor deze paragraaf is gebruik gemaakt van bijdragen van M.C. van Hoorn, oud-leraar van een pedagogische academie en oud-rector van een scholengemeenschap, en dr. E.W.A. de Moor, voormalig wiskundeleer, rector, opleider, onderzoeker en leerplanontwikkelaar.

alleen het wiskundeonderwijs op de middelbare school, maar ook de rekenles in het primair onderwijs gebaseerd op verzamelingenleer en logica. Een gangbare term voor dit onderwijs is *New Math*.

In Nederland werd in 1961 door de overheid de Commissie Modernisering Leerplan Wiskunde (CMLW) opgericht, die de ideeën van de conferentie in Royaumont moest vertalen in leerplanwijzigingen. Binnen de CMLW werd de werkgroep Wiskobas (wiskunde basisschool) opgericht voor de verbetering van het rekenonderwijs. In 1971 werd Wiskobas onderdeel van een nieuw Instituut voor Ontwikkeling van Wiskunde Onderwijs (IOWO), waarvan prof. dr. Hans Freudenthal hoogleraar-directeur werd. De opdracht was de ontwikkeling van een nieuw leerplan voor de basisschool, een nieuw programma voor de pabo en het opzetten van de bijscholing van leraren basisonderwijs. Het Wiskobas-team ontwikkelde een nieuw leerplan, dat werd neergelegd in elf leerplanpublicaties (1975-1980).

Aanvankelijk was in dit programma de invloed van de *New Math* nog merkbaar, maar al spoedig sloegen de ontwikkelaars een eigen weg in. De basisvaardigheden van het rekenen bleven daarbij een belangrijk doel, hoewel voor de cijferalgoritmen gedifferentieerde doelen werden aanbevolen. Nieuwe elementen waren *handig rekenen*, *schattend rekenen*, *meten en meetkunde*. In 1979 voerde prof. dr. Adri Treffers de term 'realistisch reken- en wiskundeonderwijs' in, omdat in dit nieuwe rekenen het leren veelal begon in contexten waar de leerlingen zich iets bij konden voorstellen. Deze contexten moesten de kinderen in staat stellen zelf kennis te construeren en het rekenen betekenisvoller maken.

Toen de leerplanpublicaties gereed waren, startten leerboekauteurs met het omzetten van de ideeën in nieuwe leerboeken. Tevens moesten de leraren zich in de nieuwe opzet bekwamen door middel van een uitgebreid nascholingsprogramma. Leraren aan de toenmalige pedagogische academies speelden in dit geheel een cruciale rol. Zij moesten de nascholing verzorgen en zij konden de leraren begeleiden. Aanvankelijk bestond hiervoor groot enthousiasme en werden op grote schaal cursussen georganiseerd. Na verloop van tijd ontstonden echter ook problemen. Het bleek voor leraren niet eenvoudig om de ambitieuze idealen in de praktijk toe te passen. Het nieuwe onderwijs stelde hoge eisen aan de leraar en velen zetten vraagtekens bij de haalbaarheid.

Er waren verschillende ontwikkelingen in het onderwijs die de invoering van realistisch rekenen beïnvloedden.

Eind jaren zeventig is de structuur van het basisonderwijs sterk gewijzigd. In 1984 fuseerden het lager onderwijs en het kleuteronderwijs tot basisonderwijs. Kleuterleidsters werden na een korte cursus bevoegd verklaard voor alle groepen van de basisschool. De pedagogische academies en de opleidingen voor kleuterleidsters werden omgevormd tot de pedagogische academie basisonderwijs (pabo). Op de pabo veranderde de rol van en waardering voor de vakdocenten. Niet de vakdocenten maar de leraar pedagogiek werd de spil van de opleiding. Het lesbezoek op stagescholen door vakleraren verdween nagenoeg en de beschikbare tijd voor rekenen en andere vakken verminderde sterk.

Op de basisschool werd de aandacht van de leraren door een grotere diversiteit van leerlingen voor een breder scala aan problemen gevraagd dan voorheen. De aandacht voor het vak rekenen verdween daardoor soms naar de achtergrond.

Het zelfstandig werken deed zijn intrede, een verzamelnaam voor onderwijsvernieuwingen waarbij de leerlingen individueel en zelfstandig hun leertraject afleggen. Hierbij is minder ruimte voor interactief onderwijs, klassikaal of in groepen. Later komt dit terug in de vorm van onderwijs-op-maat.

Er werden kerndoelen geformuleerd, waaraan het traditionele rekenonderwijs niet voldeed. Scholen die met een traditionele methode lesgeven zagen zich daarom genoodzaakt hun methoden aan te passen.

Er verschenen steeds meer nieuwe rekenboeken, die weliswaar alle het predicaat 'realistisch' droegen, maar onderling grote verschillen vertoonden. Na de invoering van de euro in 2002 waren er uitsluitend nog realistische rekenboeken op de markt (zie hoofdstuk 5). Het hanteren van een realistisch rekenboek betekende echter niet dat het daarmee uitgevoerde rekenonderwijs realistisch was. De manier waarop leraren daarmee omgingen, verschilde enorm.

De opkomst van de rekenmachine heeft geleid tot nieuwe en zeer verschillende opvattingen onder alle groepen belanghebbenden. Bij sommigen heerste de mening dat leerlingen niet meer hoefden te cijferen vanwege de opkomst van de rekenmachine, anderen wilden de rekenmachine verbieden in het basisonderwijs, weer anderen zagen de rekenmachine vooral als een didactische uitdaging.

Tijdens de Panamacaferentie van 2007 stelde prof. dr. Jan van de Craats de uitgangspunten en de kwaliteit van het realistische rekenonderwijs ter discussie. Dit maakte veel reacties los. Een maatschappelijke discussie over het rekenonderwijs werd via de media gevoerd. Wat de uitkomst ook zal zijn, meer aandacht voor rekenvaardigheid en rekendidactiek zal op den duur gunstig zijn voor de kwaliteit van het rekenonderwijs.

2.2 Wat is traditioneel rekenen?

De destijds gangbare reken- en wiskundendidactiek, waartegen de moderne of realistische benadering zich afzet, wordt meestal aangeduid met termen als 'traditioneel' of 'mechanistisch'.

Het omschrijven van mechanistisch of traditioneel rekenonderwijs is moeilijk, omdat er niet echt sprake is van een uitgewerkte theorie of een geëxpliciteerde visie. Vooral op basis van beschrijvingen die we vinden bij prominente voorstanders van een terugkeer naar het mechanistische of traditionele rekenen (Braams en Milikowski, 2008; Van de Craats, 2008), kan het kortweg als volgt getypeerd worden.

Onder traditioneel rekenen verstaat men rekenen waarbij de leraar de klas één efficiënte standaardmethode om een bepaald type opgave op te lossen aanreikt (in concreto: het standaardalgoritme) en uitlegt en door alle leerlingen intens laat inoefenen tot ze die beheersen. Men is ervan overtuigd dat leerlingen – en zeker de modale en de

zwakke leerlingen – in verwarring worden gebracht wanneer er bij elk type van rekenbewerking allerlei verschillende strategieën of methoden door en naast elkaar worden gepresenteerd. Bovendien zijn veel van die ‘handige’ (hoofd)rekenstrategieën slechts bij een beperkt aantal berekeningen werkelijk handig (Van de Craats, 2008, p. 25).

Volgens de traditionalisten moet de nadruk dus gelegd worden op het stap-voor-stap aanleren en inoefenen van die standaardrecepten. En dat zijn er precies twaalf, namelijk voor het *optellen*, *afrekken*, *vermenigvuldigen* en *delen* van achtereenvolgens natuurlijke getallen, kommagetallen en breuken. De gehele rekendidactiek moet erop gericht zijn die twaalf procedures door en door te beheersen. Het verdere rekenonderwijs kan aan deze twaalf kapstukken worden opgehangen.

Per standaardrecept ziet het efficiënte traditionele onderwijsleerproces er in principe steeds hetzelfde uit. Centraal staat steeds het uitgebreid, individueel en op papier inoefenen van de door de leraar gedemonstreerde en uitgelegde standaardaanpak voor de betreffende opgave. Men gaat ervan uit dat juist als gevolg van dat oefenen ook het begrip van en inzicht in de geleerde kennis en vaardigheden vanzelf ontstaan en toenemen. Voor concreet materiaal en (voorbeeld)contexten is er een beperkte plaats, namelijk in een korte oriënteringsfase. Leerlingen moeten daar in de daaropvolgende oefenfase zo snel mogelijk van loskomen en naar de kern van de opgave gaan, het vlot leren uitvoeren van het recept. Tijdens dat oefenen – eerst met gemakkelijker opgaven, daarna met moeilijker – hebben contexten geen nut, omdat die afleiden van de essentie. Om diezelfde reden is er geen plaats voor gevarieerd en flexibel strategiegebruik. Als het niveau van vlotte beheersing van de standaardprocedure bereikt is, is er weer ruimte voor contexten, namelijk als toepassingen achteraf van het geleerde recept. Over *meten*, *meetkunde*, *schattend rekenen*, *procenten*, *verhoudingen*, *toepassingen* en *gebruik van de rekenmachine* doen de huidige pleitbezorgers van een terugkeer naar de traditionele rekendidactiek weinig of geen uitspraken.

2.3 Wat is realistisch rekenen?

Realistisch rekenen (RR) is een didactische theorie die in de zeventiger jaren is ontwikkeld als reactie op het toen gangbare traditioneel rekenen (zie vorige paragraaf). Uitgangspunt was het idee van Freudenthal dat het bij het leren van wiskunde niet gaat om het verwerven van een kant-en-klaar-product, een verzameling weetjes, maar dat wiskunde een menselijke activiteit is, die haar oorsprong vindt in alledaagse situaties en die erop gericht is deze situaties beter hanteerbaar te maken en rekenproblemen die zich daarin voordoen op te lossen (Freudenthal, 1973). Treffers introduceert voor deze menselijke wiskundige activiteit de term ‘mathematiseren’ en onderscheidt daarin een horizontale en een verticale component. Horizontaal mathematiseren is het zodanig transformeren van een reëel probleem, dat dit met beschikbare wiskundige middelen aangepakt kan worden en het interpreteren van het antwoord. Verticaal mathematiseren verwijst naar de generalisatie van de oplossing, de niveauverhoging, de verkorting, de verdergaande formalisering van de actuele rekenkundige handelingen, het ontdekken van structuren en patronen (Treffers, 1978).

Treffers formuleerde vijf karakteristieke grondprincipes van het realistisch rekenen (Treffers, 1987):

1. *Zelf kennis construeren*: Begrip en inzicht ontstaan doordat kinderen onder begeleiding van een deskundige leraar gestimuleerd en geholpen worden om uitgaande van een reëel probleem zelf kennis te construeren. Dit probleem kan een verschijnsel zijn uit het dagelijks leven van de leerling, maar ook reeds verworven wiskundige kennis kan uitgangspunt zijn van het construeren van nieuwe wiskundige kennis (Freudenthal, 1991). Belangrijk is dat leerlingen zich bij dit probleem iets kunnen 'realiseren' of voorstellen; vandaar de term realistisch reken- en wiskundeonderwijs. Leerlingen exploreren het probleem en verzamelen intuïtieve noties waarmee essentiële aspecten van begrippen en structuren gevormd worden. Die eigen ideeën en informele oplossingsmanieren van de kinderen vormen het uitgangspunt van het realistische rekenonderwijs; dit zijn de zogenoemde 'eigen producties'. Het rekenonderwijs erkent en bouwt voort op deze informele oplossingsmanieren, sluit aan bij wat kinderen kunnen en waar ze op dat moment aan toe zijn.
2. *Niveaus en modellen*: Modellen, schema's, structuren, diagrammen en symbolen vormen de brug om informele eigen aanpakken van kinderen geleidelijk te ontwikkelen tot meer gestructureerde en uiteindelijk abstracte (formele) manieren.
3. *Reflectie op eigen producties*: Kinderen worden uitgedaagd te reflecteren op hun eigen producties, hun eigen handelen. Door het stellen van vragen, confronteren met alternatieven of contradicties en aanwakkeren van discussie stimuleert de leraar het kind een volgende stap in het leerproces te zetten. Freudenthal sprak van *guided reinvention*.
4. *Interactie*: In het realistisch rekenonderwijs leren leerlingen van en met elkaar door hun oplossingsmanieren te verwoorden, te vergelijken en eventueel te verdedigen of juist aan te passen. Onder leiding van de leraar wordt hierdoor het proces van verkorten en niveauverhoging gestimuleerd.
5. *Verstrengeling van leerlijnen*: Leerlingen worden gestimuleerd om dwarsverbanden en samenhang binnen de leerstof te ontdekken. Het doel van het rekenonderwijs is een samenhangend, toepasbaar, geïntegreerd geheel van kennis, inzichten en vaardigheden.

2.4 Traditioneel versus realistisch rekenen: discussie

Traditioneel rekenen (TR) en realistisch rekenen (RR) kunnen beide een didactiek genoemd worden, maar de commissie constateert dat er bij deelnemers aan de discussie bepaald geen eenduidig beeld bestaat van wat TR is en van wat RR is. Hiervoor zijn twee redenen. Ten eerste is TR in de loop der tijden ontstaan en alleen achteraf op hoofdlijnen als didactiek beschreven. RR is daarentegen ontworpen als didactiek en gebaseerd op een geëxpliciteerde theorie; de vijf principiële uitgangspunten van deze didactiek zijn uitvoerig beschreven (zie vorige paragraaf). In de tweede plaats bestaat iedere didactiek op verscheidene niveaus: conceptueel in het hoofd van de

ontwikkelaar, op schrift in het leerboek, in de perceptie van de leraar, en in de praktijk van de klas. De verschillen tussen deze niveaus zijn enorm. In het bijzonder RR heeft bij de concretisering, implementatie en realisatie in de praktijk vele gezichten gekregen. Men kan kortom niet over *het* RR of *het* TR spreken.

Een kanttekening bij de ontwikkeling van RR in de tijd is hier op zijn plaats. In het oorspronkelijke concept van RR bestond er een evenwicht tussen horizontale en verticale mathematisering. In de uitwerking van RR in rekenmethoden en in de concrete praktijk van de klas kreeg verticale mathematisering steeds minder aandacht en kwam de nadruk te sterk te liggen op tijdrovende verkenningen van contexten en op oppervlakkige toepassingen. De Duitse wiskundendidacticus Wittmann (2005) waarschuwde in dit verband dat RR *RR-light* dreigde te worden en riep daarom op om de verticale dimensie opnieuw de plaats te geven die het in het oorspronkelijke concept innam. In dit verband dient benadrukt te worden dat het traditionele rekenonderwijs zoals thans door haar pleitbezorgers wordt beschreven (zie bijvoorbeeld Braams en Miliowski, 2008) het risico in zich draagt op een nog geringere aandacht voor de verticale mathematiseringsdimensie.

Het publieke debat negeert de nuances en trekt TR en RR in karikaturale extremen. De discussie heeft betrekking op verschillende didactische elementen, zoals de rol van context in het rekenonderwijs, de balans tussen het krijgen van inzicht en het opdoen van routine, de balans tussen uit het hoofd rekenen en het opschrijven van tussenresultaten, de aansluiting bij strategieën die kinderen van nature al toepassen, de wenselijkheid om voor een bewerking vanuit diverse strategieën tot één standaardalgoritme te komen, en de plaats van de rekenmachine in het rekenonderwijs. Sommige deelnemers aan de discussie vertonen de neiging door hen niet gewenste elementen in negatieve zin te vervormen. Oefenen wordt stampen, de inzet van betekenisvolle contexten wordt een cursus begrijpend lezen genoemd. "Leidt routine tot begrip of leidt begrip tot routine?" wordt een kernvraag, terwijl het een zinloze simplificatie is; leerprocessen zijn complexer. Dergelijke typeringen maken het onmogelijk om de verschillende standpunten tot elkaar te brengen.

Terwijl de discussie in de media zich toespitst op de tegenstellingen tussen TR en RR ontstaat in de praktijk van het onderwijs toenadering. Realistische rekenmethoden kiezen voor meer oefening, minder snelle afwisseling van onderwerpen, visuele contexten en meer aandacht voor schriftelijk rekenen. De traditionele methoden die naar verwachting op de markt gaan komen (zie hoofdstuk 5) moeten voldoen aan de kerndoelen en daarom ook aandacht besteden aan *hoofdrekenen*, *meetkunde*, *schattend rekenen* en de ontwikkeling van begrip, inzicht, flexibiliteit, denken en redeneren; zij kunnen daarom elementen uit RR niet negeren. Het is te verwachten dat er een bandbreedte zal ontstaan waarin scholen kunnen kiezen voor een meer traditionele of meer realistische benadering.

De commissie is van mening dat sommige didactische aspecten van RR bij realisering in de praktijk hoge eisen stellen aan de leraar. Denk bijvoorbeeld aan het aansluiten bij oplossingsmanieren van leerlingen, inspelen op verschillende niveaus

van leerlingen, interactieve besprekingen houden en gezamenlijk kennis construeren. Tevens streeft men in RR naar meer en hogere doelen dan in TR, zoals redeneren, reflecteren, leren leren en strategisch denken. Dat maakt dat het uitvoeren van RR wellicht meer kennis en vakmanschap van de leraar vereist dan het realiseren van TR.

Ten slotte merken we op dat het onderwijs in het algemeen gebaat is bij kleine aanpassingen. Gegeven de inertie van het onderwijsveld zijn drastische wijzigingen soms nodig, maar zij hebben een hoge prijs en vereisen dan ook overtuigende argumentatie.

2.5 Samenvatting en conclusies

Rond 1960 kwam er in het Westen een herbezinning over het reken- en wiskunde-onderwijs op gang. De opbrengsten van het bestaande onderwijs vielen tegen en de lancering van de Spoetnik in 1957 suggereerde dat de Sovjet-Unie een wetenschappelijke en technologische voorsprong bezat. In verscheidene landen leidde dit tot een invoering van de *New Math*, met een nadruk op formalisme en abstractie. Nederland volgde, onder invloed van Freudenthal en zijn medewerkers, een eigen weg, die leidde tot de realistische rekendidactiek.

De implementatie van de realistische rekendidactiek, van ambitie naar praktijk, was niet eenvoudig. Men plaatste vraagtekens bij de haalbaarheid, mede in verband met de hoge eisen die de nieuwe didactiek aan de leraar stelde. Uiteindelijk is grootschalige invoering gerealiseerd.

Ondertussen onderging het onderwijs andere ingrijpende wijzigingen. Lager onderwijs en kleuteronderwijs werden samengevoegd tot basisonderwijs, pedagogische academie en opleiding tot kleuteronderwijzer fuseerden tot pabo, met een kleinere rol voor de vakdocent en een grotere voor de pedagoog. Er ontstonden zorgen over het niveau van de instroom. Er werden kerndoelen geformuleerd, waaraan het traditionele rekenonderwijs niet voldeed. Het zelfstandig werken werd ingevoerd, waardoor er minder ruimte kwam voor interactief klassikaal onderwijs. De diversiteit van de leerlingen nam toe. De goedkope rekenmachine kwam beschikbaar.

Sinds de invoering van de euro in 2002 presenteren alle beschikbare rekenmethoden zich als realistisch, al zijn er onderling veel verschillen.

De traditionele rekendidactiek, zoals sommige pleitbezorgers die voorstaan, bestaat uit het systematisch aanleren en intensief oefenen van één standaardmethode voor elk van twaalf bewerkingen: *optellen*, *afrekken*, *vermenigvuldigen* en *delen* voor gehele getallen, decimale getallen en breuken. De rol van de context bestaat uit een korte motivatie vooraf en een korte toepassingsfase achteraf.

De realistische rekendidactiek ziet wiskunde als een menselijke activiteit, met een oorsprong in de realiteit en met als doel die realiteit beter hanteerbaar te maken. Contexten spelen daarom een prominente rol in het onderwijsleerproces. Leerlingen worden aangemoedigd hun eigen oplossingsstrategieën te ontwikkelen en daar gezamenlijk op te reflecteren.

Iedere didactiek heeft vele gedaantes, van ideëel concept via methode in het boekje en perceptie van de leraar tot de praktijk in de klas. Sommige aspecten van realistisch rekenen zijn minder goed ingevoerd dan de ontwikkelaars oorspronkelijk voor ogen stond. In de praktijk is er vaak te weinig aandacht voor inzicht en abstractie. Traditioneel rekenen anno 2010 zal op het vlak van zowel inhouden als vaardigheden minder beperkt moeten zijn dan hierboven beschreven en zal kerndoelen als *meetkunde*, *hoofdrekenen* en *schatten* en hogere doelen als toepassingsvaardigheid, redeneren, reflecteren en abstractievermogen niet kunnen negeren.

CONCLUSIE 2.1

Traditioneel en realistisch rekenonderwijs zijn geen eenduidige begrippen. De uitgangspunten van het realistisch rekenen worden op uiteenlopende manieren vormgegeven, waarbij de verticale mathematisering vaak onvoldoende aan bod komt. Hoe de realisatie van het nieuwe traditioneel rekenen zal zijn, zal de nabije toekomst leren.

CONCLUSIE 2.2

Realistisch rekenonderwijs stelt wellicht hogere eisen aan de leraar dan traditioneel rekenonderwijs.

3. REKENVAARDIGHEID

3.1 Nationaal en internationaal empirisch onderzoek

Voordat we in het volgende hoofdstuk ingaan op de relatie tussen rekendidactiek en rekenvaardigheid, wordt in dit hoofdstuk de stand van zaken met betrekking tot de rekenvaardigheid van leerlingen in het Nederlandse basisonderwijs beschreven. Het uitgangspunt hierbij is het empirische, kwantitatieve onderzoek op het gebied van nationale peilingen (PPON) en internationale vergelijkingen (TIMSS). Beide typen van onderzoek worden door de Commissie Parlementair Onderzoek Onderwijsvernieuwingen (commissie Dijsselbloem, 2008, p. 116 e.v.) als richtinggevend genoemd om uitspraken te doen over onderwijsprestaties.

In het basis- en speciaal onderwijs wordt sinds 1986 in opdracht van het Ministerie van OCW *de periodieke peiling van het onderwijsniveau* (PPON) uitgevoerd door Cito (Van der Schoot, 2008). De peilingen worden uitgevoerd voor verschillende leerstofgebieden zoals Nederlandse taal, rekenen-wiskunde, wereldoriëntatie, verkeer, Engels en lichamelijke opvoeding. Voor rekenen-wiskunde zijn inmiddels vier peilingen uitgevoerd, in 1987, 1992, 1997 en 2003/2004, steeds in groep 5 en in groep 8 van de basisschool. Wij beperken ons in het vervolg tot de rekenpeilingen in groep 8, omdat die de rekenvaardigheid op het einde van de basisschool aangeven.

TIMSS (*Trends in international mathematics and science study*) is een periodiek internationaal georganiseerd vergelijkend onderzoek naar prestaties van leerlingen op het gebied van *science* en *mathematics* in *grade 4* (groep 6) en *grade 8* (tweede klas middelbare school). TIMSS is eerder afgenomen in 1995 en 2003. In 2007 is TIMSS weer uitgevoerd in 36 landen, Nederland heeft daarbij alleen meegedaan met rekenen-wiskunde en natuurkunde in groep 6. In het vervolg beperken we ons daarom tot de prestaties van leerlingen in groep 6 voor het onderdeel rekenen-wiskunde.

Voor verschillende onderdelen van de rekenvaardigheid zijn de groottes van de veranderingen uitgedrukt in effectgroottes (*effect size*, ES). Voor de definitie zie bijlage 2. De interpretatie van de absolute waarde is als volgt gekozen (zie bijvoorbeeld Cohen, 1988):

- tussen 0,00 en 0,20: effect is verwaarloosbaar tot klein;
- tussen 0,20 en 0,50: effect is klein tot matig;
- tussen 0,50 en 0,80: effect is matig tot groot;
- groter dan 0,80: effect is groot.

3.2 PPON: Periodieke peiling van het onderwijsniveau

PPON-2004² omvatte de rekenvaardigheid van de leerlingen eind groep 8 uitgesplitst naar 22 verschillende onderwerpen, verdeeld over drie gebieden. Het gebied *getallen en bewerkingen* telt tien onderwerpen, het gebied *verhoudingen, breuken en procenten* telt vier onderwerpen; *meten en meetkunde* telt acht onderwerpen.

In het algemeen zijn er drie manieren om het niveau van de rekenvaardigheid te bespreken: Zijn er veranderingen in de tijd? Wat is het niveau vergeleken met een vastgestelde standaard? Welke relevante verschillen zijn er tussen leerlingen?

a. Veranderingen in de tijd

Uit de voorlaatste kolom van tabel 3.1 blijkt dat bij negen onderwerpen het prestatiepeil de afgelopen twintig jaar vrij stabiel is gebleven (effectgrootte tussen $-0,19$ en $+0,19$). Bij zes onderwerpen waren er slechts kleine tot matig grote veranderingen (effectgrootte $0,20 - 0,49$ in positieve dan wel negatieve zin). Bij zeven onderwerpen hebben zich matig grote tot grote (effectgrootte $0,50 - 0,79$, idem) respectievelijk (zeer) grote (effectgrootte vanaf $0,80$, idem) veranderingen in de rekenvaardigheid voorgedaan. Al deze veranderingen zijn statistisch significant. In figuur 3.1 zijn de grootste veranderingen weergegeven. De prestaties op het gebied van *getallen en getalrelaties* en van *schattend rekenen* zijn sterk vooruitgegaan, terwijl die van alle *bewerkingen* (optellen, aftrekken, delen en vermenigvuldigen van gehele getallen en kommagetallen, vroeger cijferend rekenen genoemd) sterk achteruit zijn gegaan. De daling is het grootst bij *vermenigvuldigen* en *delen* waar het peil in toenemende mate is gaan dalen. Bij *hoofdrekenend optellen en aftrekken* en bij *procenten* is er ten slotte sprake van een matig grote verbetering (Janssen, Van der Schoot en Hemker, 2005, p. 231-235).

² Dit onderzoek werd uitgevoerd op een naar formatiegewicht gestratificeerde steekproef van 122 scholen en 3078 leerlingen. Een basissteekproef van 130 scholen is per school met 3 à 4 vooraf getrokken reservescholen met een vrijwel gelijk formatiegewicht uitgebreid, zodat in tweede instantie in totaal 404 scholen zijn benaderd; 30,2 procent daarvan heeft meegedaan aan het onderzoek.

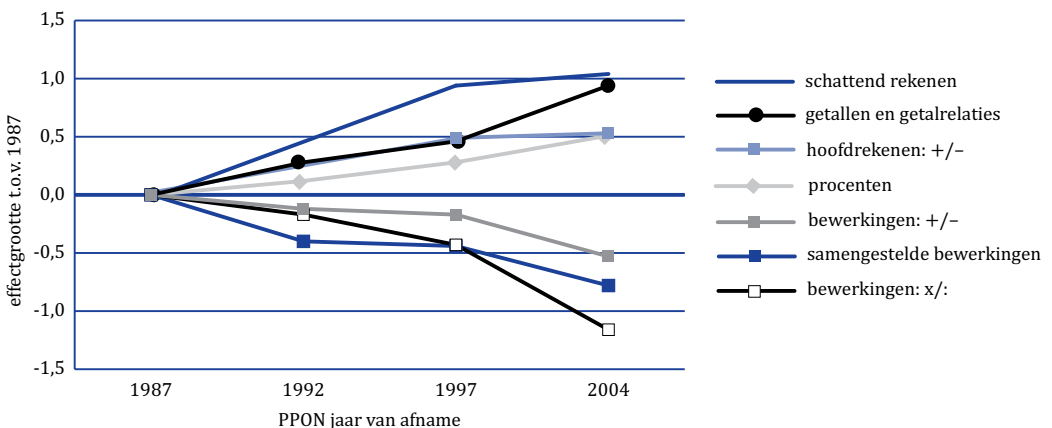
Tabel 3.1 PPON effectgroottes en vergelijking met standaarden

Onderwerpen in PPON-2004 einde basisonderwijs	Effectgrootte verschil in peil 2004-1987		Percentage leerlingen dat de standaard 'voldoende' haalt		
	< - 0,19	> + 0,19	< 40%	40-60%	> 60%
GETALEN EN BEWERKINGEN					
1. getallen en getalrelaties		+ 0,94		42	
2. basisoperaties: + / -		+ 0,24			76
3. basisoperaties: x / :	- 0,20				66
4. hoofdrekenen: + / -		+ 0,53		50	
5. hoofdrekenen: x / :		- 0,11			66
6. schattend rekenen		+ 1,04		42	
7. bewerkingen: + / -	- 0,53			27	
8. bewerkingen: x / :	- 1,16			12	
9. samengestelde bewerkingen	- 0,78			16	
10. rekenen met een zakrekenmachine		+ 0,26		34	
VERHOUDINGEN, BREUKEN EN PROCENTEN					
11. verhoudingen		+ 0,14			66
12. breuken		+ 0,15			60
13. procenten		+ 0,51		58	
14. tabellen en grafieken		+ 0,10		50	
METEN EN MEETKUNDE					
15. meten: lengte		- 0,13		38	
16. meten: oppervlakte		+ 0,05		21	
17. meten: inhoud		- 0,03			42
18. meten: gewicht			+ 0,33		58
19. meten: toepassingen	- 0,25				50
20. meetkunde		- 0,08			62
21. tijd		0,00			50
22. geld	- 0,31				42

De effectgroottes betreffen periode 2004-1987 met uitzondering van de effectgroottes voor onderwerpen 2+3 basisoperaties en 10 rekenen met een zakrekenmachine (2004-1992); 14 tabellen en grafieken (2004-1997); 22 geld (1997-1987).

Bron: Van der Schoot (2008).

Figuur 3.1 Belangrijkste veranderingen in het rekenpeil in groep 8 van 1987 tot 2004. (Janssen e.a. 2005, Cito)



Verklaring van veranderingen in het peil

Cito stelt het peil vast van het onderwijsniveau en door dit periodiek te doen worden veranderingen in het peil zichtbaar. Er is echter nog maar weinig onderzoek gedaan om dergelijke veranderingen in het peil te verklaren. Aanvullend onderzoek aan de Universiteit Leiden naar het strategiegebruik van leerlingen in de PPON-toetsboekjes van de peilingen uit 1997 en 2004 laat zien dat de sterke prestatiedaling bij de bewerkingen voor *vermenigvuldigen en delen* voor een groot deel verklaard kan worden door een verandering in het notatiegedrag van de leerlingen (Hickendorff e.a., 2009; Van Putten en Hickendorff, 2009). Opvallend was dat schriftelijke uitwerkingen van realistische oplossingsstrategieën relatief weinig voorkwamen en dat zij nauwelijks toenamen tussen 1997 en 2004. Traditionele cijferstrategieën (zoals bijvoorbeeld staartdelen) werden in 2004 veel minder vaak gebruikt dan in 1997, terwijl het beantwoorden van de opgaven zonder gebruik te maken van een schriftelijke berekening of een kladnotitie juist sterk toenam, vooral bij de jongens. Geschreven strategieën (traditioneel en realistisch) waren redelijk succesvol, maar het rekenen zonder schriftelijke uitwerking bleek riskant, zodat de toename daarvan de daling van het peil voor een deel kan verklaren. Jongens lijken door minder op schrift te rekenen bij de bewerkingen hun gemiddelde prestatievoorsprong op de meisjes (zichtbaar bij alle andere onderdelen van het rekenen) te verliezen. Ook bij de *bewerking aftrekken* in de peilingen van 2004 leidde het antwoorden zonder schriftelijke uitwerking tot slechtere resultaten, vooral vergeleken met de traditionele aanpak met notatie van het lenen (Van Putten en Hickendorff, 2009).

Ten slotte nam tussen 1997 en 2004 het succes van alle afzonderlijke strategieën bij de *bewerkingen* voor *delen en vermenigvuldigen* significant af (Hickendorff e.a., 2009). Voor een verklaring van dit laatste aspect van de prestatiedaling is nieuw onderzoek nodig.

b. Niveau ten opzichte van vastgestelde standaarden

Naast het periodiek vaststellen van het rekenpeil en de veranderingen daarin, bepaalt Cito ook de absolute waarde van het peil. Dit gebeurt met behulp van standaarden voor 'minimale', 'voldoende' en 'gevorderde' beheersing van elk onderwerp. Het bepalen van deze standaarden wordt regelmatig gedaan via een zorgvuldige bevraging van panels van rekendeskundigen en ervaren leraren uit groep 8 (Van der Schoot, 2008). Bij iedere peiling wordt vervolgens nagegaan hoeveel procent van de Nederlandse leerlingen aan deze standaarden voldoet. Wij beperken ons in onderstaande beschrijving tot de standaard 'voldoende' die zo gesteld is dat 70 tot 75 procent van de leerlingen die zou moeten halen.

Zoals weergegeven in de laatste kolom van tabel 3.1 blijkt dat in PPON-2004 alleen bij het onderwerp *basisoperaties: optellen en aftrekken* de Nederlandse leerlingen deze standaard halen (met 76 procent); bij dit onderwerp is ook het peil verbeterd sinds

de peiling van 1997. Bij enkele andere onderwerpen komt men in de buurt van de standaard met 66 procent van de leerlingen, waarbij het bij het onderwerp *basisoperaties: vermenigvuldigen en delen* opvallend is dat het peil licht is afgenomen, zodat aangenomen mag worden dat bij de eerdere peilingen van 1987 en 1992 de standaard 'voldoende' hier wel bereikt werd. Bij driekwart van de onderwerpen is er echter een aanzienlijk verschil tussen de standaard 'voldoende' en het daadwerkelijk bereikte peil. Het grootste gat is zichtbaar bij de drie bewerkingsonderwerpen waarvan het peil sterk gedaald is, terwijl bij het *meten van lengte* en van *oppervlakte* de kloof tussen standaard en peil onveranderlijk groot is. *Getallen en getalrelaties* en *schattend rekenen* zijn qua peil weliswaar duidelijk vooruitgegaan maar blijven met 42 procent toch nog steeds sterk achter bij de gestelde standaard 'voldoende'.

Standaarden bevatten onvermijdelijk elementen van beoordeling (Hambleton & Pitoniak, 2006) en zowel de procedures waarmee standaarden worden vastgesteld als de uiteindelijke standaarden roepen daarom soms vragen op. Bijvoorbeeld, legt men de lat niet te hoog (Treffers, 2007) en houdt men voldoende rekening met de grote verschillen in leervermogen tussen kinderen? Toch is het bepalen van standaarden noodzakelijk om uit peilingsonderzoek conclusies te kunnen trekken. In paragraaf 3.4 komen de recent voorgestelde referentieniveaus ter sprake. Deze referentieniveaus zijn in feite een stelsel van prestatiestandaarden voor verschillende leeftijden en subgroepen van leerlingen. De bepaling van de referentieniveaus is onder andere gebaseerd op de empirische gegevens uit het periodieke peilingsonderzoek (Expertgroep doorlopende leerlijnen, 2008).

Overigens laat het PPON-rapport zien dat de rekenresultaten halverwege groep 8 over vrijwel de gehele linie gemiddeld beter waren en de standaarden dichterbij benaderden dan aan het eind van groep 8.

c. Verschillen tussen leerlingen

In PPON-2004 zijn net als in bij eerdere peilingen verschillen tussen groepen leerlingen en tussen scholen vastgesteld. De onderzochte leerlingkenmerken zijn het formatiegewicht (indicatie van de sociaal-economische achtergrond), geslacht en de leertijd (is de leerling vertraagd of niet). Op schoolniveau zijn dat de sociaal-economische samenstelling van de schoolbevolking (stratum) en de gebruikte rekenmethode. Dit laatste aspect, verschillen in rekenprestaties naar rekenmethode, zal in hoofdstuk 4 besproken worden.

De hieronder beschreven verschillen zijn *gezuiverd*, wat wil zeggen dat de overige variabelen statistisch constant zijn gehouden.

Formatiegewicht

Formatiegewicht is gerelateerd aan sociaal-economische achtergrond en is ingedeeld in drie categorieën. Leerlingen met formatiegewicht 1,25 zijn Nederlandse arbeiderskinderen, leerlingen met formatiegewicht 1,90 komen uit gezinnen met ten minste één

ouder van niet-Nederlandse herkomst en beperkingen in opleidings- en beroepsniveau. De overige leerlingen hebben formatiegewicht 1,00.

Op vrijwel alle onderwerpen hebben zowel 1,25-leerlingen als 1,90-leerlingen een achterstand ten opzichte van 1,00-leerlingen. Op de gebieden *getallen en bewerkingen* en *verhoudingen, breuken en procenten* verschillen de prestaties van 1,90-leerlingen niet van die van 1,25-leerlingen, bij *meten en meetkunde* hebben de 1,90-leerlingen wel een lichte achterstand.

Jongens-meisjes

Op de meeste onderdelen presteren jongens beter dan meisjes, met verwaarloosbaar kleine tot matige effectgroottes. Op de complexere bewerkingsopgaven die op papier mogen worden uitgerekend (onderwerpen 7, 8 en 9 uit tabel 3.1) presteren meisjes echter beter dan jongens, met effectgroottes tussen 0,12 en 0,33.

Leertijd

Leerlingen die vertraagd zijn in hun schoolloopbaan hebben een matige achterstand ten opzichte van de reguliere, niet vertraagde leerlingen.

Stratum van de school

Scholen zijn ingedeeld in drie strata, op basis van de formatiegewichten van hun leerlingpopulatie. Stratum 1 omvat scholen met overwegend kinderen van ouders met afgeronde voortgezette opleiding en weinig allochtone leerlingen. In stratum 2 zitten scholen met relatief meer Nederlandse arbeiderskinderen maar ook weinig allochtone kinderen. In stratum 3 ten slotte zitten scholen met vooral Nederlandse arbeiderskinderen en allochtone kinderen.

Additioneel op het effect van formatiegewicht van de leerling, presteren leerlingen op stratum 2-scholen iets beter dan op stratum 1-scholen. Leerlingen in stratum 3-scholen presteren licht tot matig slechter dan leerlingen op scholen van beide andere strata.

3.3 TIMSS: Trends in international mathematics and science study

Hieronder worden de resultaten van TIMSS-2007 samengevat van de prestaties van Nederlandse leerlingen in groep 6 op het onderdeel rekenen en wiskunde, ook in vergelijking met andere landen. Deze worden vergeleken met TIMSS-prestaties in 2003 en 1995.

Opzet TIMSS

Bij de afname in 2007 zijn verschillende instrumenten gebruikt: een toets met rekenopgaven en een aantal contextvragenlijsten. De rekentoets bevatte 179 opgaven uit de

inhoudelijke domeinen *getallen* (93), *geometrische vormen en meten* (60) en *gegevensweergave* (26).

In totaal zijn de rekenitems afgenomen bij 4300 leerlingen. De contextvragenlijsten bevatten vragen om zicht te krijgen op verschillende achtergrondkenmerken van de leerlingen, de leraren en de schoolpopulatie.

De landen die meedoen aan TIMSS moeten voldoen aan strenge responseisen: 85 procent van de scholen uit de originele steekproef moet de toets afnemen. Indien dit niet haalbaar is, mag elke niet-deelnemende school vervangen worden door een vooraf bepaalde reserveschool. Dan nog moet minimaal 50 procent van de oorspronkelijk geselecteerde scholen meedoen. De uiteindelijke respons moet minimaal 85 procent zijn.

In Nederland is deelnamebereidheid van scholen altijd al laag, net als bij PPON. Daardoor is het moeilijk om aan de gestelde eisen te voldoen. Hoewel bij TIMSS-2007 inclusief de benadering van reservescholen een respons van 95 procent is gehaald, is net niet de genoemde 50 procent (49 procent) van de oorspronkelijke steekproef behaald³. Nederland is derhalve wel opgenomen in het internationale rapport, maar heeft een aantekening gekregen voor *nearly satisfied response rates with replacement schools*.

Prestaties leerlingen in groep 6 op de reken- en wiskundetoets

De scores op de gehele toets worden in TIMSS weergegeven op een internationaal gestandaardiseerde schaal met een gemiddelde van 500 en een standaarddeviatie van 100 (500 is vastgesteld op basis van het internationale gemiddelde van de deelnemende landen aan de meting in 1995). In tabel 3.2 is de verdeling gegeven van de scores per land.

Nederland heeft een gemiddelde score van 535. De spreiding is klein, de toetsscores van Nederlandse leerlingen in groep 6 liggen relatief dicht bij elkaar. Nederlandse leerlingen presteren gemiddeld significant boven het TIMSS-schaalgemiddelde van 500. Nederland scoort voor rekenen en wiskunde in de subtop van de tien best presterende landen, zoals dat ook het geval was in 2003 en 1995. Alleen Aziatische landen presteren significant beter. Het hogere gemiddelde van Engeland (541) en Letland (537) is niet significant hoger dan dat van Nederland. Het lagere gemiddelde van Litouwen (530) en de Verenigde Staten (529) is eveneens niet significant lager dan dat

3 Het ligt voor de hand om dan te twifelen aan de representativiteit van de Nederlandse steekproef. In tabel 1.2 van het TIMSS-rapport is de verdeling van leerlinggewicht over de scholen af te lezen, zowel de verdeling over de (gerealiseerde) originele steekproefscholen als de gerealiseerde steekproef met vervanging door reservescholen. Immers zijn de scholen uit de reservepool qua rekenvaardigheid vergelijkbaar met de weigerende oorspronkelijke scholen uit de steekproef? De reservescholen waren al geselecteerd ten tijde van de selectie van de steekproef: voor elke originele steekproefschool waren twee vervangende scholen achter de hand die wat betreft stratum en schoolgrootte (en daarmee waarschijnlijk ook in niveau) vergelijkbaar waren met de originele school. Het is dus onwaarschijnlijk dat scholen met juist lagere of juist hogere scores als vervangende scholen zijn geselecteerd. Ook bleken de weigerende scholen niet anders van stratum en schoolgrootte dan scholen die wel toezegden mee te doen.

Tabel 3.2 Verdeling en gemiddelde van de totale rekenscores per land, TIMSS-2007 (gebaseerd op Mullis e.a., 2008).

Landen	Schaalscore rekenen	Gemiddelde rekenscore (s.e.)	Jaren formeel genoten onderwijs	Gemiddelde leeftijd
Hong Kong		▲ 607 (3,6)	4	10,2
Singapore		▲ 599 (3,7)	4	10,4
Chinees Taipei		▲ 576 (1,7)	4	10,2
Japan		▲ 568 (2,1)	4	10,5
Kazakstan		▲ 549 (7,1)	4	10,6
Russische Federatie		▲ 544 (4,9)	4	10,8
Engeland		▲ 541 (2,9)	5	10,2
Letland		▲ 537 (2,3)	4	11,0
Nederland		▲ 535 (2,1)	4	10,2
Litouwen		▲ 530 (2,4)	4	10,8
Verenigde Staten		▲ 529 (2,4)	4	10,3
Duitsland		▲ 525 (2,3)	4	10,4
Denemarken		▲ 523 (2,4)	4	11,0
Australië		▲ 516 (3,5)	4	9,9
Hongarije		▲ 510 (3,5)	4	10,7
Italië		▲ 507 (3,1)	4	9,8
Oostenrijk		▲ 505 (2,0)	4	10,3
Zweden		503 (2,5)	4	10,8
Slovenië		502 (1,8)	4	9,8
TIMSS gemiddelde		500		
Armenië		500 (4,3)	4	10,6
Slowakije		496 (4,5)	4	10,4
Schotland		▼ 494 (2,2)	5	9,8
Nieuw-Zeeland		▼ 492 (2,3)	4.5 - 5.5	10,0
Tsjechië		▼ 486 (2,8)	4	10,3
Noorwegen		▼ 473 (2,5)	4	9,8
Oekraïne		▼ 469 (2,9)	4	10,3
Georgië		▼ 438 (4,2)	4	10,1
Iran		▼ 402 (4,1)	4	10,2
Algerije		▼ 378 (5,2)	4	10,2
Colombia		▼ 355 (5,0)	4	10,4
Marokko		▼ 341 (4,7)	4	10,6
El Salvador		▼ 330 (4,1)	4	11,0
Tunesië		▼ 327 (4,5)	4	10,2
Koeweit		▼ 316 (3,6)	4	10,2
Qatar		▼ 296 (1,0)	4	9,7
Jemen		▼ 224 (6,0)	4	11,2

0 100 200 300 400 500 600 700 800

▲ Landgemiddelde significant hoger dan TIMSS-gemiddelde
 ▼ Landgemiddelde significant lager dan TIMSS-gemiddelde

BRON: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

van Nederland. Overigens is de dekingsgraad van de TIMSS-toets vergeleken met het Nederlandse rekenonderwijs slechts 65 procent. Rekenonderdelen als *hoofdrekenen* en *schattend rekenen* kwamen in de TIMSS-toets niet aan bod.

Voor rekenen lijkt er sinds 1995 sprake te zijn van een zeer geleidelijke afname in toetsprestaties van de Nederlandse leerlingen in groep 6. De afname van 2007 ten opzichte van 2003 is niet significant, maar in 2007 én 2003 is de afname wel significant ten opzichte van 1995.⁴ Als deze trend zich doorzet, dreigt Nederland langzamerhand door een aantal landen te worden ingehaald. Dit komt mede doordat landen zoals Engeland en de Verenigde Staten verbeteringen laten zien in hun toetsscores.

Verschillen in prestaties tussen leerlingen

De prestaties van verschillende groepen leerlingen in Nederland laten enkele aanzienlijke verschillen zien.

Jongens-meisjes

Nederlandse jongens presteren op TIMSS-2007 voor rekenen en wiskunde significant beter dan meisjes. Het internationaal gemiddelde laat geen verschillen in toetsscores tussen jongens en meisjes zien. In twaalf landen presteren jongens gemiddeld beter dan meisjes, in acht landen is dat omgekeerd. Zowel van jongens als van meisjes zijn de Nederlandse prestaties ten opzichte van 2003 gedaald, voor meisjes echter significant. Ten opzichte van de meting in 1995 laten de jongens in 2007 echter wel een sterkere daling zien dan de meisjes.

Autochtone-allochtone leerlingen

Nederlandse autochtone leerlingen presteren significant beter dan allochtone leerlingen. De gemiddelde toetsscore van allochtone leerlingen ligt rond het TIMSS-schaalgemiddelde. Het verschil tussen autochtone en allochtone leerlingen is iets toegenomen vergeleken met 2003. Dit komt vooral door een sterke achteruitgang van de prestaties van allochtone meisjes. Waren er in 2003 nauwelijks verschillen tussen de gemiddelde toetsscores van allochtone jongens en meisjes, in 2007 scoren de allochtone meisjes gemiddeld 23 punten lager dan allochtone jongens. De achterstand is het grootst op het onderdeel *getallen*.

Prestaties naar inhoudelijk en cognitief domein

Zoals hierboven beschreven is, zijn de toetsopgaven ingedeeld naar drie inhoudelijke domeinen (*getallen*, *geometrische vormen/meten* en *gegevensweergave*) en naar drie

⁴ De vergelijking met de resultaten van 1995 vragen een kleine noot: In het voorgaande is toegelicht dat Nederland een extra aantekening in de internationale lijst met TIMSS resultaten 2007 krijgt, omdat het net niet voldeed aan de verplichte steekproef van originele scholen van 50 procent; dit was 49 procent. In 1995 voldeed Nederland eveneens niet helemaal aan de steekproefvereisten: in totaal moest 85 procent van de steekproef, inclusief de reservescholen, meedoen; dit werd net niet gehaald. Echter de steekproef die uiteindelijk deelnam, was wél representatief voor de Nederlandse populatie leerlingen in groep 6. Daarom is Nederland wel in het rapport opgenomen, maar heeft samen met enkele landen een aparte plaats. Volgens de bij TIMSS betrokken onderzoekers zijn desondanks betrouwbare vergelijkingen te maken met de resultaten van 2003 en 2007.

cognitieve domeinen (*weten, toepassen en redeneren*).

Nederlandse leerlingen scoren op alle inhouden gemiddeld ver boven het internationaal schaalgemiddelde van 500. Voor *gegevensweergave* scoren zij zelfs zeer hoog. Ze hebben relatief meer moeite met *geometrische vormen/meten*. Jongens doen het significant beter op het domein *getallen* dan meisjes. Op alle domeinen presteren autochtone leerlingen significant beter dan allochtone leerlingen; allochtone jongens blijken significant beter te presteren op de domeinen *getallen* en *geometrische vormen en meten* dan allochtone meisjes. Deze laatste groep presteert op het domein *getallen* maar liefst 32 punten lager dan allochtone jongens.

Op de cognitieve gebieden *weten* en *redeneren* scoren de Nederlandse leerlingen net onder de goed scorende Aziatische landen. Op het gebied *toepassen* doen Nederlandse leerlingen het relatief minder goed.

Niveaus van leerprestaties ten opzichte van referentiepunten

In TIMSS zijn op basis van referentiepunten vier verschillende niveaus geformuleerd: het geavanceerde niveau, het hoge niveau, het middenniveau en het lage niveau. In Nederland haalt 98 procent van de leerlingen het laagste niveau, een basaal kennisniveau. Het geavanceerde niveau wordt door slechts 7 procent van de Nederlandse leerlingen gehaald, terwijl dit in Singapore door de helft van de leerlingen wordt gehaald. Er zijn dus weinig leerlingen die het basisniveau niet halen, maar ook weinig leerlingen die het hoogste niveau halen.

Resultaten anders dan leerprestaties

Leerlingen en leraren hebben vragenlijsten voorgelegd gekregen over verschillende aspecten die samenhangen met het rekenonderwijs.

Resultaten van leerlingen op vragenlijsten

Aan leerlingen is gevraagd naar hun houding ten opzichte van het vak rekenen en wiskunde. Uit de resultaten blijkt dat de leerlingen evenals in 2003 een licht positieve houding ten opzichte van het vak rekenen en wiskunde hebben, al is die wel iets minder positief geworden en nemen Nederlandse leerlingen internationaal gezien een erg lage positie in. Allochtone leerlingen hebben significant meer plezier in rekenen en wiskunde dan autochtone leerlingen al zijn hun leerprestaties lager. Nederlandse leerlingen hebben vooral veel vertrouwen in hun eigen rekencapaciteiten, jongens significant meer dan meisjes. Ook al hebben allochtone meisjes significant een lagere toetscore dan autochtone meisjes, ze hebben wel ongeveer evenveel zelfvertrouwen in hun rekencapaciteiten.

Resultaten van leraren op vragenlijsten

De vragenlijst voor leraren van de getoetste leerlingen bevatte zowel vragen naar achtergrondkenmerken van de leraar en de school, als naar het gegeven rekenonderwijs.

Leraren in de groepen 6 voelen zich gemiddeld meer dan voldoende toegerust om les te geven in rekenen en wiskunde. Leraren met vijf jaar of minder ervaring en vrouwelijke leraren voelen zich relatief het minst toegerust in *getallen* en *geometrische vormen/meten*. Aan dit laatste domein wordt door Nederlandse leraren ook minder lestijd besteed in vergelijking met de andere landen. Aan het domein *getallen* wordt meer tijd besteed dan in andere landen.

Van alle TIMSS-landen ervaren Nederlandse groep 6-leraren de minste knelpunten in het omgaan met verschillen tussen leerlingen of probleemleerlingen tijdens de reken- of natuuronderwijslessen (tweederde wordt er niet door belemmerd). Men voelt zich relatief nog het meest belemmerd door de verschillen in leerniveaus van leerlingen: 41 procent een klein beetje, 29 procent enigszins en 12 procent voelt zich hierdoor erg belemmerd. In vergelijking tot TIMSS-2003 hebben ze aanmerkelijk minder na- of bijscholing gevolgd op het gebied van rekenonderwijs, ICT, pedagogiek/didactiek en het meten en beoordelen van leerprestaties; dit is bijna gehalveerd.

Wat betreft de aard en kwantiteit van het rekenonderwijs aan leerlingen zijn ook enkele gegevens verzameld over de Nederlandse situatie. Het aantal uren dat leerlingen per week aan rekenonderwijs besteden is sinds 1995 nauwelijks veranderd, nog steeds gemiddeld 4,5 uur. Ten opzichte van 2003 heeft het onderwijs in groep 6 een meer 'leerlinggeoriënteerd' karakter gekregen. Dit betekent ondermeer dat leerlingen gemiddeld minder vaak en tegelijkertijd met dezelfde leerstof bezig zijn, ze kijken gemiddeld vaker hun eigen werk of elkaars werk na, werken vaker met elkaar samen zonder tussenkomst van de leraar en werken vaker aan vakoverstijgende taken of projecten. In 2007 wordt in de helft van de klassen in groep 6 beperkt gebruik van de rekenmachine toegestaan. Sinds 1995 vindt dit in meer klassen plaats. De afgelopen twaalf jaar is de hoeveelheid huiswerk voor rekenen die leerlingen in groep 6 kregen verminderd. In 1995 gaf nog 50 procent van de leraren huiswerk op, in 2007 is dit nog 35 procent.

3.4 Commissie Meijerink: Over de drempels met rekenen

Naar aanleiding van tegenvallende reken- en taalresultaten en schooluitval in het voortgezet onderwijs heeft het Ministerie van OCW in 2007 een expertgroep verzocht referentieniveaus te ontwikkelen voor taal en rekenen voor de verschillende overgangen in het onderwijssysteem, van primair onderwijs tot en met hoger onderwijs. Het doel was het verhogen van de kwaliteit en het verbeteren van de overgangen tussen de verschillende onderwijsniveaus via het waarborgen van doorlopende leerlijnen. Het resultaat van deze commissie Meijerink (2008) 'Doorlopende leerlijnen taal en rekenen' is het rapport *Over de drempels met rekenen en taal* en voor rekenen en wiskunde het aanvullende rapport *Over de drempels met rekenen*.

De referentieniveaus zijn beschreven voor de leeftijden van ongeveer 12 jaar, 16 jaar en 18 jaar. Op elk van de referentiemomenten zijn twee verschillende kwaliteiten beschreven, een 'fundamentele kwaliteit' en een 'streefkwaliteit'. Voor het eind van het basisonderwijs houdt dit in dat er een fundamenteel niveau 1F en een streefniveau 1S

zijn geformuleerd. Deze niveaus beschrijven wat kinderen aan het eind van de basisschool moeten kennen en kunnen. Het gaat dus om een beheersingsniveau.

De referentieniveaus zijn beschreven in vier subdomeinen van samenhangende rekenfeiten, begrippen en operaties: (1) *getallen*, (2) *verhoudingen*, (3) *meten en meetkunde*, en (4) *verbanden*. Hierbij wordt onderscheid gemaakt in: 'paraat hebben', 'toepassen' en 'weten waarom'. De beschrijvingen van de referentieniveaus zijn op inhoudelijke overwegingen geformuleerd. Voor de referentieniveaus 1F en 1S is enerzijds gebruik gemaakt van empirisch kwantitatief onderzoek zoals PPO en L(O)VS (van Cito) en TIMSS, anderzijds is gekeken wat kinderen zouden moeten kennen en kunnen op twaalfjarige leeftijd op het gebied van rekenen. Een combinatie van haalbaarheid en wenselijkheid.

Fundamenteel niveau 1F wordt (in 2008) door 25 procent van de leerlingen niet gehaald. De ambitie is dat zoveel mogelijk van deze kinderen met extra inspanningen dit niveau wel gaan bereiken. Leerlingen voor wie 1F maximaal haalbaar is, zullen veelal instromen in de basisberoepsgerichte en kaderberoepsgerichte leerwegen van het vmbo, nu 35 procent van de totale reguliere leerlingenpopulatie in groep 8. Een schatting is dat zo'n 10 procent van de basisschoolpopulatie 1F niet zal kunnen halen, ook niet met extra inspanning. Voor deze groep wordt aanbevolen een afzonderlijk leertraject te formuleren. Naar schatting is er ook zo'n 20 procent van de leerlingen dat onder haar mogelijkheden presteert. Voor hen dient ook een extra aanbod gerealiseerd te worden.

Streefniveau 1S is het niveau voor leerlingen die naar vmbo theoretische leerweg (voorheen mavo), havo en vwo gaan. Dit niveau wordt thans door 50 procent van de leerlingen gehaald, maar moet volgens de commissie Meijerink in de toekomst door 65 procent van de leerlingen gehaald worden.

De bedoeling is dat de referentieniveaus bij wet worden vastgelegd. Ze zijn daarmee een aanvulling op de kerndoelen die 'slechts' een inspanningsverplichting formuleren. De referentieniveaus formuleren een opbrengstverplichting. In diverse projecten, subsidies en aanbevelingen van OCW wordt voor de komende tijd hoog ingezet op het concretiseren en behalen van de referentieniveaus voor taal en rekenen, ten einde de kwaliteit van het rekenonderwijs en de opbrengsten te verhogen en de overgang tussen de verschillende schooltypes te vereenvoudigen. Dit betekent dat in de toekomst de resultaten van nationale en internationale peilingen tegen deze referentieniveaus kunnen worden afgezet.

3.5 Samenvatting en conclusies

Nationale peilingen laten zien dat op veel onderdelen van het rekenen de prestaties de afgelopen twintig jaar betrekkelijk stabiel zijn gebleven en dat tegenover onderdelen die achteruit zijn gegaan andere onderdelen staan waarop vooruitgang is geboekt. In internationale vergelijkingen neemt Nederland nog steeds een sterke positie in. Toch is dit geen reden tot tevredenheid.

De commissie maakt zich grote zorgen om bepaalde ontwikkelingen in het rekenonderwijs. Het Nederlands rekenonderwijs kan en moet beter. Te weinig Nederlandse leerlingen behalen een geavanceerd rekenniveau. Een klein land als Nederland moet het van zijn innovatiekracht hebben. Het kan het zich niet veroorloven om niet vooruit te gaan in prestatiepeil in een wereld waarvan het economisch zwaartepunt zich naar het Oosten beweegt en waar de Aziatische landen qua rekenvaardigheid ver op ons land vooruit zijn. China en India ontbreken nog in TIMSS, maar de rekenaars uit die landen zijn niet weg te cijferen.

Het is ook niet verstandig om positieve en negatieve ontwikkelingen in het nationale rekenpeil tegen elkaar weg te strepen. De grote vooruitgang bij *getallen en getalrelaties* en bij *schattend rekenen* in PPON-2004 is op zich een fraaie realisatie, maar neemt niet weg dat het peil bij deze relatief nieuwe onderwerpen nog steeds achterblijft bij de gestelde standaard. Bij driekwart van de rekenonderdelen is er trouwens een grote kloof tussen het gerealiseerde peil en de standaard voor voldoende beheersing. De sterke achteruitgang bij *bewerkingen op grote getallen en kommagetallen* mag niet worden gerelativeerd door te wijzen op de rekenmachine en op een vorm van consensus over 'minder aandacht voor cijferen, meer aandacht voor inzicht en hoofdrekenen'. Leerlingen kiezen vaker voor uit het hoofd rekenen op een moment dat schriftelijk rekenen gewenst is, met slechtere resultaten als gevolg. Schriftelijk rekenen volgens standaardprocedures op de vier bewerkingen *optellen, aftrekken, vermenigvuldigen* en *delen* behoort tot de nationale kerndoelen.

De sterke positie van de Nederlandse rekenprestaties in West-Europa wordt bedreigd, enerzijds door de zeer geleidelijke daling van het Nederlands gemiddelde, anderzijds door de stijging van andere landen zoals Engeland. Overigens hebben niet alle West-Europese landen aan TIMSS meegedaan.

De Commissie Meijerink heeft in 2008 op verzoek van de minister van OCW doorlopende leerlijnen voor taal en rekenen voorgesteld. OCW wil hiermee de kwaliteit en de opbrengsten van het onderwijs verhogen en de drempels tussen onderwijstypes – van primair tot en met hoger onderwijs – verlagen. De leerlijnen zijn vastgelegd in termen van referentieniveaus. Waar de kerndoelen een inspanningsverplichting inhouden, formuleren de referentieniveaus een resultaatverplichting.

CONCLUSIE 3.1

De rekenvaardigheid van de Nederlandse leerlingen stemt niet tot tevredenheid, zowel in nationaal als internationaal perspectief. Voor bepaalde aspecten gaat het wel goed, maar er is een groot aantal aspecten waarbij het prestatiepeil en de ontwikkeling daarvan niet voldoen aan standaarden en referentieniveaus.

CONCLUSIE 3.2

Het is nuttig te beschikken over meetbare criteria als referentieniveaus. Daarbij merkt de commissie op dat alleen het vastleggen van dergelijke standaarden en niveaus niet voldoende zal zijn om het realiseren van de referentieniveaus te bewerkstelligen. Daarvoor zijn vervolgstappen door OCW, De Inspectie van het Onderwijs en schoolbesturen nodig.

4. RELATIE TUSSEN REKENDIDACTIEK EN REKENVAARDIGHEID

4.1 Gevolgde werkwijze

Een belangrijk gedeelte van de opdracht aan de commissie was: *Breng in kaart wat er wetenschappelijk gezien bekend is over de relatie tussen rekendidactiek en rekenvaardigheid.*

Voor de beantwoording van deze vraag heeft de commissie de volgende aanpak gevolgd.

De commissie heeft aan de hand van een aantal overzichtsartikelen en hoofdstukken uit toonaangevende handboeken een beknopte inventarisatie gemaakt van resultaten en inzichten uit de internationale literatuur over de effectiviteit van instructieprogramma's. Deze inventarisatie is uitgevoerd om het Nederlandse onderzoek in een ruimer kader te plaatsen en wordt in paragraaf 4.2 gepresenteerd.

Daarnaast is een grotere inventarisatie gemaakt van nationaal onderzoek dat rechtstreeks relevant is voor de beantwoording van de vraag. Om de kans te verkleinen dat belangrijke onderzoeken zich aan het zicht van de commissie zouden onttrekken, is aan ongeveer 50 onderzoekers gevraagd om de commissie te attenderen op relevante onderzoeken. Daarbij is gevraagd naar studies die voldoen aan de volgende voorwaarden:

- Het moet gaan over onderzoek dat zich specifiek of uitsluitend op rekenen richt, of waarbij men in ieder geval de effecten voor rekenen apart kan nagaan.
- Het moet gaan over onderzoek waarbij men de effecten voor kinderen in de leeftijdsgroep 4-12 jaar apart kan nagaan.
- Het moet gaan over onderzoek dat in de laatste twintig jaar is verricht.

- Het moet gaan over onderzoek dat in Nederland is verricht, in Nederlandse scholen en klassen, of om internationaal onderzoek waarbij men de bestudeerde effecten voor Nederland apart kan nagaan.
- Het moet gaan over empirisch onderzoek, d.w.z. om onderzoek waarin op basis van empirische data, van welke aard ook, uitspraken worden gedaan over de relatie tussen onderwijsaanpak en rekenvaardigheid van leerlingen.
- Het moet gaan om gepubliceerd onderzoek. Hierbij denken wij vooral aan publicaties in nationale of internationale wetenschappelijke tijdschriften, boeken en proefschriften.

Deze inperkingen zijn deels door praktische omstandigheden ingegeven. De commissie beschikte over de tijd noch de middelen om een uitgebreidere studie van alle direct en indirect relevante onderzoeksliteratuur te maken. De commissie realiseert zich dat er ook veel relevant onderzoek is dat niet aan bovenstaande voorwaarden voldoet en dus niet is opgenomen in de literatuurstudie. Vanwege het sterk leeftijdsgebonden karakter van onderwijsprocessen en de sterke historische, sociale en culturele bepaaldheid van deze processen achtte de commissie deze inperking echter in de gegeven omstandigheden verantwoord en noodzakelijk. Daarbij heeft de commissie geprobeerd de veel langere geschiedenis van het internationale onderzoek naar het leren en onderwijzen van rekenen en wiskunde in het algemeen niet te veronachtzamen.

De deskundigen hebben in totaal 76 publicaties aangedragen. Daarvan voldeden er 27 aan de gestelde voorwaarden. Na aanvullend literatuuronderzoek door de commissie zijn uiteindelijk 53 publicaties in detail geanalyseerd. In 34 publicaties (over 28 studies) werd een vergelijking tussen twee of meer instructieprogramma's of curricula gemaakt. Van deze studies zijn de groottes van de verschillen in rekenprestaties in kaart gebracht. De overige studies maakten geen vergelijking en zijn daarom niet in de overzichtsstudie opgenomen.

We hebben alleen onderzoeken opgenomen waarin resultaten over rekenprestaties of rekenvaardigheid werden gerapporteerd. Als daarnaast ook onderzoeksresultaten met betrekking tot strategiegebruik of motivatie beschikbaar waren, worden die ook besproken maar daar zijn geen kwantitatieve effecten voor berekend.

Paragrafen 4.3 en 4.4 beschrijven de resultaten van de literatuurstudie die door de commissie is uitgevoerd. De gevonden studies zijn ingedeeld in twee categorieën:

1. **Interventiestudies:** studies met als doel het vaststellen van het effect van één of meer ontwikkelde rekenprogramma's, waarin ofwel iets aan het reguliere curriculum wordt toegevoegd, ofwel een deel van het reguliere curriculum wordt vervangen. Het kenmerk van een interventiestudie is dat de onderzoekers interveniëren in het onderwijs. Deze studies worden beschreven in paragraaf 4.3. Daarbij zijn nog drie soorten studies onderscheiden: studies waarin minstens twee ontwikkelde programma's die een (deel van) een leergang vervangen met elkaar en eventueel nog met een controleprogramma vergeleken worden (paragrafen 4.3.1 en 4.3.2),

studies waarin het effect van één of meer remediële of extra oefenprogramma's wordt vastgesteld (paragraaf 4.3.3) en studies waarin een experimenteel programma wordt vergeleken met het zelfgekozen standaard curriculum (paragraaf 4.3.4).

2. **Curriculumstudies:** studies met als doel het vaststellen van verschillen in rekenprestaties bij verschillende curricula, meestal geoperationaliseerd als rekenmethoden. Het kenmerk van curriculumstudies is dat er door de onderzoekers niet wordt geïntervenieerd in het onderwijs; de gebruikte rekenmethoden zijn door de scholen zelf gekozen. Deze studies worden beschreven in paragraaf 4.4. De curriculumstudies zijn vervolgens nog verdeeld in domeinspecifieke studies (paragraaf 4.4.1) en grootschalige studies waarin algemene rekenvaardigheden worden vastgesteld (paragraaf 4.4.2).

Per samenhangend cluster van onderzoeken zijn in de paragrafen 4.3 en 4.4 conclusies weergegeven.

Om de kwantitatieve resultaten van verschillende studies onderling te kunnen vergelijken, geven we resultaten zoveel mogelijk in effectgroottes (*effect size*, ES) weer. Hierbij is, waar mogelijk, statistisch gecorrigeerd voor aanvangsverschillen, verschillen in rekenprestaties of andere relevante kenmerken tussen de groepen leerlingen die de onderzochte rekenprogramma's of curricula gingen volgen. Deze correctie wordt bijvoorbeeld ook ten sterkste aanbevolen door Slavin (2008). De interpretatie van de absolute grootte van de ES is als gebruikelijk gekozen:

- tussen 0,00 en 0,20: effect is verwaarloosbaar tot klein;
- tussen 0,20 en 0,50: effect is klein tot matig;
- tussen 0,50 en 0,80: effect is matig tot groot;
- groter dan 0,80: effect is groot.

Een ES van 0,20 of hoger wordt door Slavin (2008) in onderwijsstudies als praktisch relevant gekwalificeerd.

De verschillende studies hebben meestal betrekking op specifieke rekendomeinen, verschillende typen interventies en doelgroepen van leerlingen. De paragrafen 4.3 en 4.4 bevatten daardoor een hogere mate van detail dan de rest van het rapport. De lezer die vooral in de hoofdlijnen is geïnteresseerd kan zich desgewenst beperken tot de bespreking van het artikel van Slavin e.a. in paragraaf 4.2 en de samenvatting, conclusies en aanbevelingen in paragraaf 4.5.

4.2 Internationale literatuur

Deze paragraaf geeft een beknopt overzicht van de meest toonaangevende internationale literatuur.

Effective programs in elementary mathematics: A best-evidence synthesis (Slavin e.a., 2008)

In dit artikel worden de resultaten van een meta-analyse van (Amerikaans) onderzoek naar het effect van verschillende typen interventies op leerprestaties beschreven. Het gaat om de volgende drie types van interventies gericht op het verbeteren van het elementair wiskundeonderwijs:

1. aanpassingen in het reken- en wiskundecurriculum (bijvoorbeeld werken met een nieuw leerplan, een nieuw rekenboek);
2. computer-ondersteund onderwijs;
3. programma's die er primair op gericht zijn om via de beïnvloeding van de leraar het onderwijsleerproces te verbeteren.

In deze meta-analyse werden 78 studies betrokken die moesten voldoen aan een aantal strenge methodologische eisen. De belangrijkste conclusies waren de volgende:

- Er is weinig empirische evidentie voor een differentieel effect van het curriculum of de rekenmethode. In de bespreking van dit resultaat gaan de auteurs uitvoerig in op de veelgehoorde kritiek dat de meeste studies gebruik maken van gestandaardiseerde testen die vooral de meer traditionele vaardigheden meten en niet andere vaardigheden zoals bijvoorbeeld probleemoplossend vermogen. De auteurs erkennen dat dit een serieus probleem is maar geven aan dat de studies waarin geprobeerd is ook deze hogere-orde-vaardigheden in de evaluatie te betrekken, geen duidelijke winst ten voordele van de nieuwere programma's blijken op te leveren.
- De meeste studies rapporteren een positief effect van de inzet van computer-ondersteund onderwijs (*computer-assisted instruction, CAI*), vooral wat betreft de technische rekenvaardigheid. Geen enkele studie rapporteert een negatief effect voor CAI. Omdat het in de meeste gevallen gaat om een interventie die zelden meer inhoudt dan een *CAI supplement* van hoogstens een half uur per week naast het gebruikelijke rekenonderwijs, beschouwen de auteurs dit toch als een betekenisvol en bruikbaar resultaat, ook al is de grootte van de gevonden effecten meestal beperkt.
- De grootste effecten werden gevonden bij het derde type interventie, vooral bij de programma's gericht op coöperatief leren, op de verbetering van de klassenmanagementsvaardigheden van de leraar, op het verhogen van de motivatie bij de leerlingen of van *supplemental tutoring programs*. De algemene conclusie van deze meta-analyse is dan ook dat programma's die erop gericht zijn bepaalde aspecten van de dagelijkse klassenpraktijk te veranderen meer kans maken op een positief

effect op de leerresultaten van leerlingen, dan programma's die zich beperken tot het aanbrenge van veranderingen in het curriculum, het invoeren van een nieuw rekenboek of het invoeren van computerondersteunde leer- en oefenprogramma's. De auteurs concluderen dat de sleutel tot het verbeteren van de rekenvaardigheid ligt in het verbeteren van de interactie tussen leraar en leerling. De programma's die in deze analyse als succesvol naar voren kwamen, waren gericht op de manier waarop leraren les gaven en richtten zich op effectief tijdgebruik, het productief en betrokken houden van de leerlingen, de manier waarop kinderen werden gestimuleerd om elkaar te helpen en van elkaar te leren, en op de manier waarop leerlingen werden gemotiveerd.

The effects of classroom mathematics teaching on students' learning **(Hiebert e.a., 2007)**

In dit hoofdstuk van het *Second handbook of research on mathematics teaching and learning* wordt onderzoek naar de invloed van instructie op het wiskundeleren van leerlingen besproken. De auteurs wijzen allereerst op het aloude maar nog steeds geldende gegeven dat *opportunity to learn* de meest dominante invloedsfactor is op de prestaties van een leerling. Zij benadrukken daarbij echter dat *opportunity to learn* – in tegenstelling tot vroeger – niet alleen opgevat moet worden als het aanbieden van een bepaald onderwerp maar dat er kwaliteitseisen gesteld moeten worden aan het moment en de wijze waarop de leerling met dat leermateriaal wordt geconfronteerd.

De auteurs maken een onderscheid tussen onderwijzen voor het opdoen van vaardigheden en onderwijzen voor het creëren van conceptueel begrip. Voor het onderwijzen voor het opdoen van specifieke wiskundige vaardigheden bestaan er opsommingen van effectieve instructiekenmerken. Deze opsommingen zijn hoofdzakelijk afkomstig uit de algemene literatuur over schooleffectiviteit. De auteurs vatten deze kenmerken als volgt samen: *“Teaching that facilitates skill efficiency is rapidly paced, includes teacher modeling with many teacher-directed product-type questions, and displays smooth transition from demonstration to substantial amounts of error free practice. Noteworthy in this set of features is the central role played by the teacher in organizing, pacing, and presenting information to meet well-defined learning goals”*. Dit zijn echter niet de kenmerken die leiden tot wiskundige begripsvorming, zo concluderen deze auteurs op basis van hun overzicht van de beschikbare literatuur. Begripsvorming vereist andere instructiekenmerken, die de auteurs als volgt samenvatten: *“Teachers and students attend explicitly to concepts”* wat zij verderop omschrijven als: *“to connections among mathematical facts, procedures, and ideas in an explicit and public way”* (p. 383); *“students struggle and wrestle with important mathematics”* wat zij verder op omschrijven als: *“meaning that students expend effort to make sense of mathematics, to figure something out that is not immediately apparent”* en *“being the opposite of simply being presented information to be memorized or being asked to practice what has been demonstrated”* (p. 383-384).

How curriculum influences student learning (Stein e.a., 2007)

Dit hoofdstuk uit het *Second handbook of research on mathematics teaching and learning* geeft een overzicht van onderzoek over hoe het wiskundecurriculum het leren van leerlingen beïnvloedt. Het overzicht is grotendeels beperkt tot het onderzoek in de VS.

In de analyse hebben de auteurs een 35-tal gepubliceerde reken- en wiskunde-methoden betrokken die in de periode 2004-2005 in de VS gebruikt werden, op het niveau van het basisonderwijs, het lager secundair en het hoger secundair onderwijs. Iets meer dan de helft van deze methoden kon beschouwd worden als sterk geïnspireerd door realistische en constructivistische principes terwijl de andere helft als meer traditioneel of mechanistisch beschouwd kon worden.

In curriculumonderzoek moet, aldus de auteurs, onderscheid gemaakt worden tussen het *geschreven* curriculum, het *bedoelde* curriculum, en het *gerealiseerde* curriculum. Dus als men de vraag stelt 'Is curriculum x een goed curriculum?' of 'Is curriculum x effectiever dan curriculum y?', dan moet het duidelijk zijn welk soort curriculum men bedoelt en moet men beseffen dat er tussen het geschreven curriculum en het uiteindelijk leerresultaat van leerlingen heel wat tussenstappen zitten die deze relatie compliceren.

Effect van het geschreven curriculum

In het recente verleden zijn er in de VS drie grote studies verricht waarin men de inhoud (geschreven curriculum) van allerhande curriculummaterialen voor wiskundeonderwijs geanalyseerd, vergeleken en beoordeeld heeft: een review door the American Association for Advancement of Science (AAAS), een review van het Amerikaanse ministerie van onderwijs, en één uitgevoerd door de organisatie die pleit voor het behoud van meer traditioneel wiskundeonderwijs en die als de *Mathematically Correct* vereniging bekend staat. De auteurs van de eerste twee reviews waren realistisch geïntereerd, de laatste niet. De resultaten van deze drie reviews waren – niet verrassend – erg verschillend: Terwijl in de eerste twee reviews de vernieuwingsgezinde wiskundemethoden de beste beoordelingen kregen, was het in de laatste review andersom.

Wat valt te leren uit onderzoek naar het effect van het geschreven curriculum op de leerresultaten van leerlingen? Een eerste samenvatting van het relevant evaluatieonderzoek van de *reform-based curriculum materials* is te vinden in Senk en Thompson's (2003) volume *Standards-based mathematics curricula. What are they? What do students learn?* De globale conclusie van al die onderzoeken – gaande van basisschool tot de bovenbouw van het secundair onderwijs – is dat leerlingen die volgens de principes van de *reform-based math approach* onderwijs hadden gekregen het door de band genomen minstens even goed deden op de klassieke criteria van rekenkennis en -vaardigheid en het daarenboven beter deden op het vlak van de hogere-orde-doelen zoals inzicht, flexibiliteit, probleemoplossen, attitudes en *beliefs*. De auteurs voegen er

echter direct aan toe dat veel van deze studies de gouden standaard van het gerandomiseerd onderzoek met controlegroep niet haalden en dat er bovendien vaak een belangenconflict was en wel in die zin dat veel van deze evaluatiestudies verricht werden door onderzoekers die zelf meegewerkt hadden aan de ontwikkeling en implementatie van de vernieuwingsgezinde methode.

Dit leidde tot een tweede generatie van vergelijkende onderzoeken die wel aan de gouden standaard van het gerandomiseerd onderzoek met controlegroep beantwoordden en die bovendien verricht werden door belangeloze externe onderzoekers. Het aantal van dit soort onderzoeken is echter zeer beperkt; de auteurs bespreken er slechts twee. Hun conclusie is dat het patroon van de bevindingen van de externe reviewers overeenkomt met de eerder genoemde bevindingen van Senk en Thompson. In beide gevallen waren de resultaten van leerlingen die met meer vernieuwende curricula waren onderwezen gelijkwaardig op klassieke rekenvaardigheden en veel beter op redeneren en conceptueel begrip.

Stein e.a. wijzen erop dat sommige van de besproken vergelijkende onderzoeken wel beantwoorden aan de klassieke gouden standaard van deugdelijk vergelijkend methodenonderzoek, maar blind zijn voor de processen die zich afspelen tussen de wiskundemethode enerzijds en het uiteindelijke leerresultaat anderzijds. In deze onderzoeken wordt namelijk geen informatie ingewonnen over hoe een curriculum door de leraren begrepen en vertaald is en dan uiteindelijk in de klas gerealiseerd is.

Het hoofdstuk beschrijft verder onderzoek waarin duidelijk wordt dat er grote variaties zijn in de manier waarop leraren het curriculum interpreteren en zich eigen maken. Zij merken op dat de getrouwheid van de implementatie van een geschreven curriculum veel gemakkelijker is voor een traditioneel curriculum dan voor een meer vernieuwend curriculum. Ze merken ook op dat het als gevolg van deze variaties ook moeilijker is om de leeropbrengsten van deze vernieuwende curricula goed te meten.

Whole number concepts and operations (Verschaffel e.a., 2007)

Dit hoofdstuk uit het *Second handbook of research on mathematics teaching and learning* geeft een overzicht van de wetenschappelijke stand van zaken m.b.t. het leren omgaan en uitvoeren van bewerkingen met gehele getallen. Bij het begin van het hoofdstuk benadrukken de auteurs dat steeds meer onderzoekers en ontwikkelaars van wiskundeonderwijs van oordeel zijn dat leerlingen adaptieve expertise moeten ontwikkelen en niet alleen routinematige. Hieronder verstaan ze dat leerlingen ook in staat moeten zijn om geleerde procedures inzichtelijk, flexibel en creatief toe te passen. Het leren omgaan met gehele getallen wordt in dit verband gezien als een ervaring die de basis legt voor rekenvaardigheden zoals probleem oplossen, modelleren, patronen ontdekken, structuren analyseren en logisch redeneren.

Het eerste deel van het hoofdstuk gaat in op constaterend onderzoek. Dit is onderzoek dat onder gegeven omstandigheden plaatsvindt en waarbij de onderzoekers zelf niet interveniëren. Het tweede deel beschrijft onderzoek in de klas waarin men

vooral ingaat op ontwikkelingsexperimenten waarbij de onderzoekers wel ingrijpen in het onderwijsgebeuren. In het laatste deel plaatsen de auteurs het besproken onderzoek in een bredere historisch, sociale en politieke context. Interessante bevindingen uit het onderzoek zijn de volgende.

De literatuur van de laatste tien jaar heeft aanwijzingen opgeleverd dat het leren uitvoeren van bewerkingen met ééncijferige getallen veel meer omvat dan memoriseren. Dit domein van rekenen met gehele getallen demonstreert (1) dat de verschillende componenten van rekenkundige vaardigheden (feitenkennis, procedures, begripkennis) met elkaar in een voortdurende wisselwerking staan en elkaar wederzijds beïnvloeden, (2) dat kinderen bij de start van het onderwijs vaak al beschikken over rijke intuïtieve inzichten in en informele strategieën voor de diverse rekenkundige bewerkingen, die ze vervolgens geleidelijk aan uitbouwen tot meer efficiënte werkwijzen, en (3) dat leerlingen vaak adaptief weten te kiezen tussen de verschillende strategieën waarover ze beschikken, rekening houdend met de kenmerken van de opgave.

De belangrijkste thema's uit het onderzoek naar het rekenen met meercijferige getallen zijn min of meer dezelfde als bij het leren rekenen met ééncijferige getallen: (1) de grote variatie in de procedures die leerlingen hanteren bij het hoofdrekenend oplossen van deze sommen, (2) het flexibele of adaptieve gebruik van de verschillende hoofdrekenstrategieën, (3) de nauwe band tussen procedurele en conceptuele kennis bij ontwikkeling van de vaardigheid in het hoofdrekenen, en (4) het dispositioneel karakter van hoofdrekenen (d.w.z. dat hoofdrekenen geen louter cognitief proces is maar iets waar ook allerhande motivationele en affectieve aspecten mee gemoeid zijn). Bij dit laatste punt merken de onderzoekers op dat is aangetoond dat veel traditioneel onderwezen kinderen hun standaardalgoritmen stereotiep blijven toepassen ook in situaties waarin hoofdrekenstrategieën meer voor de hand liggen. Wat de relatie tussen hoofdrekenen en cijferrekenen betreft, signaleren ze dat de vraag naar het belang van cijferrekenen in de curricula nog steeds erg actueel is. Maar Verschaffel e.a. (2007, p. 576) schrijven '*... the idea that drilling students for long periods on solving problems involving large numbers by means of these written algorithms seems to be more appropriate to the twentieth century than the twenty-first (Kilpatrick et al., 2001, p. 214) is gaining ground.*'

In het onderdeel over *estimation and number sense* wordt gesteld dat schattend rekenen het curriculumonderdeel *par excellence* is waarin de bovenvermelde kenmerken van (leren) rekenen tot uitdrukking komen.

Het tweede deel van het hoofdstuk van Verschaffel e.a. (2007) handelt zoals gezegd over construerend onderzoek en meer bepaald over ontwikkelingsexperimenten. Bij het begin daarvan sommen de auteurs een aantal terugkerende kenmerken op van de recente onderzoekingen: (1) de erkenning dat leren rekenen een actief en constructief gebeuren is en van de cruciale rol die voorkennis daarin speelt; (2) de nadruk die er in de instructie op begripmatige kennis en op probleemoplossend denken wordt gelegd; (3) de reserves tegen het opleggen van één uniform denkmodel of één vaste oplossingsmethode aan alle leerlingen; (4) het intensief gebruik van interactieve

werkvormen; (5) de inspanningen die men zich getroost om het leren van rekenen in te bedden in betekenisvolle, authentieke contexten; (6) het zoeken naar een goed evenwicht tussen horizontale en verticale mathematisering; (7) het focussen op de grote ideeën die aan de specifieke leerinhouden ten grondslag liggen; (8) de overtuiging dat de bovenstaande principes niet enkel relevant zijn voor de sterkere leerlingen, maar tevens voor zwakkere rekenaars; (9) de aandacht voor de cruciale rol van de leraar bij de realisatie van onderwijs dat volgens de bovenstaande principes is opgevat.

In het laatste deel van het hoofdstuk gaan de onderzoekers onder andere in op de cruciale plaats en rol van de leraar bij de vernieuwing van het rekenonderwijs. Ze wijzen erop dat een verbetering van het onderwijs in bewerkingen met gehele getallen alleen mogelijk is door een radicale verandering in de scholing en professionele ontwikkeling van de leraren. Het grootste knelpunt dat ze hierbij zien is *“the common perception that elementary mathematics and the teaching of elementary mathematics are simple”*.

Foundations for success (US Department of Education, 2008)

De huidige discussie in Nederland over de voor- en nadelen van realistisch wiskundeonderwijs doet sterk denken aan wat er zich de voorbije jaren in de VS heeft afgespeeld. Vanuit een onvrede over de zwakke kwaliteit van het wiskundeonderwijs aldaar (o.a. gevoed door de relatief zwakke scores van leerlingen uit de VS. in internationale vergelijkende studies, zoals TIMSS) en een compleet verschillende duiding daarvan door de twee kampen die in de VS een *math war* aan het uitvechten zijn – de traditionalisten aan de ene kant en de hervormingsgezinden (*reformers*) aan de andere kant – heeft het Amerikaanse ministerie van onderwijs enkele jaren geleden een *National Mathematics Advisory Panel* (NMAP) ingesteld dat een rapport moest uitbrengen met aanbevelingen ter verbetering van het wiskundeonderwijs niveau K-8 (d.w.z. voor leerlingen van ongeveer 4 tot 14 jaar) gebaseerd op het beschikbare onderzoek. Concreet kreeg het panel de opdracht om advies uit te brengen over *“how to foster greater knowledge of and improved performance in mathematics among American students with respect to the conduct, evaluation, and effective use of the results of research relating to proven-effective and evidence-based mathematics instruction.”*

Het panel werd voorgezeten door Larry R. Faulkner, een gerenommeerde emeritus hoogleraar in de scheikunde. Omdat het panel deze belangrijke en omvangrijke opdracht binnen twintig maanden diende af te werken, werd het grootste deel van het werk verdeeld over een aantal subcommissies en werkgroepen, die deelrapporten hebben opgesteld die uiteindelijk samengevat en geïntegreerd zijn in een eindrapport. Op 13 maart 2008 gaf het NMAP het eindrapport vrij onder de titel *Foundations for Success; the final report of the National Mathematics Advisory Panel*.

Heel globaal samengevat zou men kunnen stellen dat het rapport pleit voor meer aandacht voor bepaalde inhoudelijke onderwerpen (bijv. gehele en rationale getallen,

algebra), voor meer focus en samenhang bij de behandeling van de betreffende wiskunde-inhouden, voor de geïntegreerde verwerking van rekenfeiten, -concepten en -procedures, voor een gezond en flexibel evenwicht tussen meer docent- en leerling-gerichte instructiemethoden (met een voorkeur – zeker voor de zwakke rekenaars – voor een meer directieve en sobere aanpak dan door vele hervormingsgezinden wordt voorgestaan), voor een stel stevige vakinhoudelijke en vakdidactische competenties bij (toekomstige) rekenleraren, voor meer middelen voor onderzoek (met name voor rigoureu (quasi-)experimenteel onderzoek naar de effectiviteit van bepaalde vormen van wiskundeonderwijs).

Al deze principes en aanbevelingen zijn gebaseerd op wat in het rapport omschreven wordt als de best beschikbare wetenschappelijke evidentie. Het panel legde de lat daarvoor – naar eigen zeggen – erg hoog door enkel rekening te houden met studies *that have been carried out in a way that manifested rigor and could support generalization at the level of significance to policy*. De commissie ontwikkelde hiervoor een serie algemene principes op basis van aspecten zoals excellentie van de opzet, validiteit en betrouwbaarheid, omvang en diversiteit van de onderzoekspopulatie.

Deze criteria werden verder verfijnd voor verschillende soorten onderzoek, o.a. voor interventiestudies waarvoor onderscheid gemaakt werd tussen studies van hoge, middelmatige en lage kwaliteit. Om tot de hoogste categorie te behoren moest het gaan om experimentele studies met willekeurige toewijzing aan condities, lage uitval en valide en betrouwbare metingen. Van middelmatige kwaliteit werden studies beschouwd waarbij het ontbreken van gerandomiseerde toewijzing van subjecten aan interventietypes gecompenseerd werd door koppelingstechnieken of statische controle achteraf; de andere methodologische en psychometrische eisen bleven dezelfde als voor de hoogste categorie. Interventiestudies die niet aan deze hoge eisen beantwoordden, werden als studies met lage kwaliteit bestempeld en ook als dusdanig behandeld. Alles bij elkaar zou men kunnen stellen dat volgens het panel slechts experimentele en quasi-experimentele (interventie)studies wetenschappelijk bewijs omtrent de waarde van bepaalde methoden of aanpakken van wiskundeonderwijs konden genereren. Van de meer dan 16.000 artikelen en andere onderzoeksrapporten die het panel doornam voldeed slechts een klein percentage aan de standaarden van wetenschappelijke evidentie. Sommige werkgroepen kwamen tot de conclusie dat er voor hun onderwerp nauwelijks relevant onderzoek beschikbaar was dat aan de hoge kwaliteitseisen voldeed. Zo stelde de werkgroep voor instructiemethoden dat ze uiteindelijk slechts acht bruikbare studies gevonden hadden. Wellicht mede doordat het panel slechts over een bedroevend beperkte set van bruikbare studies beschikte, besliste men om in bepaalde gevallen ook uitspraken en aanbevelingen te doen die niet gebaseerd waren op de bovenvermelde standaarden.

Het eindrapport kreeg veel kritiek te verduren. De kritische kanttekeningen slaan op de eenzijdige samenstelling van de commissie, de theoretische/inhoudelijke en methodologische uitgangspunten van het panel, en de simplistische wijze waarop het panel de relatie tussen onderzoek en ontwikkeling heeft opgevat. Vanuit theoretisch/

inhoudelijk oogpunt heeft de kritiek vooral betrekking op het feit dat men vooral van de cognitieve psychologie als theoretisch kader is uitgegaan en op het ontbreken van een grondige reflectie op doelen en inhouden van wiskundeonderwijs.

De scherpste kritiek gaat echter over het 'methodologisch credo' van het panel, in het bijzonder het grote belang dat is gehecht aan de 'gouden standaarden van de experimentele methode' (gerandomiseerd design, psychometrische tests, statistische onderbouwde conclusies). Hierdoor is volgens de critici met een te beperkte blik gezocht naar wetenschappelijke evidentie en is belangwekkend onderzoek buiten beschouwing gebleven. Dit heeft ook gevolgen voor de aanbevelingen over het soort onderzoek waar in de toekomst behoefte aan is. Een uitvoerige bespreking van de kritische reacties op dit eindrapport is te vinden in Verschaffel (2009).

Teachers knowledge and its impact (Fennema e.a., 1992)

Dit hoofdstuk uit het *Handbook of research on mathematics teaching and learning* beschrijft het effect van de vak kennis van de leraar bij het rekenonderwijs. De auteurs maken daarbij onderscheid tussen drie soorten vakspecifieke kennis van de leraar: kennis van rekenen en wiskunde, kennis van wiskundige representaties en kennis van leerlingen. De laatste twee soorten kennis worden door veel onderzoekers beschouwd als hoofdcomponenten van *pedagogical content knowledge* (PCK) of, met andere woorden, vakdidactische kennis. Daarnaast is er natuurlijk ook nog de algemene pedagogische kennis die niet specifiek is voor wiskunde.

Veel onderzoek heeft aangetoond dat de wiskundige kennis van leraren beperkt of ontoereikend is. Er is echter niet zoveel onderzoek beschikbaar dat overtuigend aantoont dat er een oorzakelijk verband is tussen het niveau van wiskundige kennis van de leraren en het prestatieniveau van hun leerlingen. Wel zijn er tal van kleinschalige *case studies* die laten zien hoe de inhoudelijke kennis de beslissingen die leraren nemen bij hun instructie in de klas beïnvloedt.

Volgens de auteurs levert het beschikbare onderzoek wel overtuigende aanwijzingen dat de PCK van de leraren hun feitelijke manier van onderwijzen en de resultaten van leerlingen bepaalt. Een bekend onderzoek in dit verband is het *Cognitively Guided Instruction* project in de VS dat in diverse deelstudies heeft aangetoond dat kennis die is verworven in onderzoek naar het denken van de leerlingen, door leraren gebruikt kan worden op een manier die de onderwijsresultaten beïnvloedt. Door metingen van de vakdidactische kennis van leraren, observaties van hun onderwijsgedrag en door toetsen bij leerlingen, voor en na de interventie, heeft men het positieve effect kunnen aantonen van een nascholing van leraren gericht op het verhogen van de PCK en vooral hun kennis van het denken van leerlingen. Dit positieve effect bleek niet alleen uit de verhoging van hun vakdidactische kennis, maar ook uit de verbetering van hun onderwijsgedrag én uit prestatieverbeteringen bij de leerlingen.

Assessing teachers mathematical knowledge. What knowledge matters and what evidence counts? (Hill e.a., 2007)

Dit hoofdstuk uit het *Second handbook of research on mathematics teaching and learning* geeft een overzicht van de geschiedenis van het evalueren van de verschillende aspecten van de professionele kennis en vaardigheden van rekenleraren. Dit gebeurt echter sterk vanuit een VS-standpunt en de relatie met de effectiviteit van hun onderwijs komt slechts zijdelings aan bod.

Verwijzend naar ouder onderzoek naar de relatie tussen rekenkennis van leraren en rekenleerprestaties van leerlingen concluderen de auteurs dat deze studies meestal een positief verband vinden. Daarbij gaat het louter om de rekenkennis van leraren in relatie tot de effectiviteit van hun onderwijs en dus niet om de vakdidactische kennis. Naderhand is er ook onderzoek verricht dat de relatie tussen de PCK van leraren in relatie tot de effectiviteit van hun onderwijs heeft onderzocht. De auteurs verzochten dan: *"If we had assessments that reliably and validly measure these different conceptions of mathematical knowledge for teaching, and could construct models of student achievement using teachers' scores on such assessments, we could specify more precisely the nature of knowledge that makes a difference for instructional quality and student learning. Without this precise specification, we cannot design pre-service and in-service learning experiences that equip teachers with this knowledge"*.

Zij beschrijven tot slot diverse lopende grootschalige projecten die dit doel trachten te realiseren en besluiten met de opmerking dat er nog veel onderzoek nodig is om te bepalen wat precies de vereiste rekenkennis voor het onderwijs is.

Mathematics interventions for children with special educational needs (Kroesbergen e.a., 2003)

In deze meta-analyse stond de vraag centraal wat de kenmerken van meest effectieve interventies op het gebied van rekenen zijn, voor kinderen met *special educational needs*. Dat zijn leerlingen die meer moeite hebben met rekenen dan hun leeftijdgenoten, die lager presteren dan hun leeftijdgenoten, of die speciale instructie nodig hebben om op een adequaat niveau te kunnen presteren. Deze leerlingen kunnen zowel op het regulier als op het speciaal basisonderwijs zitten.

Er werden interventies op drie domeinen van het basisschoolrekenen onderscheiden: (1) vroege rekenvaardigheden, zoals getalbegrip en tellen, (2) het automatiseren van basisvaardigheden, uiteenlopend van optellen tot 10 naar meercijferig delen, en (3) probleemoplossen, het weten hoe en wanneer bepaalde kennis toe te passen.

Na een uitgebreide zoekprocedure zijn 58 studies geselecteerd die aan de gestelde eisen voldeden. Voor elk van deze studies werden verschillende kenmerken op het gebied van methodologie, steekproef en interventiekenmerken gecodeerd. Ook werd voor elke studie de effectgrootte (*effect size*, ES) berekend. In het geval van een onderzoeksopzet met een vergelijking *tussen* groepen was dit het gestandaardiseerde

verschil tussen de gemiddeldes van de experimentele groep (met interventie) en de controlegroep (met een andere interventie of zonder interventie). In het geval van een vergelijking *binnen* leerlingen (herhaalde metingen volgens een pretest – interventie – posttest design) was dit het gestandaardiseerde verschil tussen de gemiddeldes van nameting en voormeting.

Met statistische meta-analysetechnieken werd vervolgens het effect van de verschillende kenmerken vastgesteld. Eerst werd dat met steeds één kenmerk tegelijk gedaan waardoor de samenhang tussen de verschillende kenmerken werd genegeerd. Vervolgens werden alle kenmerken tegelijk geanalyseerd, waardoor met de onderlinge samenhang tussen de kenmerken rekening kon worden gehouden, wat een statistisch adequatere aanpak is. We bespreken twee relevante conclusies met betrekking tot het effect van kenmerken van de interventie.

1. Er werden drie instructievarianten onderscheiden: directe instructie, zelfinstructie en *mediated/assisted* instructie. *Zowel directe instructie als zelfinstructie maken gebruik van een vast stappenplan voor het aanleren van een bepaalde strategie, die in het geval van directe instructie door de leerkracht wordt geïnstrueerd waarbij deze zijn eigen oplossingswijze overbrengt. Zelfinstructie is een procedure in vier stappen waarin de leerling afhankelijk van het getoonde begrip met weinig of veel verbale hulp – al dan niet met materiaal of schema ondersteund – een op zijn voorkennis afgestemde adequate oplossingsstrategie leert. In mediated/assisted instructie probeert de leerkracht uitgaande van de representaties/strategieën van leerlingen bij een gewenste oplossingsstrategie te komen, wat dus niet via een vast stappenplan kan.* In het algemeen was zelfinstructie het meest effectief vergeleken met directe instructie en *mediated/assisted* instructie. Maar op het gebied van de basisvaardigheden was directe instructie het meest effectief. Instructie m.b.v. de computer was minder effectief dan instructie door de leerkracht. *Peer tutoring*, waarbij leerlingen van elkaar leren was ook minder effectief dan instructie door de leerkracht. De auteurs concluderen dat de rol van de leerkracht cruciaal is.
2. Ten slotte zeggen de auteurs over de effecten van interventies die gebaseerd zijn op het realistisch rekenen, dat *mediated/assisted* instructie minder effectief was dan zelfinstructie en directe instructie. Verder zijn er geen significante effecten gevonden van het al dan niet aanwezig zijn van kenmerken van het realistisch rekenen.

Geconcludeerd werd dat directe instructie en zelfinstructie effectiever zijn dan *mediated/assisted* instructie. De eerste twee types vallen meer onder de traditionele aanpak, terwijl het laatste type vergelijkbaar is met de *guided reinvention* benadering van het realistisch rekenen. Er zijn dus aanwijzingen dat op dit aspect het realistisch rekenen voor de zwakke rekenaars minder goed uitpakt dan de traditionelere, meer voorgestructureerde, aanpak. Naast dit instructie-effect hadden de overige kenmerken van realistisch rekenen geen positief en geen negatief effect op de prestaties bij zwakke rekenaars.

4.3 Interventiestudies

Zoals beschreven in paragraaf 4.1 heeft de commissie een uitgebreide literatuurstudie laten uitvoeren naar de relatie tussen didactiek en rekenvaardigheid in de Nederlandse onderwijssituatie en deze vervolgens geanalyseerd. Deze paragraaf beschrijft de resultaten de analyse van de interventiestudies. De curriculumstudies worden besproken in paragraaf 4.4.

Het doel van interventiestudies is het vaststellen van het effect van één of meer rekenprogramma's. Een dergelijk programma komt in de plaats of voegt iets toe aan het curriculum zoals het gewoonlijk gevolgd zou worden. De implementatie van de programma's staat onder controle van de onderzoeker, al is het in verschillende gradaties. Zo kan het zijn dat externe trainers de programma's uitvoeren, wat de mogelijkheid biedt tot een relatief grote controle op de implementatie. In andere gevallen voert de hiertoe getrainde reguliere leraar het programma uit. Ook zijn combinaties mogelijk, bijvoorbeeld als een externe trainer het experimentele of remediële programma uitvoert, terwijl de reguliere leraar het standaardcurriculum uitvoert. Toewijzing van leerlingen aan de programma's gebeurt ofwel op het niveau van de individuele leerling, ofwel op het niveau van de klas of school. Deze toewijzing kan random gebeuren (experimenteel design), of niet random (quasi-experimenteel design). Er wordt al dan niet een voortoets (vóór aanvang van het programma) afgenomen.

De didactische invulling van de programma's kan sterk uiteenlopen tussen studies. Meestal is het programma slechts op een beperkt, afgebakend, domein van het rekenen gericht, zoals optellen en aftrekken tot 100. We merken daarom op dat geen van de studies mag worden gezien als een vergelijking tussen 'de' didactieken van het realistisch rekenen en traditioneel rekenen; zie ook conclusie 2.1.

Het trekken van eenduidige conclusies is moeilijk. In de onderzochte programma's worden vaak meerdere aspecten tegelijk gevarieerd, zoals de gehanteerde modellen (bijvoorbeeld de getallenlijn), het type instructie en de rol van de leraar (variërend van heel vrij tot heel directief), het type aangeboden opgaven (open problemen, contextopgaven, kale getalsopgaven) en de behandelde oplossingsstrategieën (standaardalgoritmen of informele strategieën). Het is dan niet mogelijk vast te stellen welk van de afzonderlijke aspecten de gevonden effecten heeft veroorzaakt.

In de bespreking van de onderzoeken worden drie varianten van interventiestudies onderscheiden:

1. studies waarin minstens twee ontwikkelde programma's die een (deel van) een leergang vervangen, met elkaar en eventueel nog met een controleprogramma vergeleken worden (paragraaf 4.3.1 en 4.3.2);
2. studies waarin het effect van één of meer remediële of extra oefenprogramma's wordt vastgesteld (paragraaf 4.3.3);
3. studies waarin een experimenteel programma wordt vergeleken met het zelfgekozen standaard curriculum (paragraaf 4.3.4).

4.3.1 Het effect van *guided* versus directe instructie bij zwakke rekenaars

Type onderzoek:	Interventiestudies waarbij vergelijking wordt gemaakt tussen twee of meer ontwikkelde programma's.
Onderzochte relatie:	Het effect bij zwakke rekenaars van <i>guided</i> versus directe instructie op automatiseren, rekenvaardigheid, transfer, strategiegebruik en motivationele en affectieve variabelen.
Overzicht:	Tabel A in bijlage 3.

In totaal zijn zes studies gevonden waarin twee of meer instructieprogramma's met elkaar vergeleken werden bij zwakke rekenaars, zowel uit het reguliere als uit het speciale basisonderwijs. In al deze studies bestonden de instructieprogramma's uit een variant met *guided* of banende instructie en één of meer varianten van directe of sturende instructie. Bij *guided* instructie dragen leerlingen zelf mogelijke oplossingsstrategieën aan, of worden er meerdere oplossingsstrategieën aangeboden door de leraar. De leerlingen kiezen zelf een strategie bij elke opgave. Bij directe instructie krijgen de leerlingen training in één standaardoplossingsstrategie. Dit wordt ook wel expliciete of gestructureerde instructie genoemd. De instructieprogramma's besloegen tussen de 26 en 34 lessen. Eén studie was gericht op vroege rekenvaardigheden bij kleuters. De overige vijf studies hadden betrekking op leerlingen met gemiddelde leeftijden tussen negen en tien jaar en richtten zich op *vermenigvuldigen* (twee studies) en op *optellen en aftrekken* tot 100 (drie studies). Qua effecten is vaak gekeken naar automatiseren/tempotoetsen, vaardigheidstoetsen (prestaties op het onderwezen leerstofgebied), en transfertoetsen (prestaties op een leerstofgebied dat in het verlengde ligt van wat onderwezen is).

Alle zes studies waren opgezet met een voormeting, interventie en nameting zodat verschillen op de nameting gecorrigeerd konden worden voor eventuele verschillen op de voormeting. Toewijzing aan condities gebeurde random op klasniveau, of op basis van *matching* van leerlingen binnen klassen. In Milo, Ruijsenaars en Seegers (2005) was de manier van selecteren en toewijzen van de leerlingen onduidelijk.

Omdat al deze studies het effect van *guided* instructie (GI) en directe instructie (DI) onderzochten en de opzet erg vergelijkbaar was, zijn de meest relevante resultaten weergegeven in tabel 4.1.

Gevonden effecten op automatiseren

In vier van de zes studies werden resultaten met betrekking tot automatiseren gerapporteerd. In twee studies werd een klein tot matig verschil in prestaties in het nadeel van *guided* instructie gevonden: in Kroesbergen en Van Luit (2002) met $ES = -0,51$ (maar het verschil op voormeting was eigenlijk te groot om statistisch voor te corrigeren, bovendien geheel op conto van de leerlingen in het speciaal basisonderwijs) en in Timmermans en Van Lieshout (2003) met geschatte $ES = -0,23$. In de overige twee studies werd een verwaarloosbaar klein verschil gevonden: in Kroesbergen, Van Luit

en Maas (2004) met ES = +0,03 en in Timmermans, Van Lieshout en Verhoeven (2007) met ES = +0,05.

Samenvattend: Voor automatiseren is er bij zwakke rekenaars ofwel geen verschil, ofwel een verschil in het nadeel van de *guided* instructie gevonden. Verschillen ten nadele van *guided* instructie werden alleen bij zwakke rekenaars uit het speciaal onderwijs gevonden.

Tabel 4.1 Resultaten van de interventiestudies bij zwakke rekenaars (*guided* instructie vs. directe instructie), in effectgroottes, gecorrigeerd voor verschillen op voormeting.

Studie	type basis-onderwijs	ES: GI – DI automatiseren	ES: GI – DI vaardigheid	ES: GI – DI transfer
Kroesbergen en Van Luit (2002)	reg. + spec.	[-0,51]	+0,43	+0,52
	<i>speciaal</i>	[-2,42]	+0,32	+0,36
	<i>regulier</i>	[+0,64]	[+0,85]	[+1,03]
Kroesbergen e.a. (2004)	reg. + spec.	+0,03	-0,30	n.v.t.
Milo e.a. (2005)	speciaal	n.v.t.	-0,73 (DI-rijg)	+0,07* (DI-rijg)
			-0,21 (DI-splits)	+0,59* (DI-splits)
Timmermans en Van Lieshout (2003)	speciaal	-0,23 [#]	0,00 [#]	-0,57*
Timmermans e.a. (2007)	regulier	+0,05	+0,13	n.v.t.
	<i>meisjes</i>	+0,07	+0,84	n.v.t.
	<i>jongens</i>	+0,03	-0,51	n.v.t.
Van de Rijt en Van Luit (1998)	regulier	n.v.t.	+0,20	n.v.t.

Noot 1. Waarde effectgrootte tussen []: verschil op voormeting groter dan 0,5 standaarddeviatie: adequate statistische correctie niet goed mogelijk. Gerapporteerde gecorrigeerde ES is onbetrouwbaar.

Noot 2. Waarde effectgrootte met #: verschil geschat op basis van beschikbare gegevens, waarbij ES op 0 is gezet als slechts gerapporteerd is dat het verschil niet significant was, maar geen gemiddelden zijn gegeven.

Noot 3. Waarde ES met *: geen statistische correctie voor aanvangsverschillen.

Gevonden effecten op rekenvaardigheid

Alle zes studies rapporteerden effecten met betrekking tot vaardigheid – de prestaties op het onderwezen leerstofgebied. In twee studies werd een verschil in het voordeel van de *guided* instructie gevonden: in Kroesbergen en Van Luit (2002) een klein tot matig effect (ES = +0,43) en in Van de Rijt en Van Luit (1998) een klein effect (ES = +0,20). In twee studies werd een verwaarloosbaar tot klein verschil gevonden: in Timmermans en Van Lieshout (2003) was geen ES te schatten, en in Timmermans e.a. (2007) met een ES van +0,13. In deze laatste studie was wel een relevant verschil tussen jongens en meisjes aanwezig: voor meisjes was een groot verschil in het voordeel van *guided* instructie (ES = +0,84), terwijl er voor jongens een matig verschil in het voordeel van directe instructie was (ES = -0,53). In twee studies werd een verschil in gemiddelde prestaties ten nadele van de *guided* instructie gevonden. In Kroesbergen

e.a. (2004) een ES van $-0,30$. In Milo, Ruijsenaars en Seegers (2005) leidde *guided* instructie zowel tot lagere prestaties dan directe instructie in de rijgstrategie (ES = $-0,73$) als tot directe instructie in de splitsstrategie (ES = $-0,21$).

Samenvattend: Voor rekenvaardigheid zijn de resultaten niet eenduidig. Er spelen dus andere factoren dan alleen de instructievariant een rol.

Gevonden effecten op transfer

Drie studies rapporteerden resultaten met betrekking tot transfer, het toepassen van geleerde kennis of vaardigheden in nieuwe situaties. Een voorbeeld is het toepassen van vaardigheden op een getalldomein dat buiten de instructie viel. Alleen bij Kroesbergen en Van Luit (2002) was statistische correctie voor aanvangsverschillen mogelijk, en werd er een matig verschil in het voordeel van *guided* instructie gevonden (ES = $+0,52$). In Milo e.a. (2005) werd een (ongecorrigeerd) verwaarloosbaar tot klein verschil in het voordeel van *guided* instructie in de rijgstrategie (ES = $+0,07$) gevonden en een matig tot groot positief verschil in het voordeel van *guided* instructie in de splitsstrategie (ES = $+0,59$). Timmermans en Van Lieshout (2003) ten slotte vonden een matig tot groot (ongecorrigeerd) verschil in het voordeel van de directe instructie (ES = $-0,57$).

Samenvattend: Voor transfer zijn de resultaten niet eenduidig. Ook op dit onderdeel lijken andere factoren dan alleen de instructievariant een rol te spelen.

Andere uitkomstgebieden: strategiegebruik en motivationele variabelen

Met betrekking tot strategiegebruik blijken direct geïnstrueerde leerlingen vaker de geïnstrueerde strategie te gebruiken dan *guided* geïnstrueerde leerlingen. *Guided* geïnstrueerde leerlingen waren echter niet flexibeler in hun strategiegebruik, waarmee bedoeld wordt dat ze hun grotere strategierepertoire niet adaptief inzetten op verschillende opgaven. Op motivationele en affectieve variabelen werden nauwelijks instructie-effecten gevonden.

Samenvattend: Direct geïnstrueerde leerlingen gebruikten vaker de geïnstrueerde strategie dan *guided* geïnstrueerde leerlingen. *Guided* geïnstrueerde leerlingen waren echter niet flexibeler in hun strategiegebruik. Op motivationele en affectieve variabelen werden nauwelijks instructie-effecten gevonden.

CONCLUSIE

Er is geen eenduidige conclusie te trekken over de effecten van *guided* versus directe instructie bij zwakke rekenaars. Het patroon van relatieve effecten op prestaties van *guided* versus directe instructie is wisselend. Er worden zowel positieve als negatieve effecten gevonden, en dat zowel binnen één studie op de verschillende uitkomstgebieden, als tussen de verschillende studies op hetzelfde uitkomstgebied. Kennelijk

zijn de niet gemeten factoren zoals leraar, samenstelling van de klas en implementatie van de instructie belangrijker dan het gekozen type instructie. Een onderscheid dat zeker nader onderzoek verdient is dat naar geslacht: slechts één studie rapporteerde over differentiële effecten voor jongens en meisjes en er werden grote verschillen in instructie-effecten gevonden. Verder bleken leerlingen onder *guided* instructie een groter strategierepertoire te hebben maar die verschillende strategieën niet flexibeler of adaptiever in te zetten.

4.3.2 Vergelijking tussen twee instructieprogramma's bij reguliere leerlingen

Type onderzoek:	Interventiestudies waarbij vergelijking wordt gemaakt tussen twee of meer ontwikkelde programma's.
Onderzochte relatie:	Het effect bij reguliere leerlingen van het <i>realistic program design</i> versus het <i>gradual program design</i> op rekenvaardigheid bij 2-cijferig optellen en aftrekken en het effect van twee instructieprogramma's voor het werken met procenten en grafieken.
Overzicht:	Tabel B in bijlage 3.

In twee studies werden de prestaties van alle leerlingen uit het reguliere basisonderwijs met twee ontwikkelde instructieprogramma's vergeleken. In de studie van Klein (1998) betrof dit twee instructieprogramma's voor het tweecijferig optellen en aftrekken, en in de studie van Van Dijk (Terwel, Van Oers, Van Dijk en Van Eeden, 2009; Van Dijk, Van Oers, Terwel en Van Eeden, 2003) ging het om twee instructieprogramma's voor het leren werken met procenten en grafieken.

Realistic versus gradual programma

De studie van Klein (1998; zie ook Klein, Beishuizen en Treffers, 1998; Blöte, Van der Burg en Klein, 2001) vergeleek het *realistic program design* (RPD) en het *gradual program design* (GPD) voor het tweecijferig optellen en aftrekken. In het RPD lag de nadruk op het door de leerlingen zelf creëren en bespreken van strategieën. Er werden veel contexten gebruikt en flexibel strategiegebruik werd benadrukt. In GPD was de instructie meer traditioneel met een stapsgewijze kennisopbouw beginnend bij één basale procedure voor optellen en aftrekken, het rijgen.

Er was geen voormeting, dus effecten konden niet statistisch gecorrigeerd worden voor aanvangsverschillen. Op de natoetsen bleken de verschillen in prestaties op tempotoetsen ($ES = +0,19$), strategieëntoets ($ES = +0,15$), optel- en aftrekopgaven met kladpapier ($ES = +0,10$), de gestandaardiseerde leerlingvolgsysteem-toets (ES niet te schatten, maar verschil niet significant), transferopgaven ($ES = -0,03$) en retentie ($ES = +0,20$) alle verwaarloosbaar tot klein en op één uitzondering na licht in het voordeel van RPD.

Bij de tempotoetsen, strategieëntoetsen en toetsen met kladpapier zijn de verschillen tussen RPD en GPD ook apart voor sterke en zwakke rekenaars gerapporteerd. Voor de sterke rekenaars was er een matig verschil op de tempotoetsen in het voordeel van het RPD (ES = +0,47) terwijl op de strategieëntoets en kladpapiertoets het verschil verwaarloosbaar (ES = +0,02) respectievelijk licht in het nadeel van RPD was (ES = -0,15). Hierbij moet ook opgemerkt worden dat de groep sterke RPD-leerlingen bij aanvang beter presteerde dan de groep sterke GPD-leerlingen op een algemene vaardigheidstoets van het leerlingvolgsysteem (ES = +0,50) en dat voor dit verschil niet gecorrigeerd is. Verder bleek dat voor de zwakke rekenaars de verschillen klein tot groot in het voordeel van het RPD waren (respectievelijke effectgroottes +0,57, +0,31 en +0,36), wat er volgens de auteurs op wijst dat een didactische aanpak waarbij de leerlingen zelf hun strategieën creëren en bespreken ook bij zwakkere rekenaars effectief kan zijn. Ten slotte werd gevonden (niet in tabel B) dat het RPD tot meer flexibel strategiegebruik leidde dan het GPD, zowel bij sterke als bij zwakke rekenaars. De instructieprogramma's toonden ook – meestal verwaarloosbare tot kleine – verschillen wat betreft hun effect op diverse affectieve en motivationele aspecten, die bijna altijd in het voordeel van het RPD waren, zowel bij sterke als bij zwakke rekenaars.

Samenvattend: De effecten op het gebied van prestaties waren verwaarloosbaar tot klein in het voordeel van het RPD voor de gehele groep leerlingen. De zwakke rekenaars hadden een klein tot groot voordeel bij het RPD, bij de sterke rekenaars liep dit uiteen van een matig voordeel voor het RPD tot een licht nadeel voor het RPD. Het RPD leidde ook tot meer flexibel strategiegebruik dan het GPD en tot verwaarloosbare tot iets betere effecten op affectief en motivationeel gebied, zowel bij sterke als bij zwakke rekenaars.

Instructieprogramma voor procenten en grafieken

Terwel e.a. (2009; zie ook Dijk e.a., 2003) vergeleken twee lesprogramma's op het gebied van procenten en grafieken. In het *co-constructing/designing* programma leerden de leerlingen hoe ze zelf modellen en representaties voor de aangeboden open, complexe probleemsituaties moesten ontwerpen, in samenwerking met hun medeleerlingen en onder begeleiding van de leerkracht. In het *providing* programma leerden de leerlingen werken met kant-en-klare modellen die door de leraar verschaft werden. Leerlingen werkten individueel aan de opdrachten, waarna ze met de klas besproken werden. Volgens de auteurs komt deze *providing* conditie overeen met wat gebruikelijk is in de Nederlandse schoolpraktijk.

Leerlingen in het *co-constructing/designing* programma beter presteerden dan leerlingen in het *providing* programma op opgaven over procenten en grafieken (ES = +0,32) en ook op transferopgaven (ES = +0,55).

Samenvattend: Het *co-constructing/designing* programma had een klein tot matig positief effect op de prestaties vergeleken met het *providing* programma.

CONCLUSIE

In beide studies had het meer realistisch georiënteerde programma een verwaarloosbaar klein tot matig positief effect op de prestaties.

4.3.3 Het effect van remediële programma's bij zwakke rekenaars

Type onderzoek:	Interventiestudies waarbij het effect van één of meer remediële of extra oefenprogramma's wordt vastgesteld.
Onderzochte relatie:	Het effect bij zwakke rekenaars in het speciaal basisonderwijs van remediële programma's gericht op optellen en aftrekken tot 100 en het effect bij zwakke rekenaars in het reguliere basisonderwijs van remediële programma's gericht op cijferend rekenen en productief oefenen.
Overzicht:	Tabel C (speciaal basisonderwijs) en tabel D (regulier basisonderwijs) in bijlage 3.

Er is één studie gevonden waarin het effect van remediële programma's voor leerlingen uit het speciaal basisonderwijs is onderzocht. Dit betrof een remedieel programma voor het leren optellen en aftrekken tot 100 gebaseerd op individuele oplossingsstrategieën (Harskamp en Suhre, 1995).

Er zijn vier studies gevonden waarin het effect van een remedieel (drie studies) of extra oefenprogramma (één studie) bij zwakke rekenaars uit het reguliere basisonderwijs is vastgesteld. Willemsen (1994) vergeleek eerst twee remediële programma's (studie 1) en vervolgens drie (studie 2). Harskamp, Suhre en Willemsen (1993) onderzochten de effecten van een remedieel rekenprogramma bij zwakke rekenaars, specifiek gericht op de combinatie van een realistische rekenmethode met een structuralistisch versus een realistisch remedieel rekenprogramma. Menne (2001) ten slotte onderzocht het effect van een productief oefenprogramma.

Remedieel programma voor optellen en aftrekken

Harskamp en Suhre (1995) ontwikkelden een nieuw remedieel programma op basis van het aansluiten bij individuele strategieën van leerlingen (de realistische didactiek), speciaal voor zwakke rekenaars uit het speciaal basisonderwijs (LOM en MLK). Dit programma had een groot positief effect op de prestaties op een natoets en retentietoets met kale opgaven met $ES = +3,22$ (adequate statistische correctie niet mogelijk), ook positieve effecten apart voor LOM-leerlingen $ES = +3,13$ (adequate statistische correctie niet mogelijk) en MLK-leerlingen $ES = 3,69$. Het programma had eveneens een groot positief effect op de prestaties op toepassingsopgaven bij LOM-leerlingen ($ES = +3,58$, adequate statistische correctie niet mogelijk) en MLK-leerlingen ($ES = 3,58$).

Samenvattend: Dit remediële programma had zowel bij LOM- als MLK-leerlingen grote positieve effecten vergeleken met een controlegroep op de prestaties bij optellen en aftrekken tot 100.

Remediële programma's in het regulier basisonderwijs

In twee studies vergeleek Willemsen (1994) verschillende remediële programma's op het gebied van het cijferend aftrekken: het *mapping*-programma op basis van remediëring van conceptuele misvattingen ten grondslag aan systematische rekenfouten, het kolomsgewijze programma op basis van introductie van een alternatief voor het standaardalgoritme, en een controleprogramma op basis van systematisch herhalen en oefenen van de leerstof, dat wil zeggen, het standaardalgoritme. Uit studie 1 bleken de prestaties op de natoets bij leerlingen in het *mapping*-programma beter dan die van leerlingen in het controleprogramma ($ES = +0,32$), een klein tot matig verschil. In studie 2 bleek het *mappings*programma opnieuw een positief effect op de prestaties te hebben gehad, ditmaal met matige tot grote verschillen op de natoets ($ES = +0,74$) en op de retentietoets ($ES = +0,83$). Leerlingen in het kolomsgewijze programma presteerden op de natoets wat slechter dan leerlingen in het controleprogramma ($ES = -0,17$) en op de retentietoets juist weer wat beter ($ES = +0,20$). Verder bleken leerlingen in het *mapping*-programma ook minder systematische fouten te maken dan leerlingen in het controleprogramma (studie 1 en 2) en dan leerlingen in het kolomsgewijze programma (studie 2).

Samenvattend: Het *mapping*-programma voor de remediëring van conceptuele misvattingen die ten grondslag liggen aan systematische rekenfouten had een klein tot matig positief effect op de prestaties, vergeleken met het kolomsgewijze programma en een controleprogramma.

Harskamp e.a. (1993) onderzochten de prestaties van leerlingen in zes verschillende combinaties van realistische rekenmethoden en remediëel pakket (ofwel structuralistisch, ofwel realistisch). De prestaties op kale rekenopgaven lagen bij de zes remediële condities hoger dan bij de controlegroep, zowel in groep 4 ($ES = +1,18$, maar verschil op de voormeting te groot om adequaat statistisch voor te corrigeren) als in groep 5 ($ES = +0,39$). Op toepassingsopgaven werden kleine effecten in het voordeel van de experimentele condities gevonden, in groep 4 ($ES = +0,17$) en in groep 5 ($ES = +0,24$). De zes experimentele condities verschilden onderling niet van elkaar in prestaties op kale opgaven of op toepassingsopgaven. Verder bleek uit onderzoek naar de implementatie (niet in overzichtstabel D) dat de feitelijke uitvoering van de verschillende remediële pakketten – die ofwel realistisch ofwel structuralistisch van aard waren – niet verschilde.

Samenvattend: De remediële rekenprogramma's hadden een positief effect op de prestaties van zwakke rekenaars, vooral bij de kale basisvaardigheden, ongeacht de didactische kenmerken van het remediële rekenprogramma en de combinatie met didactische kenmerken van de reguliere rekenmethode.

Oefenprogramma

Menne (2001) ontwikkelde en implementeerde een productief oefenprogramma, geënt op basaal tellend rekenen met eenheden en tientallen, en erop gericht leerlingen vlot en flexibel met sprongen van wisselende lengte over de (lege) getallenlijn te laten springen in de lijn van de aanvankelijke realistische rekendidactiek. De prestaties van leerlingen die het productieve oefenprogramma gevolgd hadden lagen hoger dan die van leerlingen die dat niet hadden gedaan: op LVS-toetsen een geschatte ES = +0,44, en op de MORE-toetsen (zie Gravemeijer e.a., 1993) een geschatte ES = +0,30. Het verschil in prestaties tussen wel en geen oefenprogramma was op de LVS-toetsen groter voor allochtone leerlingen (geschatte ES = +0,59) dan voor autochtone leerlingen (geschatte ES = +0,41). Dit patroon werd ook gevonden op de MORE-toetsen: voor allochtone leerlingen geschatte ES = +0,45; voor autochtone leerlingen geschatte ES = +0,18.

Samenvattend: Het productieve oefenprogramma had een klein tot matig positief effect op de rekenprestaties, vooral voor allochtone leerlingen.

CONCLUSIE REMEDIËLE PROGRAMMA'S EN EXTRA OEFENPROGRAMMA'S

In alle studies is een positief effect gevonden op de prestaties van remediële programma's of extra oefenprogramma in vergelijking met controlegroepen die slechts de reguliere rekenlessen volgden. Dit effect is zowel in speciaal als in regulier basisonderwijs gevonden. Qua didactiek van de remediëring lijkt het aansluiten bij individuele strategieën en systematische fouten goede resultaten op te leveren.

4.3.4 Het effect van een experimenteel programma vergeleken met het standaardcurriculum

Type onderzoek:	Interventiestudies waarbij het effect van een experimenteel programma wordt vergeleken met het zelfgekozen standaard curriculum.
Onderzochte relatie:	Effect van experimentele instructieprogramma's bij leerlingen in het speciaal onderwijs op vroege rekenvaardigheden en vermenigvuldigen en delen tot 100 en bij leerlingen in het reguliere onderwijs op schematiseren en breuken.
Overzicht:	Tabel E (speciaal basisonderwijs) en tabel F (regulier basisonderwijs) in bijlage 3.

In deze paragraaf worden studies besproken waarin prestaties van leerlingen die een experimenteel programma hebben gevolgd, worden vergeleken met die van leerlingen die het door de school zelfgekozen standaardcurriculum hebben gevolgd. Hierbij moet opgemerkt worden dat naast het instructieprogramma vaak ook de instructeur (extern persoon bij experimenteel programma, reguliere leraar bij controlegroep) en instructie-eenheid (in kleine groepjes buiten de klas bij experimenteel programma, in de klas bij controlegroep) verschillend waren. Het is daarom niet goed mogelijk vast te

stellen aan welk van deze afzonderlijke aspecten eventuele verschillen vooral zijn toe te schrijven.

4.3.4.1. Speciaal basisonderwijs

Er zijn drie studies opgenomen waarin het effect van een experimenteel instructieprogramma is vergeleken met het standaardcurriculum in het speciaal basisonderwijs. Twee studies waren gericht op vroege rekenvaardigheden bij kleuters, de andere studie richtte zich op vermenigvuldigen en delen tot 100.

VROEGE REKENVAARDIGHEDEN

In de eerste studie (Schopman en Van Luit, 1996) werd het effect onderzocht van een interventieprogramma gericht op het leren tellen tot 10 als voorbereiding op het formele rekenonderwijs in groep 3. Leerlingen in de experimentele groep⁵ volgden dit programma, en zij presteerden beter dan leerlingen uit de controlegroep op vroege rekenvaardigheden ($ES = +1,07$): een groot effect. In de tweede studie (Van Luit en Schopman, 2000) was het interventieprogramma uitgebreid naar meer sessies en naar het getalengebied tot 15. Ook in het tweede onderzoek presteerden kinderen in de experimentele groep beter dan leerlingen uit de controlegroep op vroege rekenvaardigheden ($ES = +0,73$) en ook op een transfertoets (ongecorrigeerde $ES = +0,22$).

Samenvattend: In beide studies presteerden leerlingen die het interventieprogramma gericht op het leren tellen tot 10 of 15 als voorbereiding op het formele rekenonderwijs in groep 3 volgden, beter dan leerlingen uit de controlegroep op vroege rekenvaardigheden, met een matig tot groot effect.

VERMENIGVULDIGEN EN DELEN TOT 100

Van Luit en Naglieri (1999) hebben het MASTER-programma⁶ ontwikkeld, gericht op de ontwikkeling van oplossingsstrategieën voor vermenigvuldigen en delen tot 100, gebruikmakend van realistische onderwijsprincipes als zelfinstructie, discussie en reflectie. Leerlingen die deze MASTER-training hadden gevolgd presteerden op de na-toets veel beter dan leerlingen uit de controlegroep ($ES = +2,16$). Dit gold zowel voor LOM-leerlingen ($ES = +2,50$) als voor MLK-leerlingen ($ES = +3,08$). Dit gold ook voor de vervolgoets (niet in overzichtstabel E). LOM-leerlingen bereikten met het MASTER-programma ook positieve effecten op verre transfer (opgaven die een stuk moeilijker

5 Schopman en Van Luit (1996) onderscheidden eigenlijk twee experimentele groepen: één met *guiding* instructie, en één met *directing* instructie. Volgens de auteurs bleek de praktische uitvoering van de instructie echter niet van elkaar te onderscheiden. De resultaten van de beide experimentele groepen zijn daarom samengenomen.

6 In beide studies van Kroesbergen (zie par. 4.3.1) is een bewerkte versie van het MASTER-programma voor vermenigvuldigen gebruikt. De vergelijking tussen beide experimentele condities enerzijds en de controlegroep anderzijds zou daarom ook goed in de huidige paragraaf passen. Er werden kleine, matige en grote positieve effecten van het experimentele programma gevonden, zowel bij leerlingen uit het speciaal als uit het regulier basisonderwijs.

dan de onderwezen leerstof waren), maar MLK-leerlingen niet (niet in overzichtstabel E).

Samenvattend: De MASTER-training gericht op de ontwikkeling van oplossingsstrategieën voor vermenigvuldigen en delen tot 100, gebruikmakend van zelfinstructie, discussie en reflectie, had een zeer groot positief effect op de prestaties.

4.3.4.2. Regulier basisonderwijs

Ook in het reguliere basisonderwijs zijn drie studies gevonden waarin een experimenteel programma werd afgezet tegen het door scholen zelf gekozen standaardcurriculum. Eén studie richtte zich op het schematiseren bij kleuters, en twee studies hadden breuken als leerstofdomein.

SCHEMATISEREN BIJ KLEUTERS

Poland en Van Oers (2007; zie ook Poland, 2007) ontwikkelden een experimenteel programma voor kleuters waarin schematiserende activiteiten in betekenisvolle spelsituaties werden aangeleerd. De prestaties van kinderen die het experimentele programma gericht op schematiseren gevolgd hadden in groep 2 (kleuters) verschilden nauwelijks van die van kinderen in de controlegroep op een rekentoets halverwege de interventie ($ES = -0,05$) en aan het eind van de interventie ($ES = +0,02$). Acht maanden na de interventie presteerden kinderen uit de experimentele groep hoger op de rekentoets ($ES = +0,57$), een matig tot groot effect. Aan het einde van groep 3 – twaalf maanden na de interventie – was dit verschil afgenomen tot een klein effect, nog steeds in het voordeel van de experimentele groep ($ES = +0,18$). Kinderen uit het experimentele programma lieten zowel gedurende de interventie als erna betere schematiseringsactiviteiten zien dan kinderen uit de controlegroep (niet in overzichtstabel F).

Samenvattend: Het experimentele programma voor kleuters waarin schematiserende activiteiten in betekenisvolle spelsituaties werden aangeleerd had een verwaarloosbaar klein tot matig positief effect op de rekenprestaties in groep 3.

BREUKENONDERWIJS

Streefland (1991) ontwikkelde in de jaren tachtig een nieuw realistisch programma voor instructie in en het leren van breuken, met veel aandacht voor de eigen producties en constructies van leerlingen. Uit de vergelijkende analyse tussen de leerlingen uit de experimentele klas en de leerlingen uit de controleklassen (zowel realistische als mechanistische methoden) bleek dat de prestaties op de ontwikkelde breukentoets het hoogst waren voor de leerlingen die het experimentele realistische programma gevolgd hadden, gevolgd door controleleerlingen met een realistische methode. De controleleerlingen met een traditionele methode presteerden het slechtst. Op het onderdeel *rekenen* van de Cito-eindtoets was ditzelfde patroon zichtbaar, net als op de eindtoetsopgaven die over verhoudingen, breuken en procenten gingen. Bij gebrek

aan informatie was statistische toetsing van de verschillen en het berekenen van effectgroottes niet mogelijk. Ook kon geen rekening gehouden worden met eventuele aanvangsverschillen tussen de groepen leerlingen.

Samenvattend: Het in de jaren tachtig ontwikkelde programma voor instructie in en het leren van breuken met veel aandacht voor de eigen producties en constructies van leerlingen lijkt een positief effect op de prestaties te hebben gehad.

Keijzer en Terwel (2003; zie ook Keijzer, 2003) ontwikkelden ook een nieuw programma voor het breukenonderwijs. Dit programma was op twee punten vernieuwend ten opzichte van het curriculum van de realistische rekenmethode *Wereld in Getallen* (WiG). Ten eerste het model voor breuken: in WiG was dit vooral de opgedeelde cirkel, in het experimentele programma was dit de getallenlijn. Ten tweede de manier van instructie of kennisconstructie: in het controleprogramma werkten de leerlingen voornamelijk individueel, terwijl in het experimentele programma discussies met de gehele klas over de betekenis van het geleerde centraal stond. Op gestandaardiseerde LVS-toetsen was er een verwaarloosbaar verschil in prestaties tussen de twee groepen op het onderdeel *getallen en bewerkingen* ($ES = -0,01$), maar presteerden de leerlingen die het experimentele programma hadden gevolgd beter dan de leerlingen in de controleconditie op het onderdeel *meten en meetkunde* ($ES = +0,35$): een klein tot matig effect. Op breukenopgaven afgenomen in een interviewsituatie met gestandaardiseerde hulp scoorden de leerlingen met het experimentele programma gemiddeld matig hoger (ongecorrigeerde $ES = +0,52$).

Samenvattend: Het nieuw ontwikkelde programma voor het breukenonderwijs met de getallenlijn als model voor breuken en discussies met de gehele klas over de betekenis van het geleerde hadden geen tot een matig positief effect op de prestaties.

CONCLUSIES EXPERIMENTELE PROGRAMMA'S

In de verschillende studies worden verwaarloosbaar kleine tot grote positieve effecten gevonden van experimentele programma's vergeleken met het volgen van het standaardcurriculum, zowel in speciaal als in regulier basisonderwijs. Deze experimentele programma's hadden elk verschillende kenmerken van realistische rekendidactiek als onderwerp van onderzoek: de ontwikkeling van oplossingsstrategieën d.m.v. zelf-instructie, discussie en reflectie; het schematiseren in betekenisvolle situaties; het belang van eigen producties en constructies; de getallenlijn als model, en discussies met de gehele klas over de betekenis van het geleerde.

4.4 Curriculumstudies

Het doel van curriculumstudies is het vaststellen van verschillen in rekenprestaties bij verschillende in gebruik zijnde curricula, meestal geoperationaliseerd als rekenmethode (rekenboek). De onderzoekers hebben geen controle over het curriculum waarvoor de school kiest en de uitvoering van het curriculum. Gewoonlijk vindt één meting plaats en kan dus niet gecorrigeerd worden voor aanvangsverschillen. Deze onderzoeken vallen onder de observationele designs. Een nadeel van dit soort onderzoeken is dat selectie-effecten niet uitgesloten kunnen worden: factoren die bepalen welke rekenmethode wordt gekozen kunnen heel goed samenhangen met prestaties. In de bespreking van de studies is een verdeling gemaakt in:

1. domeinspecifieke studies, gericht op een bepaald, afgebakend onderdeel van rekenen (paragraaf 4.4.1);
2. grootschalige studies, waarin algemene rekenvaardigheden worden vastgesteld (paragraaf 4.4.2).

4.4.1 Domeinspecifieke curriculumstudies

Type onderzoek:	Domeinspecifieke curriculumstudies.
Onderzochte relatie:	Optellen en aftrekken bij een structureel of realistisch curriculum bij leerlingen in het speciaal onderwijs en probleemoplossen en schriftelijk delen bij leerlingen in regulier basisonderwijs.
Overzicht:	Tabel G in bijlage 3

Er zijn drie studies gevonden waarbij de prestaties van leerlingen met verschillende curricula vergeleken werden op een specifiek domein van rekenen: één in het speciaal basisonderwijs en twee in het reguliere basisonderwijs.

Speciaal basisonderwijs

Van Luit (1994) vergeleek de prestaties op *optellen* en *afrekken* van leerlingen uit het speciaal basisonderwijs (MLK en LOM) met een structureel of realistisch curriculum.

Op de natoets⁷ presteerden MLK-leerlingen op *optellen* en *afrekken* zonder tientaloverschrijding met het realistisch curriculum wat slechter dan met het structurele curriculum ($ES = -0,22$; een klein effect), op *optellen* en *afrekken* met tientaloverschrijding was er nauwelijks verschil ($ES = +0,04$). Bij de LOM-leerlingen waren deze verschillen matig tot groot in het nadeel van het realistische curriculum, met respectievelijk $ES = -0,62$ en $ES = -1,00$. Op realistische contextopgaven verschilden de

⁷ Hoewel er ook een voortoets is afgenomen, is hiervoor niet statistisch gecorrigeerd, omdat bij de voortoets de leerlingen al zes maanden instructie in optellen en aftrekken met tientaloverschrijding hadden gehad volgens een structureel of realistisch curriculum.

prestaties van LOM-leerlingen met een realistisch of structureel curriculum nauwelijks van elkaar ($ES = -0,08$).

Samenvattend: De prestaties van leerlingen uit het speciaal basisonderwijs (MLK en LOM) met een realistisch curriculum waren gelijk aan tot slechter dan die van leerlingen met een structuralistisch curriculum.

Regulier basisonderwijs

Nelissen (1987) vergeleek het probleemoplossen van leerlingen van proefscholen die van groep 1 tot groep 8 les hadden gehad met het *Rekenwerk*-curriculum, met leerlingen van controlescholen. De methode *Rekenwerk* was nieuw ontwikkeld met als doel het rekenonderwijs te verbeteren en met name probleemoplossen en reflectie te stimuleren. De controlescholen gebruikten voornamelijk traditionele methoden (*Nieuw Rekenen*, *Naar Zelfstandig Rekenen* en *Operatoir Rekenen*). Probleemoplossen werd getoetst met acht open wiskundige vraagstukken waarbij de leerlingen werden geconfronteerd met een socio-cognitief conflict. In het oplossingsproces volgde een 'systematische samenspraak', waarbij de proefleider volgens vaste regels hulp bood. Leerlingen uit de proefscholen scoorden hierop hoger dan leerlingen uit de controlescholen ($ES = +0,82$; een groot effect), maar er kon niet gecorrigeerd worden voor eventuele aanvangsverschillen tussen de condities. Leerlingen op de proefscholen planden vaker hun handelen vooraf en reflecteerden op hoger niveau dan leerlingen in de controlegroepen (niet in overzichtstabel G).

Samenvattend: De realistische methode *Rekenwerk* had een groot positief effect op het probleemoplossen van leerlingen, vergeleken met controlescholen.

In het onderzoek van Van Putten, Van den Brom-Snijders en Beishuizen (2005) naar schriftelijk delen werden groep 6 leerlingen die les kregen met de methode *Rekenen & Wiskunde* (R&W) vergeleken met leerlingen die les krijgen met *Wereld in Getallen* (WiG). Beide methoden kunnen gekarakteriseerd worden als realistisch van aard, waarbij WiG een meer (voor)gestructureerde aanpak bij het leren delen voorstaat. Halverwege groep 6 bleken leerlingen met R&W lager te presteren dan leerlingen met WiG ($ES = -0,43$). Aan het eind van groep 6 was dit patroon omgekeerd en presteerden leerlingen met R&W beter dan leerlingen met WiG ($ES = +0,35$). Tevens bleek het strategiegebruik (niet in overzichtstabel G) zich positief te ontwikkelen op de aspecten 'mate van schematisering' (R&W meer toename dan WiG) en 'mate van gebruik van getalrelaties' (R&W en WiG dezelfde toename; WiG wel in totaal hogere score).

De resultaten van de Nederlandse leerlingen in groep 6 zijn ook vergeleken met Engelse leerlingen van vergelijkbare leeftijd (Anghileri e.a., 2002; niet in tabel G). Het Engelse onderwijs op het gebied van delen wordt gekenmerkt door een vrij abrupte overgang van de informele strategieën naar het traditionele algoritme (de staartdeling). In de Nederlandse realistische methoden R&W en WiG wordt progressief

voortgebouwd op de informele strategieën richting meer gestructureerde en efficiëntere aanpakken. Aan het einde van groep 6 presteerden de Nederlandse leerlingen beter dan de Engelse: een aanwijzing dat het progressief voortbouwen op de informele hoofdrekenstrategieën van de leerlingen beter heeft gewerkt.

Samenvattend: Leerlingen met de methode *Rekenen & Wiskunde* hadden halverwege groep 6 een lager niveau dan leerlingen met *Wereld in Getallen*, maar zetten deze achterstand om in een voorsprong aan het einde van groep 6.

CONCLUSIES DOMEINSPECIFIEKE CURRICULUMSTUDIES

In het speciaal basisonderwijs presteerden leerlingen met een realistisch rekencurriculum slechter dan leerlingen met een structureel curriculum. In het reguliere basisonderwijs had de methode *Rekenwerk* een positief effect op het hogere-orde doel probleemoplossen. De prestaties bij het onderdeel (schriftelijk) delen met de realistische methoden R&W en WiG lieten in de loop van groep 6 een wisselend patroon zien, maar beide methoden bleken effectiever dan een Engelse traditionele aanpak.

4.4.2 Grootschalige curriculumstudies naar reken- en wiskundevaardigheden

Type onderzoek:	Grootschalige curriculumstudies.
Onderzochte relatie:	Vergelijking van de prestaties bij verschillende rekenmethoden in resp. MORE-onderzoek, onderzoek van Harskamp, en PPON en onderwijskenmerken die samenhangen met rekenprestaties in het Inspectierapport basisvaardigheden

In de jaren tachtig zijn twee grootschalige curriculumstudies uitgevoerd in het reguliere basisonderwijs: het MORE-project van het Freudenthal Instituut (1993) en dat van Harskamp (1988). In beide onderzoeken werden de destijds moderne methoden met traditionele methoden vergeleken. Het MORE-onderzoek volgde leerlingen van groep 3 tot en met groep 5; Harskamp onderzocht uitsluitend groep 8 (zie tabel H). Periodieke peilingsonderzoeken (PPONs) worden sinds 1987 uitgevoerd door Cito, zowel halverwege het basisonderwijs (groep 5) als aan het einde van het basisonderwijs (groep 8). In deze PPONs, waarin het niveau van rekenvaardigheden wordt vastgesteld, worden meestal de verschillen tussen methoden gerapporteerd. In het Inspectierapport (2008) ten slotte werd gezocht naar schoolfactoren (geen rekenmethoden) die samenhangen met leerlingprestaties.

MORE-onderzoek

In het MORE-onderzoek (Gravemeijer e.a., 1993) werden de prestaties van leerlingen met twee rekenmethoden vergeleken: de (destijds) moderne methode *Wereld in Getallen – versie 1* (WiG-1) met het traditionele *Naar Zelfstandig Rekenen* (NZR). De

resultaten zijn gecorrigeerd voor beginniveau in groep 3, leerlinggewicht en intelligentie. Op algemene rekenvaardigheden presteerden leerlingen met WiG-1 in groep 3 ongeveer gelijk ($ES = -0,02$), in groep 4 iets slechter ($ES = -0,10$; verwaarloosbaar tot klein verschil) en in groep 5 nog iets slechter ($ES = -0,32$; klein tot matig verschil) dan leerlingen met NZR. Op automatiseren waren de verschillen matig tot groot in het nadeel van WiG-1 met in groep 4 $ES = -0,60$ en in groep 5 $ES = -0,58$. Verder bleek uit implementatieonderzoek (niet in overzichtstabel H) dat bij de NZR-leraren het uitgevoerde onderwijs redelijk mechanistisch was, terwijl het bij de WiG-leraren slechts beperkt realistisch was, dus dat het uitgevoerde onderwijs bij NZR-leraren meer overeenkwam met de gebruikte methode dan bij WiG-leraren.

Samenvattend: De prestaties van leerlingen met de realistische methode *Wereld in Getallen – versie 1* (WiG-1) lagen verwaarloosbaar tot beduidend lager dan met de traditionele methode *Naar Zelfstandig Rekenen* (NZR), in groep 3, 4 en 5.

Harskamp

Harskamp (1988) onderzocht de rekenprestaties van leerlingen met in totaal acht verschillende methoden, waarvan hij er drie classificeerde als modern (geen WiG) en vijf als traditioneel (waaronder NZR). Gecorrigeerd voor intelligentie werden verwaarloosbare tot kleine verschillen in prestaties gevonden op de Cito-eindtoets ($ES = +0,09$) en op ontwikkelde RION-toetsen ($ES = +0,06$), in het voordeel van de moderne methoden. Verder bleken de algemene implementatiefactoren ‘aantal lessen per week’ en ‘behandelde leerstof’ (percentage leerlingen waarmee de basisstof uit de methode is behandeld) beide een positief effect op de leerprestaties te hebben (niet in overzichtstabel H). Een vakdidactisch implementatiekenmerk met een positief effect op rekenprestaties was de variatie in de leerstof, het percentage tijd besteed aan hoofdrekenen en redactiesommen. Een negatief effect had de leerstofdifferentiatie, de mate waarin leerlingen in een klas verschillende leerstof krijgen aangeboden (niet in overzichtstabel H).

Samenvattend: De prestaties van leerlingen met drie ‘moderne’ methoden lagen licht hoger dan die van leerlingen met vijf traditionele methoden.

Periodieke peiling van het onderwijsniveau

HALVERWEGE HET BASISONDERWIJS (GROEP 5)

In de meest recente peiling *halverwege* het basisonderwijs van 2003/2004 werden de rekenprestaties van leerlingen met zeven verschillende realistische methoden vergeleken (Kraemer, Janssen, Van der Schoot en Hemker, 2005). Leerlingen met *Talrijk* en *Rekenrijk* presteerden het beste, leerlingen met *WiG-1* en *Rekenen & Wiskunde* presteerden het slechtste. Het verschil tussen hoogste en de laagste prestaties naar methode is matig tot groot ($ES = +0,64$). Omdat alle methoden realistisch van aard waren is geen

vergelijking tussen realistische en traditionele methoden mogelijk. Wel kan gezegd worden dat de verschuiving in het marktaandeel van de methoden tussen 1997 en 2003 in zijn totaliteit een klein positief effect op de rekenprestaties op alle onderdelen gezamenlijk heeft gehad ($ES = +0,18$). Deze verschuiving kan niet gekarakteriseerd worden in termen van realistisch versus traditioneel, omdat in 1997 nog slechts 1,5 procent van de methoden traditioneel was.

Samenvattend: De rekenprestaties van leerlingen in groep 5 met zeven verschillende realistische methoden verschilden van elkaar. De marktaandeelverschuiving binnen het realistische spectrum van methoden tussen 1997 en 2003 heeft in zijn totaliteit een klein positief effect op de rekenprestaties in groep 5 gehad.

AAN HET EINDE VAN HET BASISONDERWIJS (GROEP 8)

De meest recente peiling aan het einde van het basisonderwijs uit 2004 maakt geen uitsplitsing naar specifieke methode, omdat 80 procent van de scholen een nieuwe rekenmethode gebruikte naar aanleiding van de invoering van de euro (Janssen, Van der Schoot en Hemker, 2005). Slechts de summatieve effecten van methodeverschuivingen werden gerapporteerd. Voor de periode 1997 tot 2004 was dit effect positief maar erg klein ($ES = +0,12$), voor de periode 1992 tot 2004 was het effect ook positief maar klein ($ES = +0,18$): in zijn totaliteit heeft de verschuiving in marktaandeel van rekenmethoden tussen 1992 en 2004 dus een kleine positieve bijdrage aan de rekenprestaties geleverd. Dit wil zeggen dat de verschuiving in dit tijdvak (van 37 naar 100 procent realistische methoden) de globale rekenprestaties licht gunstig heeft beïnvloed.

De derde peiling aan het einde van het basisonderwijs uit 1997 maakte nog wel een uitsplitsing naar specifieke methoden (Janssen, Van der Schoot, Hemker en Verhelst, 1999). Leerlingen met *Wereld in Getallen – versie 2* presteerden het beste, leerlingen met *Niveau Cursus Rekenen* en *NZR* (de twee meest typische traditionele methoden) presteerden het slechtste. Het verschil tussen de hoogste en de laagste prestaties naar methode is matig ($ES = +0,53$). De verschillen binnen een methodotype (realistisch of traditioneel) zijn groter dan tussen de methodotypen.

Samenvattend: In 1997 verschilden de rekenprestaties van leerlingen in groep 8 met acht verschillende methoden van elkaar, waarbij de verschillen binnen een methodotype (realistisch of traditioneel) groter waren dan tussen de methodotypen. Verder heeft de marktaandeelverschuiving van methoden (zowel van traditionele en hybride methoden naar realistische methoden, als binnen het realistisch spectrum zelf) tussen 1992 en 2004 in zijn totaliteit een klein positief effect op de rekenprestaties in groep 8 gehad.

Inspectierapport basisvaardigheden

De Inspectie van het Onderwijs (2008) zocht naar schoolfactoren die samenhangen met rekenprestaties. Er werd geen vergelijking gemaakt tussen verschillende curricula, dus strikt genomen past dit onderzoek niet binnen de curriculumstudies. De resultaten zijn echter recent, op een grote steekproef verkregen, en relevant. Rekenzwakke scholen bleven op negen factoren achter ten opzichte van rekensterke scholen: (1) de jaarlijkse systematische evaluatie van de resultaten van leerlingen; (2) de borging van de kwaliteit van het leren en onderwijzen; (3) rekenzwakke scholen bieden de leerstof vaker voor veel leerlingen niet aan tot en met het niveau van groep 8; (4) het realiseren van een taakgerichte werksfeer; (5) duidelijk uitleggen; (6) onderwijzen van strategieën voor leren en denken; (7) actieve betrokkenheid van de leerlingen; (8) planmatige uitvoering van zorg; en (9) nagaan van effecten van zorg. Ook blijken rekenzwakke scholen in de ontwikkeling van de tussentijdse opbrengsten achter te blijven. Scholen met achterblijvende tussenresultaten besteden minder tijd aan rekenen van groep 1 tot en met 8.

Samenvattend: Op rekenzwakke scholen was het onderwijsproces (de kwaliteitszorg, het leerstofaanbod, het didactisch handelen en de leerlingenzorg) van mindere kwaliteit.

CONCLUSIES: GROOTSCHALIGE CURRICULUMSTUDIES

Uit één van de curriculumstudies uit de jaren tachtig bleek dat de prestaties van leerlingen met de methode *Wereld in Getallen – versie 1* (WiG-1) lager lagen dan met de traditionele methode *Naar Zelfstandig Rekenen* (NZR), in groep 3, 4 en 5. In de andere jaren tachtig studie lagen de prestaties van leerlingen met drie ‘moderne’ methoden juist licht hoger dan die van leerlingen met vijf traditionele methoden.

Een eerste observatie uit de PPONs is dat er geen eenduidige resultaten over de vergelijking in rekenprestaties tussen traditionele en realistische methoden zijn gevonden. Er blijken ook substantiële verschillen in prestaties binnen beide methodetypen (traditioneel en realistisch). Verder werden in groep 5 met de nieuwere realistische methoden licht betere resultaten behaald dan met de oudere realistische methoden. In groep 8 heeft de verschuiving van traditionele, hybride en (oudere) realistische methoden naar (nieuwere) realistische methoden ook een licht positief effect op de rekenprestaties gehad. Ten slotte bleek uit het Inspectierapport dat op rekenzwakke scholen het onderwijsproces van mindere kwaliteit was.

4.5 Samenvatting, conclusies en aanbevelingen

De commissie heeft een overzichtsstudie verricht naar de relatie tussen rekendidactiek en rekenvaardigheid. De studie concentreerde zich op empirisch onderzoek dat de laatste twintig jaar in Nederland is verricht. Het omvatte echter ook een beknopte inventarisatie van buitenlands onderzoek. Op basis van deze studie doet de commissie de volgende observaties.

De resultaten geven geen eenduidig beeld. Hoewel er de afgelopen decennia in Nederland veel onderzoek is gedaan naar het wiskundig denken en leren van kinderen en naar de verbetering van het rekenonderwijs, is er opmerkelijk weinig literatuur waaruit gefundeerde en eenduidige conclusies zijn te trekken omtrent de relatie tussen rekendidactiek en rekenvaardigheid.

Interventiestudies maken een vergelijking tussen de effectiviteit van twee of meer interventies. De interventiestudies die de commissie in haar analyse heeft betrokken laten geen duidelijke conclusies toe, omdat zij beperkt zijn qua rekendomein, qua steekproefgrootte, qua omvang of duur van de interventie of qua breedte van de gemeenten uitkomstvariabelen.

Curriculumstudies zijn ruimer van opzet en maken een grootschalige vergelijking tussen verschillende curricula. Zij zijn echter weer beperkt qua controle op de feitelijke implementatie van de onderzochte didactieken en qua controle van de verstorende variabelen. Het feit dat het laboratorium hier de realiteit is brengt grote beperkingen met zich mee. Ook de resultaten van deze studies wijzen niet eenduidig in een bepaalde richting.

Veel in Nederland verricht onderzoek betreft kleinschalige *design experiments* ter ontwikkeling en verfijning van nieuwe leergangen. Het Freudenthal Instituut, het belangrijkste Nederlandse instituut op het gebied van reken- en wiskundeonderwijs, heeft zich vooral op dit type onderzoek gericht. Het werk heeft innovatieve deel-leergangen opgeleverd, die ook internationaal veel erkenning en navolging hebben gekregen. Dergelijk onderzoek beschrijft weliswaar het tot stand komen van deze deelleergangen en hoe leerlingen erop reageren, maar maakt meestal geen vergelijking met alternatieve leergangen. De commissie heeft als er geen vergelijking tussen programma's wordt gemaakt, dit werk buiten beschouwing gelaten.

De door de commissie in beschouwing genomen studies maken zelden een directe vergelijking tussen een traditionele en realistische didactiek maar betreffen wel aspecten die men als traditioneel of realistisch kan aanmerken. De gevonden effecten zijn over het algemeen klein. Alles bij elkaar genomen lijkt er sprake te zijn van een zeer klein effect in het voordeel van aspecten van het realistisch rekenonderwijs (met name wat betreft de hogere-orde vaardigheden), echter te klein om er een harde conclusie aan te verbinden. Bovendien zijn er enkele studies, waaronder het MORE-project, die eerder in het voordeel van de traditionele aanpak pleiten.

Samenvattend concludeert de commissie dat het door haar bestudeerde materiaal geen algemene wetenschappelijk gefundeerde uitspraken over de relatie tussen

rekendidactiek en rekenvaardigheid rechtvaardigt. In het bijzonder zijn er geen harde onderzoeksresultaten die de claims van enige partij in de discussie over traditioneel versus realistisch rekenen op overtuigende wijze ondersteunen.

Deze wat teleurstellende hoofdconclusie betekent niet dat er in het geheel niets te zeggen is over de relatie tussen rekendidactiek en rekenvaardigheid. Over de volgende punten bestaat wetenschappelijk gezien wel degelijk duidelijkheid.

Het is opmerkelijk dat er op het vlak van leerlingprestaties *binnen* een bepaalde rekendidactiek vaak grotere verschillen bestaan dan *tussen* verschillende rekendidactieken. De algemene didactische principes spelen kennelijk een kleinere rol dan de specifieke uitwerking van de didactiek en de interactie tussen leraar en leerling. De rol van de leraar blijkt groter dan die van de gebruikte didactiek. We komen hier in hoofdstuk 6 op terug.

Meer onderwijstijd voor rekenen leidt tot betere resultaten. Dit blijkt uit de overwegend positieve effecten van remediële en extra oefenprogramma's, waarin onderwijs wordt toegevoegd aan het reguliere curriculum. Ook geven, bij gelijkblijvende onderwijstijd, experimentele programma's die in kleine groepjes leerlingen worden geïmplementeerd positieve effecten ten opzichte van de reguliere onderwijspraktijk.

In het Nederlandse onderzoek dat de commissie heeft bestudeerd gaat de meeste aandacht uit naar *zwakke rekenaars*. Dit is begrijpelijk omdat deze leerlingen in het onderwijsproces de meeste zorgen baren. Rekenzwakke kinderen lijken minder gebaat bij een vrije vorm van instructie en hebben meer behoefte aan een sturende rol van de leraar in hun leerproces. Naar de relatie tussen rekendidactiek en rekenvaardigheid bij gemiddelde en sterke rekenaars is minder onderzoek gedaan. Ook is er weinig bekend over de verschillen in het effect van rekendidactieken tussen jongens en meisjes.

Ook *internationaal* gezien is de empirische evidentie over de relatie tussen rekendidactiek en rekenvaardigheid beperkt. Het *National Mathematics Advisory Panel* in de VS, wier opdracht sterk leek op die van deze commissie, bestudeerde 16.000 onderzoeksrapporten en concludeerde dat slechts een zeer klein deel daarvan verantwoorde uitspraken toeliet over het effect van instructievariabelen op wiskundige leeruitkomsten. In lijn daarmee concluderen Slavin en Lake dat, gegeven de beschikbare empirische evidentie, de didactiek en het curriculum van minder gewicht zijn dan de interactie tussen leraar en leerling. Hun bevinding dat computerondersteund onderwijs kleine maar consistent positieve effecten heeft verdient aandacht.

CONCLUSIE 4.1

De commissie constateert dat het bereik van het vergelijkend wetenschappelijk onderzoek op het gebied van de effectiviteit van het rekenonderwijs smal is, zowel nationaal als internationaal. De meeste interventiestudies zijn beperkt qua rekendomein, qua steekproefgrootte, qua duur van de interventie of qua breedte van de gemeten uitkomstvariabelen. De meestal ruimere curriculumstudies gaan mank aan een gebrek aan controle over de feitelijke implementatie van de onderzochte didactieken en over allerlei versturende variabelen. De ontwikkelingsexperimenten zijn gericht op de ontwikkeling van leergangen en niet op vergelijkend onderzoek.

CONCLUSIE 4.2

Het door de commissie bestudeerde materiaal leidt niet tot een eenduidig beeld en rechtvaardigt geen algemene wetenschappelijk gefundeerde uitspraken over de relatie tussen rekendidactiek en rekenvaardigheid. Het biedt, in het bijzonder, geen overtuigende empirische ondersteuning voor de claims van enige partij in de discussie over traditioneel versus realistisch rekenen.

CONCLUSIE 4.3

Binnen een bepaalde rekendidactiek bestaan er vaak grotere verschillen in de leerlingprestaties dan *tussen* verschillende rekendidactieken. De specifieke uitwerking van de didactiek en de interactie tussen leraar en leerling spelen kennelijk een grotere rol dan de algemene vakdidactische principes.

CONCLUSIE 4.4

Meer onderwijstijd en aandacht voor rekenen leidt tot betere resultaten.

CONCLUSIE 4.5

Rekenzwakke kinderen lijken minder gebaat bij een vrije vorm van instructie en meer behoefte te hebben aan een sturende rol van de leraar.

AANBEVELING 4.1

Het Ministerie van OCW dient maatregelen te nemen om het wetenschappelijk onderzoek op het gebied van het rekenonderwijs in omvang en variatie te doen toenemen. De commissie denkt daarbij aan:

- a. een reflectie op de kerndoelen van het rekenonderwijs in het licht van maatschappelijke, technologische en wetenschappelijke ontwikkelingen;
- b. ontwikkelingsexperimenten, die nieuwe leergangen en -materialen uitwerken en evalueren;
- c. methodologisch verantwoord vergelijkend onderzoek naar het effect van verschillende rekendidactieken en van verschillende realisaties van een didactiek op de leerprocessen en leeruitkomsten;
- d. nadere analyses van gegevens van grootschalige nationale en internationale peilingen van het rekenonderwijs teneinde tot gefundeerde uitspraken te komen over de relatie tussen rekendidactiek en leerlingprestatie;
- e. studies naar de rol en het effect van informatietechnologie in het rekenonderwijs, in het bijzonder van computerondersteund onderwijs.

AANBEVELING 4.2

Bij de inrichting van universitaire opleidingen voor leraren primair onderwijs dienen de universiteiten het palet van het onderzoek naar het rekenonderwijs te vergroten.

AANBEVELING 4.3

NWO dient in samenwerking met de NVORWO multidisciplinaire samenwerking tussen wiskundigen, vakdidactici, pedagogen en psychologen bij het uitvoeren van onderzoek naar het rekenonderwijs te stimuleren.

5. ONTWIKKELING IN HET AANBOD VAN DE REKENMETHODEN

De commissie was gevraagd aan te geven hoe ruimte kan worden geschapen voor leraren en ouders om te kiezen tussen rekendidaktieken. Dit korte hoofdstuk beschrijft de ontwikkeling in de beschikbaarheid van rekenmethoden (lesboeken), die tot de gewenste ruimte lijkt te leiden. Het was echter niet de taak van de commissie de rekenmethoden die op de markt komen inhoudelijk te vergelijken.

5.1 Een markt in beweging

De uitgevers van rekenmethoden brengen ongeveer iedere acht jaar herziene uitgaven op de markt. Het marktaandeel van realistische rekenmethoden is toegenomen van 15 procent in de nationale peiling einde basisonderwijs van 1987 tot 100 procent in die van 2004; zie tabel 5.1. Sinds de invoering van de euro in 2002 presenteren alle beschikbare rekenboeken zich als realistisch. Overigens vertonen deze methoden grote onderlinge verschillen, bijvoorbeeld wat betreft de hoeveelheid kale sommen en de wijze waarop oplossingsstrategieën en cijferalgoritmen worden geoefend.

Tabel 5.1 Marktaandeel van rekenmethoden in de peilingen einde basisonderwijs (Jansen e.a., 1999, 2005)

PPON eind basisonderwijs	traditioneel	operator rekenen (hybride)	realistisch	overig
1987	49%	28%	13%	10%
1992	29%	31%	37%	3%
1997	8%	17%	74%	1%
2004	0%	0%	100%	0%

Rond 2010 kunnen we weer nieuwe uitgaven van de rekenmethoden tegemoet zien. De ervaringen in de klas maar ook de discussie over rekendidactieken stimuleren een aanpassing van de rekenboeken. Er wordt gestreefd naar meer rust in de presentatie. Een snelle opeenvolging van verschillende rekenactiviteiten kan leerlingen in verwarring brengen; langer stilstaan bij één onderwerp geeft hen de kans zich in de nieuwe stof te verdiepen en zich die eigen te maken. Verder geven de lesboeken meer aandacht aan oefenen en het onderhouden van basale vaardigheden.

Daarnaast wordt er gewerkt aan nieuwe rekenmethoden die minder contexten en één standaardalgoritme per bewerking bieden. Deze methoden zullen ook allerlei elementen van het hedendaagse rekenonderwijs bevatten, want de leerlingen zullen aan de kerndoelen moeten voldoen en de referentieniveaus van de doorlopende leerlijnen moeten bereiken. Vooral de hogere-orde-doelen vragen om een goede invulling.

Deze ontwikkelingen zullen leiden tot meer diversiteit in het aanbod van rekenboeken. De commissie juicht dit toe. De keuzeruimte in de rekenboeken heeft echter een keerzijde. Een methode moet passen bij de leraren en leerlingen van een school, maar het is voor leraren nu vrijwel ondoenlijk om een gedegen vakinhoudelijke en vakdidactische analyse van rekenboeken te maken. Daarom dienen rekenmethoden aan een diepgaande vakdidactische analyse te worden onderworpen. Ook kunnen rekencoördinatoren hier een belangrijke adviesfunctie vervullen.

Alle partijen zijn van mening dat leerlingen meer moeten oefenen met rekenen. De uitgevers passen in dit opzicht hun rekenmethoden aan. Tevens heeft staatssecretaris mevrouw Dijkema, in navolging van het rapport van de Commissie Meijerink, zOEfi, de Nationale Oefenimpuls ingesteld voor onderhoud van basale rekenvaardigheden.

5.2 Conclusie en aanbevelingen

CONCLUSIE 5.1

De markt voor rekenmethoden is sterk in beweging. De toenemende diversiteit in het aanbod geeft scholen meer ruimte om te kiezen. Dit is een gunstige ontwikkeling.

AANBEVELING 5.1

Het Ministerie van OCW dient, in samenwerking met partijen in het veld, ervoor te zorgen dat rekenmethoden aan een diepgaande en objectieve vakinhoudelijke en vakdidactische analyse worden onderworpen, opdat scholen een verantwoorde keuze kunnen maken en ook ouders over gefundeerde informatie beschikken. Dit definieert een werkerrein voor de nieuwe generatie onderzoekers naar reken- en wiskundeonderwijs.

AANBEVELING 5.2

Bij de toenemende diversiteit in het aanbod dient het Ministerie van OCW het principe te blijven handhaven dat de overheid gaat over het 'wat' en niet over het 'hoe'. De Inspectie van het Onderwijs dient zich te blijven beperken tot de controle *of* de doelen worden bereikt en niet te sturen op *hoe* ze worden bereikt.

AANBEVELING 5.3

Cito dient zijn toetsen zodanig in te richten dat er een balans is tussen opgaven met en zonder context, tussen inzicht en vaardigheid. In zowel het basis- als het vervolgonderwijs is het niet gewenst dat een bepaalde vaardigheid is gekoppeld aan één context.

6. ROL, OPLEIDING EN NASCHOLING VAN DE LERAAR

6.1 Rol van de leraar

Vrijwel alle geïnterviewde personen noemden de cruciale rol van de leraar in het leerproces van de leerling. Ook verscheidene van de onderzoeken samengevat in hoofdstuk 4 wijzen in deze richting; zie bijvoorbeeld Slavin e.a. (2008). Voor een nadere wetenschappelijke onderbouwing van deze stelling heeft de commissie gebruik gemaakt van een literatuuronderzoek door dr. Bieke De Fraine van de Katholieke Universiteit Leuven.

Er blijkt weinig consensus te zijn over de criteria om leraren te beoordelen. Gaat het erom dat een leraar een grote leerwinst weet te behalen of dat hij zijn leerlingen optimaal weet te motiveren? Veel onderzoek valt in de categorie lerareffectiviteitsonderzoek. Daarbij worden leraren beoordeeld op basis van de resultaten die ze met hun leerlingen weten te bereiken. Die resultaten kunnen uiteenlopen van kennis en vaardigheden tot motivatie en schools welbevinden. Uit deze onderzoeken komen een aantal interessante bevindingen naar voren.

De leraar maakt verschil. De meeste onderzoekers stellen vast dat er wel degelijk effectiviteitsverschillen bestaan tussen leraren. Nye, Konstantopoulos en Hedges (2004) bijvoorbeeld ordenen leraren volgens de mate van leerwinst die ze bereiken met hun leerlingen. Een leraar die op percentiel 25 scoort wordt beschouwd als minder effectief, terwijl een leraar op percentiel 75 gezien wordt als effectief. Nye e.a. (2004) stellen vast dat het verschil tussen deze beide leraren relevant is: het gaat om een derde van een standaarddeviatie voor lezen en bijna een halve standaarddeviatie voor wiskunde per schooljaar. Er bestaat ruime consensus over het bestaan van leraareffecten.

Effectgrootte. Minder eenduidig zijn de meningen over de grootte van de effecten. Dit heeft te maken met een aantal aspecten die de grootte van effecten beïnvloeden zoals de keuze van het effectiviteitscriterium, de aard van de leerlingenpopulatie en het al dan niet in rekening brengen van schooleffecten. Het onderzochte criterium heeft een grote invloed op de schatting van de grootte van het effect. Er wordt bijvoorbeeld meestal vastgesteld dat effecten groter zijn voor wiskunde dan voor taal (Nye e.a., 2004), maar sommige onderzoeken vinden het tegenovergestelde effect (Palardy e.a., 2008).

Het is bovendien te eenvoudig om te spreken over ‘de effectieve leraar’. Binnen het onderwijs van eenzelfde leraar kunnen er effectiviteitsverschillen optreden afhankelijk van het effectiviteitscriterium, het soort leerlingen en de context. Kortom: er kan sprake zijn van differentiële leraareffecten. In elke geval bevestigt wetenschappelijk onderzoek de cruciale rol van de leraar bij het reken- en wiskundeonderwijs.

De leraar is de spil in het onderwijsleerproces. Er worden hoge eisen aan hem gesteld, vooral bij onderwijs dat ruimte laat aan de inbreng van en de interactie tussen leerlingen, zoals realistisch rekenen. Een verhoging van de kwaliteit van het Nederlandse rekenonderwijs vraagt om investeringen in de opleiding en nascholing van leraren in zowel de eigen rekenvaardigheid als de kennis en beheersing van verschillende didactische inzichten.

Sinds de jaren negentig is onder invloed van het *onderwijs-op-maat*-beleid het onderwijs op de basisschool steeds meer geïndividualiseerd. Zelfstandig werken, alleen of in kleine groepjes, is een veel gebruikte werkvorm, die de rol van de leraar verkleint. In TIMSS-2007 scoort Nederland wat betreft klassenorganisatie het laagst op het *uitleggen van antwoorden* en het hoogst op *zelfstandig werken aan opgaven*. Engeland heeft daarentegen met de *National Number Strategy* juist gekozen voor meer klassikaal onderwijs.

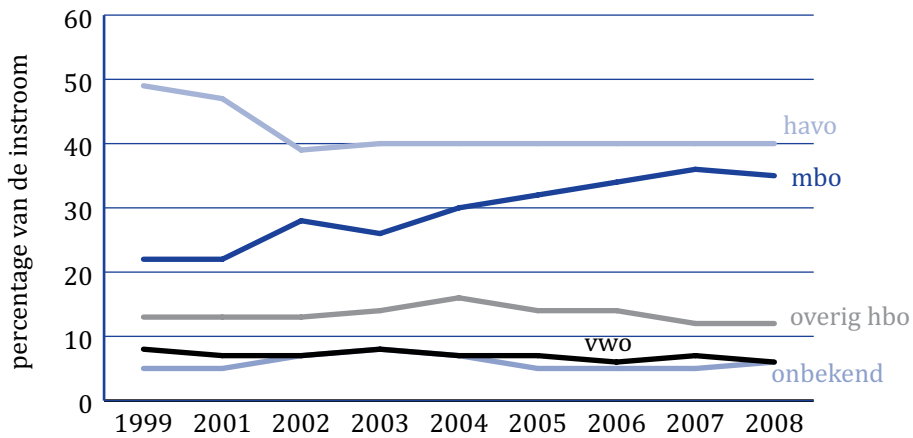
Om rekenprestaties te verhogen lijkt het van belang dat de leraar meer en vaker inhoudelijk betrokken is bij het rekenwerk van de leerlingen. Dit kan in de vorm van instructie, oefening en interactie in de gehele klas of in groepen. Onderwijs waarin elke leerling zelfstandig zijn eigen individuele traject doorloopt biedt minder kansen voor inhoudelijke begeleiding, wat de rekenprestaties nadelig kan beïnvloeden. Dit betekent natuurlijk niet dat er in goed rekenonderwijs geen ruimte zou zijn voor weloverwogen en gedoseerde vormen van differentiatie en individualisering. Goede scholing en nascholing van leraren op het gebied van rekenonderwijs is van belang, maar de opleiding op de pabo lijkt tekort te schieten en de deelname aan nascholing op het gebied van het rekenonderwijs is beperkt. Het percentage leraren dat de afgelopen jaren nascholing in rekenen heeft gevolgd is sterk afgenomen; in TIMSS-2007 is Nederland op dit punt hekkensluiter.

6.2 Opleiding: de situatie op de pabo

Onze gesprekspartners uitten zonder uitzondering hun zorg over de situatie op de pabo. De volgende punten werden naar voren gebracht.

1. Het niveau van de vooropleiding van de instromende studenten neemt af. Cijfers die de commissie heeft verzameld laten zien dat het aandeel mbo-abituriënten in de instroom van de pabo's is gegroeid naar 34 procent ten koste van het aantal vwo-en havo-abituriënten (zie figuur 6.1). Uit onderzoek van het Sectorbestuur Onderwijsarbeidsmarkt (SBO) blijkt dat zowel de medewerkers van de pabo's als de studenten zelf van mening zijn dat studenten met een mbo-vooropleiding een aanzienlijke achterstand hebben op vakinhoudelijk terrein. Dit uit zich vooral bij vakken zoals taal, rekenen en wiskunde en wereldoriëntatie (Van Bergen, 2006). Dit is uitermate zorgelijk omdat een goede docent vakinhoudelijk boven de stof moet staan. Om deze populatie op het gewenste vakinhoudelijke niveau te krijgen is meer tijd en inspanning vereist.
2. Het aantal contacturen voor rekenen op de pabo's kent een grote variatie en een laag gemiddelde. Gemiddeld besteden de opleidingen in de majorfase 54 contact(klok)uren aan rekenen en wiskunde. Dit varieert echter per opleiding van 5 tot 120 contacturen (Keijzer, 2009). De eigen rekenvaardigheid van de studenten krijgt vrijwel uitsluitend aandacht in het eerste studiejaar. Na de verplichte reken-toets wordt de krappe rekentijd grotendeels besteed aan kennis van en vaardigheid in rekendidactiek (Keijzer en Van Os, 2002). Studenten krijgen weinig gelegenheid een basale rekenvaardigheid uit te bouwen tot een professionele rekenvaardigheid. Dit laatste houdt onder andere in dat studenten vraagstukken op verscheidene manieren en niveaus kunnen oplossen, de juistheid van oplossingsmanieren kunnen beargumenteren, de strategieën die leerlingen kiezen kunnen volgen, voltooien en beoordelen, en contexten bij een gegeven bewerking kunnen bedenken en omgekeerd (Kool 2009).
3. De stage neemt in de opleiding een belangrijke plaats in, maar vakdocenten zijn niet betrokken bij het observeren en beoordelen van de rekenlessen van hun studenten.
4. *Assessments* hebben vaak de plaats van toetsen ingenomen. Ook waar dit niet het geval is wordt er veelal integraal getoetst, zodat resultaten voor andere vakken mindere rekenresultaten kunnen compenseren.
5. Er is geen landelijke afstemming wat betreft de te stellen eisen. Onlangs heeft de HBO-raad een kennisbasis laten ontwikkelen. Een volgende stap kan een landelijk eindexamen zijn.

Figuur 6.1 Instroom pabo's naar vooropleiding (STAMOS, jaarboekonderwijsmarkt.nl)



6.3 Nascholing en begeleiding

De rekenvaardigheid en de vakdidactische kwaliteiten van pabostudenten zijn van evident belang voor de kwaliteit van toekomstige rekenonderwijs. Voor de leraren die al werkzaam zijn in het basisonderwijs is echter vooral nascholing van belang. Uit TIMSS-2007 blijkt dat leraren van groep 6 zich meer dan gemiddeld voelen toegerust om les te geven op het gebied van *getallen, geometrische vormen en meten* en *gegevensweergave*, onderdelen van rekenen en wiskunde. Van alle deelnemende landen aan het TIMSS-onderzoek ervaren Nederlandse leraren de minste knelpunten in het omgaan met verschillen tussen leerlingen of probleemleerlingen tijdens de rekenles. In vergelijking met 2003 is het aantal leraren dat nascholing heeft gevolgd met betrekking tot rekenonderwijs gehalveerd. Het beeld ontstaat hierdoor dat leraren in groep 6 geen problemen ervaren en dus ook geen behoefte aan nascholing hebben. Betrouwbare cijfers over nascholing door leraren heeft de commissie niet kunnen vinden.

Om toch meer zicht hierop te krijgen heeft de SLO-nationaal expertisecentrum leerplanontwikkeling een kleine enquête uitgevoerd onder 26 rekenadviseurs van 24 schoolbegeleidingsdiensten. Deze rekenadviseurs zijn lid van het landelijk netwerk voor rekenspecialisten. De centrale vraag bij deze enquête was: "Wat is het beeld van onderwijsbegeleiders van onderwijsadviesdiensten en pedagogische centra die specialist zijn op het gebied van rekenbegeleiding in het basisonderwijs, over het aanbod van nascholing en begeleiding en over de vraag naar nascholing en begeleiding binnen hun eigen werkveld en in Nederland in het algemeen?"

De resultaten zijn gerapporteerd in het achtergronddocument *Nascholing en begeleiding voor rekenen in de basisschool*. We geven de belangrijkste bevindingen hieronder kort weer:

- Onderwijsadviseurs zijn het er unaniem over eens dat leraren regelmatig nascholing moeten volgen om de kwaliteit van rekenonderwijs te waarborgen. Men vindt ook unaniem dat begeleiding op de werkvoet, al dan niet in combinatie met nascholing, noodzakelijk is om tot onderwijsverbetering te komen.
- In de laatste vijf jaar heeft volgens 35 procent van de adviseurs minder dan 10 procent van hun scholen nascholing en begeleiding voor rekenen gevolgd. Eveneens geeft 35 procent aan dat van hun scholen slechts tussen 10 en 25 procent van de leraren nascholing heeft gevolgd. Dit is lager dan voor taal en voor andere onderwerpen.
- De vraag waardoor er in verhouding weinig vraag naar nascholing en begeleiding is, is moeilijker te beantwoorden. Volgens de adviseurs ervaren leraren niet meer knelpunten bij taal dan bij rekenen (35 procent is het hiermee eens, 27 procent niet), en men vindt de leraren deskundiger in taal dan in rekenen. Waarom dan niet meer nascholing? Bijna de helft geeft aan dat leraren zich deskundig genoeg voelen op het gebied van rekenen (tegen 27 procent die het hier niet mee eens is). Dat betekent dat men er minder behoefte aan heeft. Ook kosten en tijd kunnen een rol spelen.
- In sommige gevallen speelt geld een rol bij de vraag om nascholing. Volgens de adviseurs (80 procent) vinden scholen vooral de begeleiding op de werkvloer te duur. Ook het moment van nascholing speelt mee: als het onder schooltijd zou zijn, zou er volgens driekwart van de adviseurs meer nascholing gevolgd worden.
- De belangstelling voor nascholing en begeleiding neemt het laatste jaar wel toe, door de belangstelling van de overheid voor rekenen en wiskunde, de rekenverbetertrajecten en de komst van nieuwe rekenmethodes die de vraag naar invoeringsprogramma's doen toenemen. Dit is een kans voor nascholing en begeleiding, zegt ruim 80 procent.
- Hoewel landelijke nascholing aangepast moet worden aan de situatie en vraag van de school (61 procent is het hiermee eens, 19 procent niet) vindt de meerderheid, bijna 70 procent, dat er meer landelijke nascholing ontwikkeld moet worden.
- Bijna alle adviseurs geven aan dat vooral de directie bepaalt wie er nascholing en begeleiding voor rekenen volgt. In veel gevallen gaat dat gepaard met advies van de intern begeleider. Zelden bepaalt de leraar dit zelf. Ook diensten worden zelden bij de beslissing betrokken.

Hoewel de enquête gezien de beperkte omvang en de specifieke doelgroep niet representatief is, bevestigt hij wel het beeld dat ook uit TIMSS ontstaat. De vraag naar nascholing en begeleiding op het gebied van rekenen en wiskunde is laag. Wel lijkt hierin verandering te komen door de landelijke aandacht voor de kwaliteit van rekenonderwijs en leerlingprestaties. Rekenverbetertrajecten en de komst van nieuwe rekenmethodes zullen de vraag doen toenemen. Dit biedt kansen om – gestimuleerd door de huidige belangstelling van overheid en beleidsmakers (en inspectie) én het op de markt komen van nieuwe of vernieuwde rekenmethoden – via scholing en

begeleiding de kwaliteit van het rekenonderwijs van leraren te vergroten. Nascholing die gepaard gaat met begeleiding op de werkvloer van leraren en teams lijkt hiervoor de beste mogelijkheden te bieden.

De inzet van rekencoördinatoren kan helpen om meer kennis over rekenvaardigheid en rekendidactiek binnen de school te brengen. Het is niet de bedoeling dat deze coördinatoren in de plaats komen van de leraren.

6.4 Samenvatting, conclusies en aanbevelingen

Wij merken nogmaals op dat onze opmerkingen over opleiding, nascholing en rol van de leraar primair zijn gebaseerd op de interviews die wij hebben gehouden. Men is over het algemeen van mening dat een stevigere rol van de leraar tot verbetering van het rekenonderwijs kan leiden. Een grondige analyse van alle factoren die hier een rol spelen behoorde niet tot onze opdracht en paste niet binnen ons tijdbestek. We hebben enig feitenmateriaal verzameld over deelaspecten van deze problematiek, die tot enkele conclusies en aanbevelingen leiden. Daarnaast is er alle aanleiding om vragen te stellen, die tot nader onderzoek kunnen leiden.

CONCLUSIE 6.1

De kwaliteit van de leraar heeft direct effect op de leerprestaties.

CONCLUSIE 6.2

Het niveau van de pabo-instroom neemt af. Het aantal contacturen voor rekenen op de pabo's kent een grote variatie maar is in het algemeen laag.

CONCLUSIE 6.3

Nascholing en begeleiding voor rekenen en wiskunde zijn noodzakelijk, maar de vraag ernaar is laag.

AANBEVELING 6.1

Scholen dienen zicht te bezinnen op het veelvuldig gebruik van niet-begeleid zelfstandig werken tijdens de rekenles. De commissie plaatst vraagtekens bij de effectiviteit van deze werkvorm en pleit voor een grotere inhoudelijke rol voor de leraar. Sturing door en interactie met de leraar en instructie, oefening en nabespreking zijn noodzakelijk.

AANBEVELING 6.2

Het Ministerie van OCW dient de situatie op de pabo's op korte termijn aan een grondig onderzoek te onderwerpen, met als doel het niveau van rekenvaardigheid en rekendidactiek van de opleiding te verhogen. Daarbij dienen de volgende vragen te worden geadresseerd:

- a. Is het mogelijk hogere eisen te stellen aan de instroom?
- b. Is er op de pabo een goede balans tussen aandacht voor onderwijskunde en pedagogiek enerzijds en vakinhoud en vakdidactiek anderzijds?
- c. Is er in de beschikbare onderwijstijd een goed evenwicht tussen aandacht voor de eigen rekenvaardigheid en voor rekendidactiek?
- d. Kan de tijd voor rekenen worden uitgebreid, met als doel te komen tot professionele gecijferdheid en kennis van verschillende didactische inzichten?
- e. Biedt de kennisbasis die de HBO-raad heeft laten ontwikkelen kans op een landelijke normering?
- f. Is het zinvol de pabo-opleiding te splitsen in specialisaties voor onderbouw en bovenbouw?
- g. Moet de ontwikkeling van universitaire pabo's worden gestimuleerd? Welke rol kunnen zij spelen bij het verbeteren van het rekenonderwijs?

AANBEVELING 6.3

Het Ministerie van OCW dient nascholing in het rekenonderwijs krachtig te stimuleren. Nascholing dient gepaard te gaan met begeleiding op de werkvloer. Nascholing en begeleiding zouden onder schooltijd moeten plaatsvinden. De bereidheid om aan nascholing deel te nemen en het effect ervan zullen hierdoor toenemen.

AANBEVELING 6.4

Scholen dienen te overwegen om rekencoördinatoren aan te stellen, eventueel in onderlinge samenwerking. Hun taak zou niet het vervangen maar het ondersteunen van de leraar zijn. Zij zouden zich kunnen organiseren in een landelijk netwerk. Is het in dit verband zinvol een masteropleiding rekencoördinatie te ontwikkelen, bijvoorbeeld op de universitaire pabo of een post-hbo-opleiding?

LITERATUUR

- Anghileri, J., M. Beishuizen en C.M. van Putten (2002). 'From informal strategies to structured procedures: Mind the gap!' *Educational Studies in Mathematics*, 49, 149-170
- Bergen, C.T.A. van, P.J. Krooneman (2006). *Mbo'ers op de pabo, knelpunten en oplossingen bij de overgang van mbo naar pabo en leraarschap*. Sectorbestuur Onderwijsarbeidsmarkt (SBO), Den Haag
- Blöte, A. W., E. van der Burg, E. en A.S. Klein (2001). 'Students' flexibility in solving two-digit addition and subtraction problems: instruction effects.' *Journal of Educational Psychology*, 93, 627-638.
- Braams, T. en M. Milikowski (2008). *De gelukkige rekenklas*. Boom
- Buijs, K. (2008). *Leren vermenigvuldigen met meercijferige getallen*. Proefschrift Universiteit Utrecht, Freudenthal Instituut.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Commissie Parlementair Onderzoek Onderwijsvernieuwingen ('Dijsselbloem') (2008). *Tijd voor onderwijs; eindrapport* (31 007, nr. 6, 13 februari 2008). 's-Gravenhage: Sdu Uitgevers.
- Craats, J. van de (2007). 'Waarom Daan en Sanne niet kunnen rekenen.' *Nieuw Archief voor Wetenschap*, 8, 132-136.
- Dijk, I.M.A.W. van, B. van Oers, J. Terwel en P. van Eeden (2003). 'Strategic learning in primary mathematics education: Effects of an experimental program in modelling.' *Educational Research and Evaluation*, 9, 161-187.
- Expertgroep doorlopende leerlijnen (Commissie Meijerink) (2008). *Over de drempels met taal en rekenen; Hoofdrapport. Over de drempels met rekenen; Consolideren, onderhouden, gebruiken en verdiepen*.
- Fennema, E. en M.L. Franke (1992). 'Teachers' knowledge and its impact.' In D. Grouws (ed.), *Handbook of research on mathematics teaching and learning*. New York: MacMillan, 147-164.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht:Reidel.
- Freudenthal, H. (1991). *Revisiting Mathematics Education*. China Lectures. Dordrecht: Kluwer.
- Gravemeijer, K., M. van den Heuvel-Panhuizen, G. van Donselaar, N. Ruesink, L. Streefland, W. Vermeulen, E. te Woerd en D. van der Ploeg (1993). *Methoden in het reken-wiskundeonderwijs, een rijke context voor vergelijkend onderzoek*. Utrecht: Freudenthal Instituut.
- Hambleton, R.K. en M.J. Pitoniak (2006). 'Setting performance standards.' In R.L. Brennan (ed.), *Educational Measurement*. Fourth Edition. Westport, CT: American Council on Education en Praeger Publishers.

- Harskamp, E.G., C.J.M. Suhre en T.W.F.P. Willemsen (1993). *Remediële rekenprogramma's voor het basisonderwijs beproefd*. Groningen: GION.
- Harskamp, E.G. (1988). *Rekenmethoden op de proef gesteld*. Proefschrift Rijksuniversiteit Groningen.
- Harskamp, E.G. en C.J.M. Suhre (1995). *Hoofdrekenen in het speciaal onderwijs*. Groningen: GION
- Heuvel-Panhuizen, M. van den (2009). *Hoe rekent Nederland? Oratie* (10 maart 2009), Universiteit Utrecht, Freudenthal Instituut. Utrecht: All Print.
- Hickendorff, M., W.J. Heiser, C.M. van Putten en N. Verhelst (2009). 'Solution strategies and achievement in Dutch written arithmetic: Latent variable modeling of change.' *Psychometrika*, 74, 331-350.
- Hiebert, J. en D.A. Grouws (2007). 'The effects of classroom mathematics teaching on students' learning' In F. K. Lester (ed.), *Second handbook of research on mathematics teaching and learning*, 371-404. Greenwich, CT: information Age Publishing.
- Hill, H.C. et al. (2007). 'Assessing teachers' mathematical knowledge. What knowledge matters and what evidence counts?' In: F. K. Lester (ed.), *Second handbook of research on mathematics teaching and learning*, 111-154. Greenwich, CT: information Age Publishing.
- Inspectie van het Onderwijs (2007). *De staat van het onderwijs; onderwijsverslag 2005/2006*. Den Haag: OBT bv.
- Inspectie van het Onderwijs (2008). *Basisvaardigheden rekenen-wiskunde. Een onderzoek naar het niveau van rekenen-wiskunde in het basisonderwijs en naar verschillen tussen scholen met lage, gemiddelde en goede reken-wiskunderesultaten*. Den Haag: Inspectie van het Onderwijs.
- Janssen, J., F. van der Schoot en B. Hemker (2005). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 4*. PPON-reeks nummer 32. Cito: Arnhem.
- Janssen, J., F. van der Schoot, B. Hemker en N. Verhelst (1999). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 3*. Cito: Arnhem.
- Keijzer, R. (2003). *Teaching formal mathematics in primary education*. (diss.) Utrecht: Freudenthal Instituut.
- Keijzer, R. (2009). *Stand van zaken bij rekenen-wiskunde en didactiek op de lerarenopleiding basisonderwijs*, in druk.
- Keijzer, R. en J. van Os (2002). 'Rekenen-wiskunde & didactiek anno 2002.' *Tijdschrift voor nascholing en onderzoek van het reken-wiskundeonderwijs*, jaargang 20, nummer 3.
- Keijzer, R. en J. Terwel (2003). 'Learning for mathematical insight: a longitudinal comparative study on modelling.' *Learning and Instruction*, 13, 285-304.
- Klein, A.S. (1998). *Flexibilization of mental arithmetic strategies on a different knowledge base: the empty number line in a realistic versus gradual program design*. Proefschrift Rijksuniversiteit Leiden.
- Klein, A.S., M. Beishuizen en A. Treffers (1998). 'The empty number line in Dutch second grades: Realistic versus gradual program design.' *Journal for Research in Mathematics Education*, 29, 443-464.
- Kool, M. (2009). 'De professionele wiskundekennis van de leraar basisonderwijs.' In R. Keijzer e.a. *Over de muurtjes heen kijken*. 54-64
- Kraemer, J.-M., J. Janssen, F. van der Schoot en B. Hemker (2005). *Balans van het reken-wiskundeonderwijs halverwege de basisschool 4*. Cito: Arnhem.
- Kroesbergen, E.H. en J.E.H. van Luit (2003). 'Mathematics interventions for children with special educational needs. A meta-analysis.' *Remedial and Special Education*, 24, 97-114.
- Kroesbergen, E.H. en J.E.H. van Luit (2002). 'Teaching multiplication to low math performers: Guided versus structured instruction.' *Instructional Science*, 30, 361-378.

- Kroesbergen, E.H., J.E.H. van Luit en C.J.M. Maas (2004). 'Effectiveness of explicit and constructivist mathematics instruction for low-achieving students in the Netherlands.' *Elementary School Journal*, 104, 233-251.
- Lipsey, M.W. en D.B. Wilson (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Luit, J.E.H. van (1994). 'The effectiveness of structural and realistic arithmetic curricula in children with special educational needs.' *European Journal of Special Needs Education*, 9, 16-26.
- Luit, J.E.H. van en J.A. Naglieri (1999). 'Effectiveness of the MASTER program for teaching special children multiplication and division.' *Journal of learning disabilities*, 32, 98-107.
- Luit, J.E.H. van en E.A.M. Schopman (2000). 'Improving early numeracy of young children with special educational needs.' *Remedial and Special Education*, 21, 27-40.
- Menne, J. (2001). *Met sprongen vooruit. Een productief oefenprogramma voor zwakke rekenaars in het getalengebied tot 100 – een onderwijsexperiment*. (diss.) Utrecht: Freudenthal Instituut.
- Milo, B.F., A.J.J.M. Ruijsenaars, en G. Seegers (2005). 'Math instruction for students with special educational needs: Effects of guiding versus directing instruction.' *Educational en Child Psychology*, 22, 70-80.
- Mullis, I.V.S., M.O. Martin, J.F. Olson, D.R. Berger, D. Milne & G.M. Stanco (2008). *TIMSS-2007 Encyclopedia. A guide to mathematics and science education around the World*. Part 2, Boston: Boston College, TIMSS & PIRLS International Study Center.
- Nelissen, J.M.C. (1987). *Kinderen leren wiskunde*. Proefschrift Universiteit Utrecht. Gorinchem: De Ruiter.
- Noteboom, A. (2008). *Fundamentele doelen rekenen-wiskunde. Uitwerking van het Fundamenteel niveau 1F voor einde basisonderwijs*, versie 1.1. Enschede: SLO.
- Nye, B., S. Konstantopoulos, en L.V. Hedges (2004). 'How large are teacher effects?' *Educational Evaluation and Policy Analysis*, 26, 237-257.
- Palardy, G.J., en R.W. Rumberger (2008). 'Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning.' *Educational Evaluation and Policy Analysis*, 30, 111-140.
- Poland, M. (2007). *The treasures of schematising. The effects of schematising in early childhood on the learning processes and outcomes in later mathematical understanding*. Proefschrift Vrije Universiteit Amsterdam.
- Poland, M. en B. van Oers (2007). 'Effects of schematising on mathematical development.' *European Early Childhood Education Research Journal*, 15, 269-293.
- Putten, C.M. van en M. Hickendorff (2009). 'Peilstokken voor Plasterk: Evaluatie van de rekenvaardigheid in groep 8.' *Tijdschrift voor Orthopedagogiek*, 48, 184-194.
- Putten, C.M. van, P.A. van den Brom-Snijders en M. Beishuizen (2005). 'Progressive mathematization of long division strategies in Dutch primary schools.' *Journal for Research in Mathematics Education*, 36, 44-73.
- Rijt, B.A.M. van de en J.E.H. van Luit (1998). 'Effectiveness of the additional early mathematics program for teaching children early mathematics.' *Instructional Science*, 26, 337-358.
- Schoot, F. van der (2008). *Onderwijs op peil? Een samenvattend overzicht van 20 jaar PPON*. Cito Arnhem
- Schopman, E.A.M. en Van Luit, J.E.H. (1996). Learning transfer of preparatory arithmetic strategies among young children with a developmental lag. *Journal of Cognitive Education*, 5, 117-131.
- Senk, S. en D. Thompson (eds.) (2003). *Standards-oriented school mathematics curricula: What does research say about student outcomes*. Mahwah, NJ: Erlbaum.

- Slavin, R.E. (2008). 'What works? Issues in synthesizing educational program evaluation.' *Educational Researcher*, 37, 5-14.
- Slavin, R.E. en C. Lake (2008). 'Effective programs in elementary mathematics: A best-evidence syntheses.' *Review of Educational Research*, 78, 427-515.
- Stein, M.K., J. Remillard en M.S. Smith (2007). 'How curriculum influences student learning.' In F. K. Lester (ed.), *Second handbook of research on mathematics teaching and learning*, 319-370. Greenwich, CT: information Age Publishing.
- Streefland, L. (1991). *Fractions in realistic mathematics education. A paradigm of developmental research*. Dordrecht: Kluwer Academic Publishing.
- Terwel, J., B. van Oers, I. van Dijk, I. en P. van Eeden, P. (2009). 'Are representations to be provided or generated in primary mathematics education? Effects on transfer.' *Educational Research and Evaluation*, 15, 25-44.
- Timmermans, R.E. en E.C.D.M. van Lieshout (2003). 'Influence of instruction in mathematics for low performing students on strategy use.' *European Journal of Special Needs Education*, 8, 5-16.
- Timmermans, R.E., E.C.D.M. van Lieshout en L. Verhoeven (2007). 'Gender-related effects of contemporary math instruction for low-performers on problem-solving behavior.' *Learning and Instruction*, 17, 42-54.
- Treffers, A. (1978). *Wiskobas doelgericht*. (diss.) Utrecht: IOWO
- Treffers, A. (1987). *Three dimensions: a Model of Goal and Theory Description in Mathematics Instruction. The Wiskobas Project*. Dordrecht: Kluwer
- Treffers, A. (2007). 'De kwaliteit van het rekenonderwijs; een virtueel vraagesprek.' *Reken-wiskundeonderwijs: onderzoek, ontwikkeling, praktijk*, 26 (4), 11-17.
- US Department of Education (2008). *Foundations for Success. The final report of the National Mathematics Advisory Panel*.
- Verschaffel, L. (2009). "'Over het muurtje kijken": Achtergrond, inhoud en receptie van het Final Report van het "National Mathematics Advisory Panel" in de U.S.' *Panama-Post – Reken-wiskundeonderwijs: Onderzoek, ontwikkeling, praktijk*, 28(1), 3-20
- Verschaffel, L., B. Greer en E. de Corte (2007). 'Whole number concepts and operations.' In F. K. Lester (ed.), *Second handbook of research on mathematics teaching and learning*, 557-628. Greenwich, CT: information Age Publishing.
- Willemsen, T.F.W.P. (1994). *Remediële rekenprogramma's voor de basisschool*. Proefschrift Rijks-universiteit Groningen: GION.
- Wittmann, E.Ch. (2005). 'Realistic Mathematics Education, past and present.' *Nieuw Archief voor Wiskunde*, 5/6 (4), 294-296.

BIJLAGE 1

GEÏNTERVIEWDE DESKUNDIGEN

Bij de totstandkoming van dit advies heeft de commissie dankbaar gebruik gemaakt van de deskundigheid en visie van de volgende personen:

M. van der Burg	Basisschool de Morgenster, Sleeuwijk
Dr. C. Buijs	SLO-nationaal expertisecentrum leerplanontwikkeling, Enschede
J. Bouman	Basisschool de Morgenster, Sleeuwijk
H. van Die	ex-Inspectie van het Onderwijs
Dr. H.A.A. van Eerde	Freudenthal Instituut, Utrecht
Dr. M. van Groenestijn	Hogeschool Utrecht
Dr. E.G. Harskamp	Rijksuniversiteit Groningen
Dr. L.S.J.M. Henkens	Inspectie van het Onderwijs
Prof. dr. M.H.A.M. van den Heuvel-Panhuizen	Freudenthal Instituut, Utrecht
Drs. K. Hoogland	APS
Drs. J. Janssen	Cito, Arnhem
Dr. J.E.H. van Luit	Universiteit Utrecht
Dr. M. Milikowski	de Rekencentrale
J. Muijres	Basisschool de Hoeven, Beuningen
Dr. J.M.C. Nelissen	Freudenthal Instituut, Utrecht
C. Notten	Basisschool de Uilenspiegel, Boekel
J. Oudshoorn	Zalmplaatschool, Hoogvliet
I. Roessingh	Zalmplaatschool, Hoogvliet
A. Ruizeveld de Winter	Zalmplaatschool, Hoogvliet
Dr. E. Simons	Uitgeverij Malmberg
M. Vos-van Driel	Basisschool de Morgenster, Sleeuwijk
Drs. M. van Zanten	Freudenthal Instituut, Utrecht

BIJLAGE 2

BEREKENING

EFFECTGROOTTE

De gehanteerde effectgrootte (ES) is het gestandaardiseerde gemiddelde verschil tussen condities (*standardized mean difference*, SMD), zie bijvoorbeeld Lipsey en Wilson (2001). Het verschil tussen de gemiddelde score in conditie 1 (\overline{X}_1) en de gemiddelde score in conditie 2 (\overline{X}_2) wordt gedeeld door de gepoolde standaarddeviatie s_p :

$$ES_{SMD} = \frac{\overline{X}_1 - \overline{X}_2}{s_p}$$

met s_p berekend als:

$$s_p = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 1}}$$

Om eventuele verschillen tussen de conditie vóór de aanvang van de interventie zo goed mogelijk te verdisconteren, zijn waar mogelijk aangepaste gemiddelden (uit bijvoorbeeld een covariantie-analyse of een multiële regressie-analyse) gebruikt, waarbij gecorrigeerd is voor relevante variabelen als voorttoetscore en achtergrondkenmerken. Als de aangepaste gemiddelden niet beschikbaar waren, is de effectgrootte van het verschil tussen condities op de voortoets afgetrokken van de effectgrootte van het verschil tussen condities op de natoets (Slavin, 2008). Slavin raadt ook aan studies waarbij de verschillen op de voortoets groter zijn dan 0,5 standaarddeviatie uit te sluiten, omdat statistische correctie voor dergelijke aanvangsverschillen niet goed mogelijk is. In hoofdstuk 4 zijn deze studies niet uitgesloten, maar is er voor gekozen om de gecorrigeerde effectgrootte (dus ES op nameting minus ES op voormeting) wel te rapporteren, tussen vierkante haken. Aan dergelijke gecorrigeerde effectgroottes kan slechts beperkte waarde worden gehecht, omdat ze statistisch onbetrouwbaar zijn.

BIJLAGE 3

TABELLEN MET

EFFECTGROOTTES

Tabel A. Interventiestudies waarbij vergelijking wordt gemaakt tussen twee of meer ontwikkelde programma's, bij zwakke rekenaars.

Noot. Als de waarde van een ES tussen [] staat, was het verschil op de voormeting groter dan 0,5 SD. Adequate statistische correctie is dan niet goed mogelijk (Slavin, 2008). De gerapporteerde gecorrigeerde ES is daarom onbetrouwbaar.

studie	domein rekenen	N	groep/lft	BO	interventie
Kroesbergen & Van Luit (2002)	vermenigvuldigen	75 lln; 7 scholen	9 jaar	regulier + speciaal BO; zwakke rekenaars	1a: guided instruction (GI); 1b: structured instruction (SI); 2. controle
Kroesbergen, Van Luit & Maas (2004)	vermenigvuldigen	265 lln; 23 scholen	9,7 jaar	regulier + speciaal BO; zwakke rekenaars	1a: constructivistic instruction (CI) 1b: explicit instruction (EI) 2. controle
Milo, Ruijsenaars & Seegers (2005)	optellen en aftrekken < 100	70 lln; 3 scholen	9,8 jaar	speciaal BO	1. directing instruction - rijg (DI-r) 2. directing instruction - splits (DI-s) 3. guiding instruction (GI)

Vervolg tabel A. op de volgende pagina's

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
4 mnd, 30 sesses van 1/2 uur; GI+SI (1a+1b) uitgevoerd door onderzoekers in kleine groepjes; controle (2) uitgevoerd door reguliere leerkracht in hele klas	voormeting – interventie – nameting; random toewijzing op ll-niveau aan exp (GI/SI) vs. controle; random toewijzing op klasniveau aan GI vs. SI	voormeting automatisering	automatisering: GI vs. controle: n.s. SI > controle GI < SI <i>regulier BO: GI > SI</i> <i>speciaal BO: GI < SI</i>	0,00 [+0,51] [-0,51] [+0,64] [-2,42]
		voormeting vaardigheid	vaardigheid: GI > controle SI > controle GI > SI <i>regulier BO: GI > SI</i> <i>speciaal BO: GI > SI</i>	+0,89 [+0,46] +0,43 [+0,85] +0,32
		voormeting transfer	transfer: GI > controle SI > controle GI > SI <i>regulier BO: GI > SI</i> <i>speciaal BO: GI > SI</i>	+0,96 [+0,44] +0,52 [+1,03] +0,36
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
4 mnd, 30 sessies van 1/2 uur; CI+EI (1a+1b) uitgevoerd door externen in kleine groepjes controle (2) uitgevoerd door reguliere leerkracht in hele klas	voormeting – interventie – nameting random toewijzing op ll-niveau aan exp (CI/EI) vs. controle; random toewijzing op klasniveau aan CI vs. EI	voortoets, eerdere instructie, geslacht, alg. rekenniveau, IQ, schooltype (reg / speciaal)	automatisering: CI > controle EI > controle CI vs. EI: n.s.	+0,35 +0,32 +0,03
		"	vaardigheid: CI > controle EI > controle CI < EI	+0,23 +0,53 -0,30
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
6 mnd, 30 lessen; uitgevoerd door getrainde studenten	voormeting – interventie – nameting; toewijzing groepjes lln. (uniform MLK of uniform LOM) aan instructeur; procedure selectie en toewijzing onduidelijk	voormeting + / - < 100	prestaties optellen / aftrekken < 100: GI < DI-r GI vs. DI-s: n.s. DI-r vs. DI-s: n.s.	-0,73 -0,21 +0,52
		geen correctie	transfer: GI vs. DI-r: n.s. GI vs. DI-s: n.s. DI-r vs. DI-s: n.s.	+0,07 +0,59 +0,52

Vervolg tabel A.

studie	domein rekenen	N	groep/lft	BO	interventie
Timmermans & Van Lieshout (2003)	afrekken < 100	16 lln; 2 scholen	10,5 jaar	speciaal BO; laag-presterend op aftrekopgaven	1. guided instruction (GI); 2. direct instruction (DI)
Timmermans, Van Lieshout & Verhoeven (2007)	afrekken < 100	40 lln; 5 scholen	9,3 jaar	regulier BO; laag-presterend op aftrekopgaven	1. guided instruction (GI); 2. direct instruction (DI)
Van de Rijt & Van Luit (1998)	vroege rekenvaardigheden	136 lln; 20 klassen	kleuters: 4-7 jaar (gem. 5,9 jaar)	regulier BO; zwakke rekenaars	1. AEM-programma – guided instruction (GI) 2. AEM-programma – structured instruction (SI) 3. controle (wel of geen rekenmethode)

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
34 lessen, waarvan 24 specifiek GI vs. DI; alle lessen door dezelfde trainer uitgevoerd	voormeting – interventie – nameting; binnen klassen IIn toegewezen aan GI vs. DI; IIn in DI en GI gematched op voortoets en leeftijd	voortoets tempo	tempotoets optellen (zonder tientaloverschrijding): GI < DI	-0,99
		voortoets tempo	overige tempotoetsen: GI vs. DI n.s.	??
		voortoets prestaties	prestatietoets: GI vs. DI n.s.	??
		geen correctie	transfer (zonder tientaloverschrijding): GI < DI	-0,96
		geen correctie	transfer (met tientaloverschrijding): GI vs. DI n.s.	-0,18
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
34 lessen, waarvan 24 specifiek GI vs. DI binnen schoolklas GI en DI door dezelfde trainer uitgevoerd	voormeting – interventie – nameting toewijzing: matched pairs van IIn. binnen klassen; binnen 'pair' random toewijzing IIn aan GI vs. DI	voortoets prestaties	prestaties: GI vs. DI n.s. <i>meisjes: GI > DI</i> <i>jongens: GI < DI</i>	+0,13 <i>+0,84</i> <i>-0,51</i>
		voortoetsen tempo	tempotoetsen: GI vs. DI n.s. <i>meisjes: GI vs. DI n.s.</i> <i>jongens: GI vs. DI n.s.</i>	+0,05 <i>+0,07</i> <i>+0,03</i>
		geen correctie	strategietoets: GI vs. DI n.s. <i>meisjes: GI > DI</i> <i>jongens: GI vs. DI n.s.</i>	+0,17 <i>+0,60</i> <i>-0,25</i>
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
13 weken, 2 x per week les van 30 minuten instructie in groepjes 4-5 IIn., waarschijnlijk door reguliere leerkracht	voormeting – interventie – nameting matching IIn. op voortoets, leeftijd, geslacht; toewijzing aan condities op leerlingniveau	voortoets vroege rekenvaardigheden	prestaties vroege rekenvaardigheden: AEM-GI > controle AEM-SI > controle AEM-GI vs. AEM-SI: n.s.	+1,06 <i>+1,26</i> <i>+0,20</i>

Tabel B. Interventiestudies waarbij vergelijking wordt gemaakt tussen twee of meer ontwikkelde programma's, bij reguliere leerlingen.

studie	domein	N	groep/lft	BO	interventie
Klein (1998) Klein, Beishuizen & Treffers (1998) Blöte, Van der Burg & Klein (2001)	2-cijferig optellen en aftrekken	275 lln; 10 klassen (9 scholen)	groep 4	regulier BO; alle lln	1. Realistic Program Design (RPD); 2. Gradual Program Design (GPD)
studie	domein	N	groep/lft	BO	interventie
Van Dijk, Van Oers, Terwel & Van Eeden (2003); Terwel, Van Oers, Van Dijk & Van Eeden (2009).	procenten en grafieken	238 lln; 10 klassen (8 scholen)	groep 7; 10-11 jaar	regulier BO; alle lln	1. programma 'co-constructing / designing' (CCD) 2. programma 'providing' (P)

Tabel C. Interventiestudies waarbij het effect van één of meer remediële of extra oefenprogramma's wordt vastgesteld, in het special BO.

Noot. Als de waarde van een ES tussen [] staat, was het verschil op de voormeting groter dan 0,5 SD. Adequate statistische correctie is dan niet goed mogelijk (Slavin, 2008). De gerapporteerde gecorrigeerde ES is daarom onbetrouwbaar.

studie	domein	N	groep/lft	BO	interventie
Harskamp & Suhre (1995)	optellen en aftrekken <100	24 lln; 4 scholen	LOM 10,2 jr; MLK 11,4 jr	speciaal BO; zwakke rekenaars	1. remediële programma; 2. controlegroep

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
1 schooljaar; implementatie door reguliere leerkracht	interventie – tussenmeting – nameting; random toewijzing op klasniveau, via gematchede paren van klassen (op LVS)	geen correctie	tempotoetsen + en –: RPD vs. GPD n.s <i>zwakke rekenaars</i> <i>sterke rekenaars</i>	+0,19 +0,57 +0,47
		geen correctie	strategieën-toets + en –: RPD vs. GPD n.s <i>zwakke rekenaars</i> <i>sterke rekenaars</i>	+0,15 +0,31 +0,02
		geen correctie	toets + en – met kladpapier: RPD vs. GPD n.s <i>zwakke rekenaars</i> <i>sterke rekenaars</i>	+0,10 +0,36 -0,15
		"	LVS-toets E4: RPD vs. GPD n.s.	??
		"	transfer (> 100): RPD vs. GPD n.s	-0,03
		"	retentie: RPD > GPD	+0,20
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
3 weken, 13 lessen van 1 uur; implementatie door reguliere leerkracht	voormeting – interventie – nameting; random toewijzing aan condities op klasniveau	gestandaardiseerde voortoets	prestaties procenten + grafieken: CCD > P	+0,33
		gestandaardiseerde voortoets	prestaties transfer: CCD > P	+0,55

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
13 weken, 2 lessen per week van 45 minuten; rem (1) uitgevoerd door remedie-assistent in tweetallen buiten de klas controle (2) uitgevoerd door reguliere leerkracht in de klas	voormeting – interventie – nameting	voormeting +/- < 100	prestaties +/- < 100 op natoets en retentietoets: rem > controle <i>LOM-IIIn</i> <i>MLK-IIIn</i>	[+3,22] [+3,13] +3,69
		voormeting +/- < 100	prestaties toepassingsopgaven: rem > controle <i>LOM-IIIn</i> <i>MLK-IIIn</i>	?? [+3,58] +3,58

Tabel D. Interventiestudies waarbij het effect van één of meer remedieële of extra oefenprogramma's wordt vastgesteld, in het regulier BO.

Noot. Als de waarde van een ES tussen [] staat, was het verschil op de voormeting groter dan 0,5 SD. Adequate statistische correctie is dan niet goed mogelijk (Slavin, 2008). De gerapporteerde gecorrigeerde ES is daarom onbetrouwbaar.

studie	domein rekenen	N	groep/lft	BO	interventie
Harskamp, Suhre en Willemsen (1993)	optellen en aftrekken < 100; tafels	357 lln (89 scholen) 242 lln (81 scholen)	groep 4 groep 5	regulier BO; zwakke rekenaars	6 remedieële condities: combinatie remedieel rekenpakket met rekenmethode 1 controleconditie (geen rem. pakket)
studie	domein rekenen	N	groep/lft	BO	interventie
Menne (2001)	basale rekenvaardigheden < 100	225 lln; 12 scholen	groep 4	regulier BO; alle lln (school moest voldoende laag-presterende rekenaars hebben)	1. productief oefenprogramma 2. controlegroep (waarschijnlijk geen extra oefensessies)
studie	domein rekenen	N	groep/lft	BO	interventie
Willemsen (1994): studie 1	cijferend aftrekken	40 lln; 5 klassen (8 scholen)	groep 6	regulier; zwakke rekenaars	1. exp. remedieel programma 'mapping' 2. controle remedieel programma 'systematisch oefenen'.
studie	domein rekenen	N	groep/lft	BO	interventie
Willemsen (1994): studie 2	cijferend aftrekken	34 lln; 3 scholen	groep 6	regulier BO; zwakke rekenaars	1. remedieel programma 'mapping' 2. remedieel programma 'koloms-gewijs' 3. controle remedieel programma 'systematisch oefenen'.

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
12 lesblokken, 16 uur in 4-5 maanden	voormeting – interventie –nameting toewijzing: onduidelijk, i.i.g. op schoolniveau	voormeting kale opgaven	prestaties kale opgaven: groep 4: rem > controle groep 5: rem > controle	[+1,18] +0,39
		voormeting toepassingsopgaven	prestaties toepassingsopgaven: gr. 4: rem vs. con n.s. gr. 5: rem vs. con n.s.	+0,17 +0,24
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
heel schooljaar; min. 3 keer per week 15 minuten; exp. oefenprogramma uitgevoerd door reguliere leerkracht (??)	voormeting – interventie –nameting toewijzing: onduidelijk, i.i.g. op schoolniveau	voormeting LVS	LVS-toetsen: oefenen > controle (NB. gemiddelden en SD's geschat) <i>allochtone ll.n.: oef > con</i> <i>autocht ll.n.: oef. vs. con n.s.</i>	+0,44 <i>+0,59</i> <i>+0,41</i>
		voormeting MORE	MORE-toetsen (NB: SD's uit MORE-onderzoek) <i>allochtone ll.n.</i> <i>autochtone ll.n.</i>	+0,30 <i>+0,45</i> <i>+0,18</i>
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
10 weken, 10 lessen van 50 minuten exp (1) uitgevoerd door remedial teacher; controle (2) uitgevoerd door reguliere leerkracht	voormeting – interventie –nameting random toewijzing aan condities op klasniveau	voortoets cijf. aftrekken	prestaties cijf. aftrekken: mapping > controle	+0,32
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
10 weken, 10 lessen van 50 minuten exp (1+2) uitgevoerd door remedial teacher; controle (3) uitgevoerd door reguliere leerkracht	voormeting – interventie –nameting random toewijzing aan condities op leerlingniveau	voortoets cijf. aftrekken + voortoets hoofdrekenen < 100	prestaties cijferend aftrekken: mapping > controle koloms vs. controle: n.s.	+0,75 -0,17
		"	retentie prestaties cijferend aftrekken: mapping > controle koloms vs. controle: n.s.	+0,84 +0,20

Tabel E. Interventiestudies waarbij het effect van een experimenteel programma wordt vergeleken met het zelfgekozen standaard curriculum in het speciaal BO.

studie	domein rekenen	N	groep/lft	BO	interventie
Schopman & Van Luit (1996)	vroege rekenvaardigheden	60 lln ?? kleuterscholen	5-7 jaar	speciaal BO; zwakke rekenaars	1. experimentele groep: exp. programma (guiding of directing*) 2. controlegroep: standaard rekencurriculum
* volgens de auteurs verschilden beide instructietypen in de praktische uitvoering niet van elkaar, daarom zijn de resultaten samengenomen					
studie	domein rekenen	N	groep/lft	BO	interventie
Van Luit & Naglieri (1999)	vermenigvuldigen en delen < 100	84 lln; ?? klassen	LOM-lln: 10,8 jaar MLK-lln: 12,7 jaar	speciaal BO; zwakke rekenaars	1. experimentele groep: MASTER trainings-programma 2. controlegroep: standaard (traditioneel) curriculum
studie	domein rekenen	N	groep/lft	BO	interventie
Van Luit & Schopman (2000)	vroege rekenvaardigheden	124 lln; 9 kleuterscholen	5-7 jaar	speciaal BO; zwakke rekenaars	1. experimentele groep: interventieprogramma 2. controlegroep: standaard rekencurriculum (Pluspunt, Rekenwerk, Remelka)

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
exp. programma: 3 maanden, 13 lessen, 2 x per week sessie 0,5 uur; in groepjes van 4 uitvoering exp (1) onduidelijk door wie; uitvoering controle (2) door reguliere leerkracht (??)	voormeting – interventie – nameting lln. gematched, toewijzing binnen matching onduidelijk, maar op leerlingniveau	voormeting vroege rekenvaardigheden	vroege rekenvaardigheden: exp > controle	+1,07

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
exp. programma: 17 weken, 3 x per week les van 45 minuten; in groepjes van 5-6 lln. uitvoering exp (1) door remedial teacher; uitvoering controle (2) door reguliere leerkracht	voormeting – interventie – nameting random toewijzing aan condities op leerlingniveau, binnen LOM en MLK	voortoets verm. + delen	prestaties vermenigvuldigen + delen: exp > controle <i>LOM-lln</i> <i>MLK-lln</i>	+2,16 +2,50 +3,08

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
exp. programma: 6 maanden, 20 lessen, 2 x per week sessie 0,5 uur; in groepjes van 3; uitvoering exp (1) door getrainde assistent; uitvoering controle (2) door reguliere leerkracht	voormeting – interventie – nameting lln. gematched, toewijzing binnen matching onduidelijk, maar op ll.-niveau	voormeting vroege rekenvaardigheden geen correctie.	vroege rekenvaardigheden: exp > controle transfer: exp. vs controle n.s.	+0,75 +0,22

Tabel F. Interventiestudies waarbij het effect van een experimenteel programma wordt vergeleken met het zelfgekozen standaard curriculum, in het regulier BO.

studie	domein rekenen	N	groep/lft	BO	interventie
Keijzer (2003) Keijzer & Terwel (2003)	breuken	20 lln; 1 school (2 klassen)	groep 6; 9-10 jaar	regulier BO; alle lln	1. experimenteel programma 2. controle-programma: Wereld in Getallen
studie	domein rekenen	N	groep/lft	BO	interventie
Poland (2007) Poland & Van Oers (2007)	algemene rekenvaardigheden	133 lln; 6 scholen	groep 2-3	regulier BO; alle lln.	1. experimenteel programma 'schematiseren' 2. controlegroep: standaard kleutercurriculum
studie	domein rekenen	N	groep/lft	BO	interventie
Streefland (1991)	breuken	215 lln; 10 scholen	groep 8	regulier BO; alle lln	1. experimenteel breukenprogramma 2. controlegroepen: 2a. realistische methoden (Rekenwerk, WiG, Taltaal) 2b. mechanistische methoden (NZR en NCR)

duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
1 schooljaar; exp (1) uitgevoerd door onderzoeker; controle (2) uitgevoerd door reguliere leerkracht	voormeting – interventie – nameting	voortoets LVS: G&B	LVS – Getallen & Bewerkingen: exp vs. controle n.s.	-0,02
	toewijzing op klasniveau, niet random,	voortoets LVS: M&M	LVS – Meten & Meetkunde: exp vs. controle n.s.	+0,35
	post-hoc matching op ll-niveau	geen correctie	breukenvaardigheid interviews: exp > controle	+0,52
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
1 schooljaar (kleuters) exp (1) uitgevoerd door reguliere leerkracht en trainer; controle (2) uitgevoerd door reguliere leerkracht	voormeting – interventie – nameting	voortoets getalbegrip	Cito-Rekenen: halverwege interventie: exp vs. controle n.s.	-0,05
	matching scholen in paren; toewijzing binnen paren onduidelijk	"	Cito-Rekenen: einde interventie: exp vs. controle n.s.	+0,02
		"	Cito-Rekenen: 8 mnd. na interventie: exp > controle	+0,57
		"	Cito-Rekenen: 12 mnd. na interventie: exp vs. controle n.s.	+0,18
duur/implementatie	procedure/design	gecorrigeerd	resultaten	ES
exp. programma: gedurende 3 jaar (groep 6-8) exp (1) uitgevoerd door reguliere leerkracht en onderzoeker; controle (2a en 2b) uitgevoerd door reguliere leerkracht	interventie – nameting	geen correctie	ontwikkelde breuken-toets: exp > real > mech (NB. niet statistisch getoetst)	??
	toewijzing: één pilotschool kreeg exp. programma; overige scholen ter vergelijking geselecteerd		"	Cito-Eindtoets Rekenen: exp > real > mech (NB. niet statistisch getoetst)

Tabel G. Domeinspecifieke curriculumstudies

studie	domein rekenen	N	groep/lft	BO	curriculum
Nelissen (1987)	probleemoplossen	studie 1: 117 lln (7 scholen) studie 2: 27 lln (2 scholen)	groep 8	regulier BO; alle lln	1. Rekenwerk (nieuw ontwikkeld curriculum) 2. controle (traditionele methoden)
studie	domein rekenen	N	groep/lft	BO	curriculum
Van Luit (1994)	optellen en aftrekken	44 lln.; ?? scholen	MLK: 11;5 jaar LOM: 9;5 jaar	speciaal BO (MLK + LOM); zwakke rekenaars	1. realistisch curriculum (MLK: Remelka; LOM: Rekenwerk) 2. structureel curriculum (MLK: SWP Lichtenvoorde; LOM: Naar Zelfstandig Rekenen)
studie	domein rekenen	N	groep/lft	BO	curriculum
Van Putten, Van den Brom-Snijders & Beishuizen (2005)	(schriftelijk) delen	259 lln; 10 scholen	groep 6	regulier BO; alle lln	1. Rekenen & Wiskunde 2. Wereld in Getallen

duur	gecorrigeerd	resultaten	ES
gehele schoolloopbaan	geen correctie	probleemoplossen (open opgaven met socio-cogn. conflict): Rekenwerk > controle	+0,82
duur	gecorrigeerd	resultaten	ES
MLK: 6 maanden; 4 x per week 45 minuten	geen correctie	optellen/afrekken zonder tentaloverschrijding: MLK: real. vs. struct. n.s. LOM: real < struct.	-0,22 -0,62
LOM: 8 maanden, 4 x per week 45 minuten	"	optellen/afrekken met tentaloverschrijding: MLK: real. vs. struct. n.s. LOM: real < struct.	+0,04 -1,00
	"	realistische contextopgaven: MLK: n.v.t. LOM: real. vs. struct. n.s.	n.v.t. -0,08
duur	gecorrigeerd	resultaten	ES
schoolloopbaan tot meetmoment	tempotoets	medio groep 6: R&W < WiG	-0,43
	"	einde groep 6: R&W > WiG	+0,35

Tabel H. Grootschalige curriculumstudies

studie	domein rekenen	N	groep/lft	BO	curriculum
Gravemeijer e.a. (1993): MORE-onderzoek	algemene rekenvaardigheden	430 lln; 18 scholen	longitudinaal: groep 3-5	regulier BO; alle lln	gebruikte rekenmethode 1. WiG – versie 1 2. NZR
Harskamp (1988)	algemene rekenvaardigheden	2579 lln; 120 scholen	groep 8	regulier BO; alle lln	gebruikte rekenmethode 1. modern (3 methoden); 2. traditioneel (5 methoden)
Inspectie v/h Onderwijs (2008)	algemene rekenvaardigheden	4500 scholen	groep 8	regulier BO; alle lln	schoolfactoren, vastgesteld in Periodieke Kwaliteitsonderzoeken
Janssen, Van der Schoot, Hemker & Verhelst (1999) = PPON 1997 – einde BO	algemene rekenvaardigheden	1987: 7890 lln. (321 scholen) 1992: 4335 lln. (241 scholen) 1997: 5314 lln. (253 scholen)	groep 8	regulier BO; alle lln	gebruikte rekenmethode (8 verschillende)
Janssen, Van der Schoot & Hemker (2005) = PPON 2004 – einde BO	algemene rekenvaardigheden	1992: 4335 lln. (241 scholen) 1997: 5314 lln. (253 scholen)	groep 8	regulier; alle lln	gebruikte rekenmethode (NB: 80% nieuwe methoden)
Kraemer, Janssen, Van der Schoot & Hemker (2005) = PPON 2003 – medio BO	algemene rekenvaardigheden	1992: 3350 lln. (164 scholen) 1997: 5972 lln. (130 scholen) 2003: 2032 lln. (77 scholen)	groep 5	regulier; alle lln	gebruikte rekenmethode (7 verschillende)

duur	gecorrigeerd	resultaten	ES
3 schooljaren	beginniveau gr. 3, leerlinggewicht, intelligentie	alg. rekenvaardigheden: groep 3 groep 4 groep 5	-0,02 -0,10 -0,32
	"	automatiseren: groep 4 groep 5	-0,60 -0,58
duur	gecorrigeerd	resultaten	ES
gehele schoolperiode	intelligentie	Cito-Eindtoets: modern vs. trad. n.s.	+0,09
	"	RION-toetsen: modern vs. trad. n.s.	+0,06
duur	gecorrigeerd	resultaten	ES
schoolloopbaan tot meetmoment	niet gecorrigeerd	aspecten waarop rekensterke scholen sign. hoger scoren dan rekenzwakke scholen; ES = verschil (sterk - zwak) jaarlijkse syst. evaluatie kwaliteitsborging leerinhouden voldoende lln taakgerichte werksfeer duidelijke uitleg strategieën leren + denken lln. actief betrokken school voert zorg planmatig uit school gaat effecten zorg na tijd besteed aan rekenen-wiskunde	+0,23 +0,23 +0,39 +0,35 +0,31 +0,23 +0,25 +0,32 +0,22 ??
duur	gecorrigeerd	resultaten	ES
schoolloopbaan tot meetmoment	formatiegewicht leerling, stratum school, geslacht, leertijd en afnamejaar	t.o.v. Naar Zelfstandig Rekenen Wereld in Getallen - 2 Nieuw Rekenen Pluspunt Rekenen & Wiskunde Operator Rekenen - 1 Wereld in Getallen - 1 NiveauCursus Rekenen Naar Zelfstandig Rekenen	+0,53 +0,31 +0,29 +0,25 +0,23 +0,22 +0,02 0,00
duur	gecorrigeerd	resultaten	ES
schoolloopbaan tot meetmoment	formatiegewicht leerling, stratum school, geslacht, leertijd	summatief methodeneffect, vgl. 2004-1997 vgl. 2004-1992 vgl. 1997-1992	+0,12 +0,18 +0,06
duur	gecorrigeerd	resultaten	ES
schoolloopbaan tot meetmoment	herkomst, geslacht, leertijd, afnamejaar	t.o.v. Rekenen & Wiskunde Talrijk Rekenrijk Wereld in Getallen - 2+3 Pluspunt - 1+2 Wereld in Getallen - 1 Rekenen & Wiskunde	+0,64 +0,62 +0,46 +0,44 +0,18 0,00
		summatief methodeneffect: vgl. 2003-1997	+0,18

