



# ‘Over het muurtje kijken’: achtergrond, inhoud en receptie van het Final Report van het ‘National Mathematics Advisory Panel’ in de U.S.<sup>1</sup>

L. Verschaffel<sup>2</sup>  
Katholieke Universiteit Leuven

*Dit artikel bevat een uitvoerige bespreking van de achtergrond, inhoud en receptie van het ‘Final Report’ van het ‘National Mathematics Advisory Panel’ in de Verenigde Staten, waarin gepoogd is een eind te maken aan de ‘math war’ die daar al meerdere jaren in alle hevigheid woedt tussen ‘traditionalisten’ en ‘reformisten’.*

*In deze bespreking wordt vooral ingegaan op de stevige kritiek die er vanuit diverse hoeken - vooral vanuit kringen van ‘reform-based math educators’, wetenschappers die tot andere theoretische strekkingen behoren dan de informatieverwerkingsbenadering, kwalitatieve onderzoekers, en ‘critical math theorists’ - is geuit op (1) de samenstelling van dit panel, (2) een aantal belangrijke inhoudelijke tekorten en eenzijdigheden van het rapport, (3) het radicale methodologische uitgangspunt van het panel, en (4) de naïviteit van het (impliciet) gehanteerde ‘R&D’ model.*

*Hoewel het een grove vergissing zou zijn om het Amerikaanse en het Nederlandse (onderzoek van het) wiskundeonderwijs over één kam te scheren, is er een aantal gelijkenissen tussen de ‘math war’; in de Verenigde Staten en de huidige commotie en discussie in Nederland omtrent de kwaliteit en toekomst van het (elementair) wiskundeonderwijs.*

## 1 Inleiding

In Nederland is sinds enige tijd een brede en hevige maatschappelijke discussie aan de gang over het soort van wiskundeonderwijs dat kinderen en jongeren nodig hebben. Aan de ene kant is er het door het Freudenthal Instituut<sup>3</sup> ontwikkelde en gepropageerde realistische onderwijsmodel, dat uitgaat van ‘wiskunde als menselijke activiteit’:

- waarin veel aandacht wordt besteed aan het begeleiden ‘her-uitvinden’ van wiskundige begrippen en structuren op basis van al aanwezige intuïtieve inzichten en informele strategieën van leerlingen;
- waarin veel belang wordt gehecht aan betekenisvolle contexten, zowel bij de vorming van de beoogde wiskundige begrippen en structuren als bij het leren toepassen ervan;
- waarin het wiskundeleren wordt opgevat als het geleidelijk voortbewegen langs verscheidene niveaus van internalisering, abstractie, en formalisering;
- waarin een centrale plaats wordt ingenomen door reflectie op het eigen handelen, gestimuleerd door interactie met de leerkracht en de medeleerlingen,
- waarin ten slotte zorgvuldig gelet wordt op de onderlinge afstemming van verschillende leerlijnen (Treffers, 1987).

Deze visie op wiskundeonderwijs, die tot ontwikkeling en bloei is gekomen vanuit een kritiek op het toentertijd

als (te) mechanistisch, empiristisch of structuralistisch ervaren wiskundeonderwijs (Treffers, 1987), heeft de afgelopen veertig jaar een grote groep Nederlandse wiskundendidactici en -practici geïnspireerd, en heeft er - volgens de voorstanders van deze visie op wiskundeonderwijs - rechtstreeks toe bijgedragen dat het begrip van wiskunde, de vaardigheid in het toepassen ervan, en het vertrouwen en plezier in wiskunde is toegenomen bij leerlingen. Ook internationaal kan deze visie op ruime erkenning, waardering en navolging rekenen (zie bijvoorbeeld Becker & Selter, 1996; De Corte, Greer & Verschaffel, 1996; Verschaffel, Greer & De Corte, 2007).

Daartegenover staat een groep van tegenstanders van het realistische onderwijsmodel - met als bekendste protagonisten Van de Craats (2007) in Nederland en Feys (2008) in Vlaanderen - die stellen dat Nederlandse kinderen - mede door toedoen van dat realistisch onderwijsmodel - minder goed kunnen rekenen dan voorheen, dat de parate kennis van rekenfeiten en de vlotte beheersing van standaardprocedures zijn verdwenen, dat de door de realisten gepropageerde instructieprincipes het rekenonderwijs onnodig chaotisch maken en wiskundige structurering en abstractie in de weg staan, en dat er dus dringend behoefte is aan een aanzienlijke koerswijziging, weg van dat ‘realistische rekenen’.

Ook hun kritiek is overigens te kaderen binnen een ruimere internationale (tegen)beweging die het vernieuwingsgezinde (in het Engels: *reform-based*) gedachtegoed, waartoe we de realistische visie kunnen rekenen, op z’n zachtst gezegd niet genegen is.<sup>4</sup>

De meningsverschillen over deze kwesties in de Nederlandse media zijn bijtijds zo heftig, lopen zo uiteen en zorgen voor zoveel verwarring onder beleidsmensen en practici, dat er door bepaalde instanties, waaronder de gerenommeerde ‘Koninklijke Nederlandse Academie voor Wetenschappen’ (KNAW), met de idee gespeeld wordt om helderheid in de discussie te scheppen door de instelling van een commissie met als opdracht te komen tot (empirisch) gefundeerde uitspraken over de relatie tussen onderwijsmethode en rekenvaardigheid en tot daaruit voortvloeiende aanbevelingen over welke richting het Nederlandse wiskundeonderwijs het best verder uitgaat.<sup>5</sup> Een moedig en belangrijk initiatief volgens sommigen, een arrogante en ronduit gevaarlijke onderneeming volgens anderen.<sup>6</sup>

Hoe dan ook, als deze commissie er komt, doet zij er allicht goed aan bij de start van haar werkzaamheden eens aandachtig ‘over het muurtje te kijken’ of - beter gezegd - ‘over de grote plas’, alwaar zich de afgelopen jaren een ontwikkeling heeft voltrokken op het vlak van het wiskundeonderwijs die sterke gelijkenissen vertoont met wat er zich thans in Nederland afspeelt en die daar ook geleid heeft tot de oprichting van een (expert)commissie die recent een rapport heeft uitgebracht dat heel uiteenlopende reacties heeft uitgelokt.

## 2 De ‘Math War’ in de Verenigde Staten

De aan de gang zijnde strijd in Nederland tussen voor- en tegenstanders van realistisch wiskundeonderwijs doet sterk denken aan wat er zich de afgelopen jaren in de Verenigde Staten heeft afgespeeld. Vanuit een onvrede over de zwakke kwaliteit van het wiskundeonderwijs aldaar (onder andere gevoed door de relatief zwakke scores van leerlingen uit de Verenigde Staten in diverse internationale vergelijkende studies, zoals ‘Trends in International Mathematics and Science Study’ (TIMSS)<sup>7</sup> en ‘Programme for International Student Assessment’ (PISA)<sup>8</sup> en een compleet verschillende duiding daarvan door de twee kampen die in de Verenigde Staten ‘math wars’ aan het uitvechten zijn - de traditionalisten aan de ene kant en de hervormingsgezinden (*reformers*) aan de andere kant - heeft het ‘U.S. Department of Education’ enkele jaren geleden een ‘National Mathematics Advisory Panel’ (verderop afgekort als: NMAP) opgericht dat tegen 2008 een rapport moest uitbrengen met aanbevelingen ter verbetering van het wiskundeonderwijs niveau K-8 (dat wil zeggen voor leerlingen van pakweg vier tot veertien jaar) gebaseerd op het beschikbare onderzoek.

In deze bijdrage geef ik eerst een korte schets van de context waarin dit panel tot stand is gekomen en van de inhoud en opzet van het eindrapport. Daarna ga ik dieper

in op de (kritische) reacties die het in onderwijskundige en vakdidactische kringen heeft uitgelokt.

### Samenstelling van de commissie

In 2006 gaf niemand minder dan toenmalig president George W. Bush een ‘National Mathematics Advisory Panel’ (NMAP) de opdracht om hem en zijn staatssecretaris van onderwijs te adviseren 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.

(Executive Order 13398, geciteerd in NMAP, 2008, p.7)

De verantwoordelijkheid voor de samenstelling van het panel lag bij de staatssecretaris van onderwijs. Het panel werd voorgezeten door Larry R. Faulkner, een gerenommeerde emeritus professor in de scheikunde. De overige panelleden waren zes psychologen, vier ‘mathematics educators’ (waarvan er een pas later werd toegevoegd), vier wiskundigen, één specialist op het terrein van het speciaal onderwijs, één *middle school mathematics teacher*, één specialist op het terrein van onderwijsbeleid en een leesonderzoeker (U.S. Department of Education, 2008).

### Werkwijze van de commissie

Omdat het panel deze belangrijke en omvangrijke opdracht binnen een beperkte tijd (van 20 maanden!) diende af te werken, werd het grootste deel van het werk verdeeld over een aantal subcommissies en werkgroepen. Er werden subcommissies opgericht rond ‘Standards of Evidence’, ‘Instructional Materials’, en ‘a National Survey of Algebra Teachers’ en werkgroepen voor ‘Conceptual Knowledge and Skills’, ‘Learning Processes’, ‘Teachers and Teacher Education’, ‘Instructional Practices’, en ‘Assessment’. Al deze subcommissies en werkgroepen hebben deelrapporten opgesteld, die uiteindelijk samengevat en geïntegreerd zijn in een *final report*.

Op geregelde tijdstippen deed het panel beroep op externe (binnenlandse en buitenlandse) experts, die op vraag van het panel bepaalde materialen bestudeerden, over bepaalde netelige kwesties advies gaven, en ontwerp teksten van commentaar voorzagen (NMAP, 2008, p.ix). 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’.

### Inhoud van het rapport

Het zou me te ver voeren alle conclusies en aanbevelingen uit het eindrapport van het NMAP panel in detail te bespreken. Geïnteresseerden kunnen het downloaden,<sup>9</sup> maar wie niet zo veel tijd heeft kan zich beperken tot het

doornemen van de zogenoemde 'Executive Summary' aan het begin van het rapport, waarin alle conclusies en aanbevelingen in 45 punten - verdeeld over een tiental rubrieken (namelijk *curricular content, learning processes, teacher and teacher education, instructional practices, instructional materials, assessment and research policy*) kernachtig gebundeld zijn.

Om toch een globale impressie te geven van de inhoud, vorm en stijl van deze 45 punten, treft u een aantal van deze punten hierna aan.

1. A focused, coherent progression of mathematics learning, with an emphasis on proficiency with key topics, should become the norm in elementary and middle school mathematics curricula. Any approach that continually revisits topics year after year without closure is to be avoided. By the term focused, the Panel means that curriculum must include (and engage with adequate depth) the most important topics underlying success in school algebra. By the term coherent, the Panel means that the curriculum is marked by effective, logical progressions from earlier, less sophisticated topics into later, more sophisticated ones. Improvements like those suggested in this report promise immediate positive results with minimal additional cost. By the term proficiency, the Panel means that students should understand key concepts, achieve automaticity as appropriate (e.g., with addition and related subtraction facts), develop flexible, accurate, and automatic execution of the standard algorithms, and use these competencies to solve problems.
4. A major goal for K-8 mathematics education should be proficiency with fractions (including decimals, percents, and negative fractions), for such proficiency is foundational for algebra and, at the present time, seems to be severely underdeveloped. Proficiency with whole numbers is a necessary precursor for the study of fractions, as are aspects of measurement and geometry. These three areas - whole numbers, fractions, and particular aspects of geometry and measurement - are the Critical Foundations of Algebra.
11. Computational proficiency with whole number operations is dependent on sufficient and appropriate practice to develop automatic recall of addition and related subtraction facts, and of multiplication and related division facts. It also requires fluency with the standard algorithms for addition, subtraction, multiplication, and division. Additionally it requires a solid understanding of core concepts, such as the commutative, distributive, and associative properties. Although the learning of concepts and algorithms reinforce one another, each is also dependent on different types of experiences, including practice.
17. Research on the relationship between teachers' mathematical knowledge and students' achievement confirms the importance of teachers' content knowledge. It is self-evident that teachers cannot teach what they do not know. However, because most studies have relied on proxies for teachers' mathematical knowledge (such as teacher certification or courses taken), existing research does not reveal the specific mathematical knowledge and instructional skill needed for effective teaching, es-

pecially at the elementary and middle school level. Direct assessments of teachers' actual mathematical knowledge provide the strongest indication of a relation between teachers' content knowledge and their students' achievement. More precise measures are needed to specify in greater detail the relationship among elementary and middle school teachers' mathematical knowledge, their instructional skill, and students' learning.

22. Explicit instruction with students who have mathematical difficulties has shown consistently positive effects on performance with word problems and computation. Results are consistent for students with learning disabilities, as well as other students who perform in the lowest third of a typical class. By the term explicit instruction, the Panel means that teachers provide clear models for solving a problem type using an array of examples, that students receive extensive practice in use of newly learned strategies and skills, that students are provided with opportunities to think aloud (i.e., talk through the decisions they make and the steps they take), and that students are provided with extensive feedback.
40. As in all fields of education, the large quantity of studies gathered in literature searches on important topics in mathematics education is reduced appreciably once contemporary criteria for rigor and generalizability are applied. Therefore, the Panel recommends that governmental agencies that fund research give priority not only to increasing the supply of research that addresses mathematics education but also to ensuring that such projects meet stringent methodological criteria, with an emphasis on the support of studies that incorporate randomized controlled designs (i.e., designs where students, classrooms, or schools are randomly assigned to conditions and studied under carefully controlled circumstances) or methodologically rigorous quasi-experimental designs. These studies must possess adequate statistical power, which will require substantial funding. Both smaller-scale experiments on the basic science of learning and larger-scale randomized experiments examining effective classroom practices are needed to ensure the coherent growth of research addressing important questions in mathematics education. Basic research on causal mechanisms of learning, as well as randomized trials, are essential, and, depending on their methodologies, both can be rigorous and relevant to educational practice. Basic research, in particular, is necessary to develop explicit predictions and to test hypotheses, which are underemphasized in current research on mathematics education.

Heel globaal samengevat zou men kunnen stellen dat het rapport pleit voor meer aandacht voor bepaalde inhoudelijke topics, bijvoorbeeld gehele en rationale getallen, algebra, voor meer aandacht voor *focus, coherence and closure* bij de behandeling van de betreffende wiskunde-inhoud, voor de geïntegreerde verwerking van rekenfeiten, -concepten en -procedures, voor een gezond en flexibel evenwicht tussen meer docent- en studentgerichte instructiemethoden (met een voorkeur - zeker voor de zwakke rekenaars - voor een meer directieve en sobere

aanpak dan door vele *reformers* wordt voorgestaan), voor een stel stevige vakinhoudelijke en vakdidactische competenties bij (toekomstige) wiskundeleerkrachten, voor meer middelen voor onderzoek (met name voor rigoureuze (quasi)experimenteel onderzoek naar de effectiviteit van bepaalde vormen van wiskundeonderwijs). Al deze principes en aanbevelingen zijn *in principe* gebaseerd op wat in het rapport omschreven wordt als ‘the best available scientific evidence’. 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. (NMAP, 2008a, p.xvi)

De subcommissie rond ‘Standards of Evidence’ ontwikkelde daartoe een serie algemene principes op basis van aspecten, zoals:

... the excellence of the design, the validity and reliability of the measures, the size and diversity of student samples’ and similar considerations of internal (scientific rigor and soundness) and external validity (generalizability to different circumstances and students). (NMAP, 2008, p.82)

Op basis daarvan onderscheidde men vijf algemene (kwaliteits)niveaus van wetenschappelijke evidentie: *strong, moderate, suggestive, inconsistent and weak evidence* (NMAP, 2008, p.83-84). Deze criteria werden verder verfijnd voor verschillende soorten onderzoek, onder andere 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 *random assignment to conditions, low attrition and valid and reliable measures* (NMAP, 2008, p.84). Van middelmatige kwaliteit werden studies beschouwd waarbij het ontbreken van gerandomiseerde toewijzing van subjecten aan interventietypes gecompenseerd werd door ‘matching’-technieken 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 *low quality* bestempeld en ook als dusdanig behandeld (NMAP, 2008, p.84).

Alles bij elkaar zou men kunnen stellen dat volgens het panel slechts experimentele en quasi-experimentele (interventie)studies *scientific evidence* omtrent de waarde van bepaalde methoden of aanpakken van wiskundeonderwijs konden genereren. Opvallend is dat van de meer dan zestien duizend artikelen en andere onderzoeksrapporten die door het panel werden doorgenomen:

Only a small percentage of available research met the standards of evidence and could support conclusions. (NMAP, 2008, p.82).

Sommige werkgroepen kwamen tot de conclusie dat er

voor hun topic nauwelijks relevant onderzoek beschikbaar was dat aan de (hoge) kwaliteitseisen voldeed. Zo stelde de werkgroep rond ‘instructional methods’ 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 van wetenschappelijke evidentie:

A small number of questions have been deemed to have such currency as to require comment from the Panel, even if the scientific evidence was not sufficient to justify research-based findings. In those instances, the Panel has spoken on the basis of collective professional judgment, but it has also endeavored to minimize the number and scope of such comments. (NMAP, 2008, p.12)

## 3 De reacties

### Algemeen

Relatief snel na de verschijning van het eindrapport van het NMAP verschenen de eerste reacties. Een eerste zeer kritische analyse verscheen in de vorm van een themanummer van ‘The Montana Mathematics Enthusiast’ (TMME) (Greer, 2008). Andere kritische reviews werden gepubliceerd in ‘MAA Focus’ van de ‘Mathematical Association of America’ (Ralston, 2008) en in ‘Teachers College Record’ (Good, 2008). Een eerder neutrale bespreking verscheen in ‘Mathematics Educator’ (Moldavan, 2008). De ‘National Council of Teachers of Mathematics’ (2008) onderschreef de aanbevelingen van het rapport grotendeels en deed tegelijk een oproep voor extra financiële middelen voor de implementatie van deze aanbevelingen én voor meer onderzoek op het domein van wiskundendidactiek. Recent verscheen in de ‘Educational Researcher’ (ER), één van de toptijdschriften binnen de onderwijskunde, een heel themanummer vol bijdragen van toonaangevende (Amerikaanse) wetenschappers waarin het NMAP rapport (zeer) kritisch geanalyseerd wordt vanuit verschillende invalshoeken (Kelly, 2008). Het themanummer wordt afgesloten met een kort weerwoord van respectievelijk de ondervoorzitter en voorzitter van het NMAP-panel (Benbow & Faulkner, 2008).

Hierna geef ik een samenvatting van de voornaamste terugkerende kritische kanttekeningen bij respectievelijk de samenstelling van de commissie, de theoretische/inhoudelijke en methodologische uitgangspunten van het panel, en de wijze waarop het panel de relatie tussen ‘research and development’ heeft opgevat, die ik in deze reviews ben tegengekomen. Hoewel ik deze kritieken

afzonderlijk bespreek, dient beklemtoond dat zij, zoals verderop herhaaldelijk zal blijken, onlosmakelijk met elkaar verweven zijn.

### Samenstelling van de commissie

Het spreekt voor zich dat de kwaliteit, geloofwaardigheid en impact van een rapport zoals het *final report* van het NMAP, in belangrijke mate medebepaald wordt door de samenstelling van de commissie. Alleen wanneer een dergelijke commissie evenwichtig is samengesteld uit experts die de verschillende relevante gebieden waarlijk, evenwichtig en met gezag vertegenwoordigen, maakt zij kans op succes (Confrey, Maloney, & Nguyen, 2008). Welnu, zodra de samenstelling van het NMAP bekend was, werd die scherp bekritiseerd, met name vanuit de hoek van de 'mathematics educators' die zich verbaasden over en ergerden aan de opvallend sterke vertegenwoordiging van (een bepaald type van) cognitieve psychologen in de commissie en het geringe aantal specialisten op het terrein van 'mathematics education'. Zo schrijven Confrey et al. (2008) in hun bijdrage aan het themanummer van de ER:

This means that, of 19 members, only 5 (the mathematics educators and the mathematics teacher) regularly had sustained interactions with mathematics instruction at the K-12 level; fewer than half the Panel members had documented academic preparation in mathematics. In the area of research, more of the experience on the Panel was based in psychology (cognitive and developmental) than in any other area. Mathematics education as a field includes experts in sociology, anthropology, and critical theory, but the Panel was deficient in expertise in those areas. (p.631)

Dezelfde kritiek weerklinkt in Greer's (2008) inleiding tot het themanummer over het NMAP-rapport in de *MME*:

Practitioners, scholars, and researchers within the field of mathematics education were underrepresented on the panel, and accorded surprisingly little input to the style and content of the report. (p.365)

Deze ondervertegenwoordiging van leidinggevende denkers en doeners *from the field of mathematics education itself* komt duidelijk tot uiting, aldus Greer (2008, p.365), niet enkel in 'The make-up of the panel', maar ook 'in uncited literature, and in the lack of reference to the most prominent scholars in the field'.

Lobato (2008, p.595) wijst erop dat de NMAP-werkgroep rond 'learning processes' zelfs zo ver ging alle belangrijke tijdschriften op het terrein van de wiskundendidactiek (inclusief de vier toptijdschriften 'Journal for Research in Mathematics Education', 'Journal of Mathematical Behavior', 'Educational Studies in Mathematics' en 'Mathematical Thinking and Learning') niet eens in haar werkzaamheden te betrekken. Tijdschriften die wel

als relevant beschouwd werden, situeerden zich vrijwel uitsluitend op het terrein van de experimentele psychologie (e.g. 'Journal of Experimental Psychology'). In dezelfde zin signaleert Greer (2008, p.368) dat Frank Lester, editor van het 'Second NCTM Handbook of Research on Mathematics Teaching and Learning' (Lester, 2006) - dat algemeen beschouwd wordt als het meest gezaghebbend internationaal standaardwerk op het terrein van het onderzoek van het wiskundeonderwijs - een brief schreef aan de voorzitter en vice-voorzitter van het panel die als volgt begon:

Dear Drs. Faulkner and Benbow, I am contacting the two of you in your roles as chair and vice-chair of the National Mathematics Advisory Panel to inform you of a resource that may be of value to the Panel in its deliberations. Specifically, I am editing a revision of the Handbook of Research on Mathematics Teaching and Learning that was originally published in 1992 by Macmillan. The 1992 handbook has been the most widely cited reference on research in mathematics education in the world and the second edition is likely to be just as valuable to the research community. (Lester, 2008).

De enige reactie die hij kreeg was een formeel ontvangstbewijs van de secretaris. Het panel was niet geïnteresseerd in het manuscript van dit ter perse zijnde handboek en in het NMAP-eindrapport wordt er dan ook geen enkele keer naar verwezen.

Greer (2008, p.366) besluit zijn commentaar op de samenstelling van de commissie en de impact daarvan op hoe de commissie met de onderzoeksliteratuur is omgegaan dan ook als volgt:

... the members (with commendable, but arguably misplaced, diligence) plowed through (and regurgitated) huge masses of empirical work, preselected according to strict criteria that excluded most of the deepest work in the field. In my opinion, a lot of their time and intellectual energy would have been much better spent on reading rich reflections on mathematics education, such as the work of Hans Freudenthal (1983, 1989).

### Theoretische en inhoudelijke uitgangspunten

#### *Cognitieve psychologie als enig uitgangspunt*

Het is gebruikelijk om binnen het onderzoek van de psychologie van leren en instructie - en van wiskundeleren en -onderwijzen in het bijzonder - onderscheid te maken tussen drie grote domeinoverstijgende<sup>10</sup> theoretische benaderingen:

- 1 de associatieve en/of behavioristische stroming, die zich richt op uiterlijk waarneembare gedragingen opgevat als reacties op prikkels uit de omgeving,
- 2 de cognitieve benadering, waarin de mens opgevat wordt als een doelgerichte verwerker van informatie; vandaar ook de veel gebruikte benaming informatieverwerkingsbenadering,

3 de situationele en/of socioculturele benadering, die de sociale en culturele context waarin dat menselijke denken en leren plaatsvinden centraal plaatst (Greeno, Collins, & Resnick, 1996; Verschaffel et al., 2007).

Bij het bekijken van de lijst van panelleden en het doornemen van het eindrapport valt meteen de oververtegenwoordiging op van de cognitieve benadering en de afwezigheid van de andere genoemde benaderingen, met name de derde benadering die de laatste tijd net op het terrein van *mathematics education* erg invloedrijk is. Een inhoudelijke argumentatie voor deze keuze wordt niet gegeven; de verantwoording geschiedt vanuit het 'methodologisch credo' van het panel dat eerder al kort aan bod kwam en dat verderop nog uitvoerig ter sprake komt. Verwijzend naar Vygotsky's socioculturele denkkader, argumenteren Geary et al. (2008, p.xiii, in Lobato, 2008, p.595) als volgt waarom zij geen (empirisch) onderzoek vanuit die theoretische invalshoek in hun analyse hebben betrokken:

A shortage of controlled experiments makes the usefulness of Vygotsky's approach for improving mathematics learning difficult to evaluate, and thus its utility in mathematics classrooms and mathematics curricula needs to be scientifically tested.

Ook Greeno and Collins (2008, p.619) betreuren dat het panel, omwille van dit (eng) methodologisch uitgangspunt (zie verder), al het onderzoek dat bekend staat onder benamingen als *sociocultural*, *activity-theoretical* of *situative* en waarin een benadering van (wiskunde)leren en onderwijzen wordt voorgestaan, vrijwel compleet genegeerd wordt:

That aims to be more comprehensive than the behaviorist and individual cognitive perspectives that have framed much previous research in education, including mathematics education.

### **Geen grondige reflectie op doelen en inhoud van wiskundeonderwijs**

Een tweede inhoudelijk tekort van het NMAP-rapport is, aldus de critici, het ontbreken van een grondige en inspirerende reflectie op de doelen en inhoud van het wiskundeonderwijs, rekening houdend met allerhande recente ontwikkelingen en uitdagingen op wetenschappelijk, maatschappelijk, en technologisch gebied. Het huidige wiskunde curriculum wordt, zo luidt de kritiek, als een vaststaand gegeven (een *fait accompli*) beschouwd, met als enige bedenking dat bepaalde (klassieke) onderdelen van het curriculum die de laatste tijd onvoldoende aandacht kregen of onvoldoende kwaliteitsvol werden onderwezen, zoals gehele en rationale getallen en algebra, meer prominent in het curriculum aanwezig moeten zijn en/of de methodische aanpak ervan een stevige oppoetsbeurt moeten krijgen. Maar mag men van

zo'n rapport niet meer verwachten, zo vragen onderzoekers als Greeno en Collins, Greer en Roth zich af. Meer bepaald, moet een dergelijk rapport niet evenzeer, en misschien zelf in de eerste plaats, de *what question of mathematics education* aanpakken (vgl. Van den Heuvel-Panhuizen, 2005)? Ik laat eerst Greeno en Collins (2008, p.620) hierover aan het woord:

We think it is time to reconsider what should be taught in the mathematics curriculum in light of the changing needs for mathematics expertise in the modern world. Certainly the need for such expertise is increasing, as the Panel's report makes clear, but at the same time the nature of the needed expertise is also changing. As one example, it seems fair to say that practically no one, including scientists and engineers, ever has occasion to compute the area of a triangle or trapezoid, use the quadratic formula, or add and multiply fractions.

Verwijzend naar enkele andere bijdragen tot hetzelfde themanummer van MME, argumenteert Greer (2008, p.366) in zijn inleiding op het themanummer dat deze inhoudelijke kritiek rechtstreeks samenhangt met de eenzijdige samenstelling van de commissie:

As pointed out by O'Brien and Smith (this issue), surprisingly few of the panel were researchers in mathematics education. Members from outside the field (including mathematicians) could not be expected to begin the process with the width and depth of multidisciplinary knowledge that those within the field attempt to achieve. (...) Hence the preponderance of psychological research contributes to the maintenance of the status quo as described by Roth (this issue).

Greer (2008, p.367) verwijst in dit verband naar een bijdrage van De Corte, Greer en Verschaffel (1996, p.492), waarin gewezen wordt op het grote verschil in benaderingswijze tussen (cognitieve) psychologen en *mathematics educators*:

A further source of tension between mathematics educators and psychological researcher is the balance among, in Bishop's words, 'what is', 'what might be', and 'what should be' (Bishop, 1992, p.714). Psychologists who take mathematics as an area of application tend to investigate the situation as they find it or perceive it, to take mathematics as an (uncontroversial) given. In the course of their research they may identify problems and make suggestions for improvements within the existing framework, but without questioning fundamental goals. Mathematics educators, by contrast, are more likely to call for radical change.

Zoals eerder gezegd, is een van de meest opvallende kenmerken van het rapport de zeer sterke nadruk die op algebra wordt gelegd. Volgens Greer (2008) is deze klemtoneel vreemd en onterecht:

The instruction by President Bush that directed the panel towards a focus on algebra may have partially contributed to its incoherence - perhaps it would have been more appropriate to call it the National Algebra Advisory Panel. In any

case, it is very noticeable that geometry, probability, and data handling receive little attention.

Moreover, geometry (in some cases) and combinatorics (throughout) are accorded attention only insofar as they feed into algebra. (p.368)

En hij gaat verder (Greer, 2008)

More serious yet, in my opinion, there is the almost total inattention to two fundamental aspects of doing mathematics, namely (a) applications and modeling, including data handling and statistical modeling, (b) problem solving (O'Brien & Smith, this issue). The first of these I consider of much more importance for responsible citizenship than a knowledge of algebra. (p.368)

Volgens Greer (2008) is het jammer dat het rapport dergelijke fundamentele kwesties, waarbinnen het wetenschapsgebied van *mathematics education* - met name door toedoen van het situationistisch perspectief - zoveel om te doen is, helemaal uit de weg gaat:

Many of the developments within the field of mathematics education manifest aspects of the general theme of mathematics as a human activity, including ethno mathematical research, studies of the history of mathematics, studies of people behaving mathematically in natural settings such as work, new directions in the philosophy of mathematics, attempts to establish connections between school mathematics and the lived experiences of the students. Within our field, it is being questioned whether mathematics as a school subject should continue to be dominated by mathematics as an academic discipline or should reflect more fully the range of mathematical activities in which humans engage. The report doesn't go there. (p.369)

En dat men Greer (2008,) er moeilijk van kan betichten om de wiskunde als hoogstaande intellectuele activiteit en wetenschappelijke discipline te willen reduceren tot een louter praktische en maatschappelijke (overlevings)vaardigheid, blijkt uit één van zijn andere verzuchtingen omtrent de inhoudelijke oriëntatie van het rapport:

Perhaps the strongest impression that the report leaves me with is a lack of any sense that mathematics is intellectually exciting, and could be taught in a way that makes it so, for at least a majority of students. (p.369)

### ***Kritiekloos inschrijven in een nationale economische rationaliteit***

Een laatste inhoudelijk kenmerk van het rapport dat in de kritische analyses aan de kaak gesteld wordt, is het kritiekloos aanvaarden van de wetmatigheden en principes van nationale economische groei en internationale wedijver als voornaamste (of zelfs enige) uitgangspunt voor het rapport. Spillane (2008) vat deze kritiek als volgt samen:

Embedded in these ongoing policy discourses, the NMAP report reflects some familiar themes. For example, mathe-

matics education is discussed chiefly in terms of social efficiency with some reference to social mobility, that is, meeting the needs of the economy and keeping the United States competitive in the world market. The democratic goal of education does not figure prominently unless reference to 'the safety of the nation and the quality of life' (NMAP, 2008, p.xi) and 'citizens and policy leaders who deal with the public interest in positions of civic leadership' (p. 3) count for that. (...) Mathematics education and its improvement in this report, and indeed in much of the current education policy conversation in the United States, are framed chiefly in terms of addressing economic needs. At the same time, the report contributes to legitimating these ideas as the 'real' or 'only' alternatives in these policy discourses. (p.640)

Ook Greer (2008, p. 367) schrijft dat het eindrapport van het NMAP op te vatten is als een antwoord op 'a number of factors combine to threaten the economic dominance of the United States' als geheel. Opvallend weinig aandacht is er, aldus Greer (2008, p.369) in het rapport voor 'the constellation of problems that arise from class and racial inequities', een kritiek die in een andere bijdrage tot het TMME themanummer verder uitgewerkt wordt.

### **Methodologische uitgangspunten**

#### ***Het 'methodologisch credo' van het NMAP***

Veruit de meeste en de scherpste kritiek vanwege de critici slaat echter niet over de inhoudelijke conclusies en aanbevelingen doch op:

- het 'methodologisch credo' van het panel en meer bepaald het vooropstellen van de 'gouden standaarden van de experimentele methode' (gerandomiseerd design, psychometrische tests, statistische onderbouwde conclusies), en de daarmee gepaard gaande miskennis van al het onderzoek dat niet aan deze gouden standaarden beantwoordt;
- de filter waarmee - als rechtstreeks gevolg daarvan - naar bruikbare empirische evidentie is gezocht;
- in het verlengde daarvan ook de (impliciete) aanbeveling met betrekking tot het soort van onderzoek waar in de (nabije) toekomst behoefte aan is en waar (dus) de onderzoeksmiddelen naar toe moeten gaan.

Voor alle duidelijkheid: niemand van de auteurs naar wie ik verderop nog (veelvuldig) zal verwijzen, stelt de waarde en de betekenis van het soort onderzoek dat door het panel als (meest) bruikbaar naar voren geschoven wordt, als zodanig ter discussie. Zij erkennen allemaal dat dergelijke onderzoeksmethoden en -technieken door een (belangrijk) deel van de gemeenschap van onderzoekers van (wiskunde-)leren en -onderwijzen gehanteerd worden en dat de bevindingen die ze opleveren bijdragen tot onze kennis van en inzicht in het vakgebied. Maar wat zij wel in scherpe bewoordingen bekritisieren, is dat het NMAP-rapport - zonder daar een overtuigende verantwoording voor te geven - die ene soort van (empirische)

evidentie bruikbaar acht (zonder te erkennen dat dergelijk onderzoek ook z'n beperkingen en gevaren kent) terwijl al het andere onderzoek als inferieur en onbruikbaar bestempeld wordt (zonder de eigen verdiensten van dit onderzoek te erkennen). Of, zoals Greeno en Collins (2008) het uitdrukken:

The report does not attempt to justify this choice of a methodological constraint. Education researchers know that this view of research quality is held by some of our colleagues, but it by no means expresses a consensus of the field. (p.618)

Men zou hiertegen kunnen opwerpen, dat het NMAP onmiskenbaar een duidelijke selectie gemaakt heeft in het bestaande onderzoek, maar daarmee heeft het panel zich toch niet uitgesproken over de waarde van de verschillende soorten van empirisch onderzoek over wiskundeleren en -onderwijzen als zodanig. Het panel heeft een beperkte, welomlijnde beleidsgerichte opdracht gekregen, namelijk de beleidsvoerders van de Verenigde Staten te adviseren over de gewenste inhoudelijke en methodische aanpak van het wiskundeonderwijs voor vier- tot veertienjarigen, uitgaande van de beschikbare wetenschappelijke kennis van wat er wel en niet 'werkt'. Het panel heeft gemeend dat de enige verdedigbare en haalbare werkwijze erin bestond enkel rekening te houden met (empirisch) onderzoek dat informatie opleverde over wat wel en niet 'werkt' en dat bovendien - volgens de gangbare wetenschappelijke maatstaven - aan de hoogste methodologische kwaliteitseisen voldeed. Greeno en Collins' (2008) antwoord daarop is echter als volgt:

'We have no quarrel with a report intended to inform the president and the public about research 'designed to investigate the effects of programs, practices, and approaches on students' mathematics learning and achievement.' (NMAP, 2008, p.2-3). On the other hand, the decision by the Panel to limit its interest to 'experimental and quasi-experimental research,' in the narrow sense of these terms adopted by the Panel, resulted in a gross misrepresentation of the body of research findings that could and should inform the nation's efforts to improve mathematics education and education in general. According to the Panel's report, the only way for research about 'programs, practices, and approaches' to contribute to improvement of educational practice is to produce a statistically significant difference between an innovative treatment and something else. This simplistic view does not represent the variety of ways in which researchers and professional educators are experimenting with arrangements to find ways to interact productively, including developing methods and evidentiary criteria for design-based research. (p.618-619)

Een echo van deze kritiek vinden we ook bij Thompson (2008):

So, by the Panel's definitions, most qualitative studies, especially teaching experiments and design experiments, fall under the Panel's low quality category and cannot provide even suggestive evidence for future research. As a result, the

report is not informed by a large portion of basic mathematics education research that investigates fundamental processes of classroom learning and teaching. (p.582)

### *De logica van 'Evidence-based Research and Development'*

Het hierboven beschreven 'methodologisch credo' van het NMAP kan niet los gezien worden van een discussie rond *evidence-based research and development* die in het Amerikaanse onderwijs- en onderwijsonderzoeksbeleid al een hele tijd aan de gang is (zie onder andere Schoenfeld, 2006, 2007).

Dit *evidence-based* principe komt uit de medische wetenschappen. In Wikipedia<sup>11</sup> vinden we volgende omschrijving ervan:

Evidence-based medicine (EBM) is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. (...) EBM seeks to clarify those parts of medical practice that are in principle subject to scientific methods and to apply these methods to ensure the best prediction of outcomes in medical treatment. (...) The foundation of evidence-based medicine is the systematic review of evidence for particular treatments, mainly randomized controlled trials.

Toen het 'U.S. Department of Education' (2002) haar strategisch plan voor 2002-2007 opstelde, trok men daarin resoluut de kaart van de *evidence-based policy*, zoals blijkt uit onderstaand citaat:

Unlike medicine, agriculture and industrial production, the field of education operates largely on the basis of ideology and professional consensus. As such, it is subject to fads and is incapable of the cumulative progress that follows from the application of the scientific method and from the systematic collection and use of objective information in policy making. We will change education to make it an evidence-based field. (p.48)

Het NMAP paste dit methodologisch principe dus bewust en nadrukkelijk toe. Met als gevolg (zie eerder) dat het panel, omwille van de vooropgestelde methodologische kwaliteitseisen, geconfronteerd werd met een bedroevend beperkte hoeveelheid bruikbare onderzoeksgegevens (NMAP, 2008):

As in all fields of education, the large quantity of studies gathered in literature searches on important topics in mathematics education is reduced appreciably once contemporary criteria for rigor and generalizability are applied. (p.xxvi)

Binnen de context van de in de Verenigde Staten aan de gang zijnde strijd tussen voor- en tegenstanders van 'reform-based mathematics education' was dit - (onderwijs- en onderzoeks)politiek gezien - een zeer belangrijke beslissing van het panel, want daarmee was meteen een negatief vonnis geveld over de wetenschappelijke kwaliteit - en dus de bruikbaarheid - van de vele *case stu-*



*dies, design experiments*<sup>12</sup> en *classroom-based research projects* die in het recente verleden vooral door *reformed-based math educators* in de Verenigde Staten verricht waren en die globaal genomen tot de conclusie hadden geleid dat leerlingen die volgens de principes van de *reform-based math approach* onderwijs hadden gekregen het door de band genomen minstens even goed doen 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* (voor overzichten en reviews van die evaluatiestudies op het niveau van het lager, middelbaar en hoger secundair onderwijs, zie respectievelijk Putnam, 2003, Chappell, 2003 en Swafford, 2003, in Senk & Thompson, 2003; zie ook Schoenfeld, 2007; Verschaffel et al., 2007),<sup>13</sup> omdat die studies over het algemeen niet beantwoordden aan de strenge methodologische criteria van *evidence based research*. En met die uitspraak nam het NMAP ineens ook een duidelijk standpunt in omtrent het soort van toekomstig onderzoek van wiskundeleren en -onderwijzen waar bovenal gebrek aan is en - derhalve - waar de onderzoeksmiddelen in de toekomst het beste naar toe moesten gaan...

#### **‘The reasonable ineffectiveness of (mathematics) education’**

De afgelopen jaren hebben verscheidene toonaangevende Amerikaanse onderwijsonderzoekers in het algemeen en onderzoekers van wiskundeleren en -onderwijzen in het bijzonder - in een reactie op de door de politiek sterk ondersteunde *evidence-based movement* - getracht om uit te leggen waarom het moeilijk en wellicht zelfs onmogelijk is om *(mathematics) education* om te vormen tot een ‘research and development’ (R&D) gebied waarvoor het *evidence based*-principe geldt. Of, anders gezegd, waarom *(mathematics) education* als een wetenschappelijke discipline niet in staat is om gegarandeerd succesvolle en algemeen geldende handelingsvoorschriften af te leveren, zoals dat in de landbouwwetenschappen of farmacie wel mogelijk blijkt te zijn. Kort en eenvoudig gesteld, worden daarvoor drie soorten verklaringen naar voren geschoven.

Een eerste reden waarom er op dit terrein nauwelijks onderzoek te doen is volgens de gouden standaard van de positieve wetenschappen, is dat het nu eenmaal - anders dan bij planten en dieren - zo lastig en meestal zelfs gewoon onmogelijk is om voldoende leerlingen of leraren zomaar *at random* te verdelen over bepaalde condities, aanpakken of methoden, tenminste wanneer men onderzoek wil doen naar interventies van enige betekenis en omvang én in een setting die enige relevantie heeft voor de reële onderwijsleersituatie. Boaler (2008) schrijft daarover het volgende:

In natural experiments, or quasi-experiments, researchers compare teaching approaches, not by sorting children into control and experimental groups and applying treatments, but by finding schools that use different approaches and

studying their effectiveness. The term *natural* should not be taken to suggest that such research is not purposeful and rigorous; it simply means that researchers find different teaching approaches and study them in their natural settings, rather than creating different conditions and studying them. (p.590)

Ten tweede is het studie-object - of beter gezegd, het studie-subject - zo immens multi-dimensioneel, zo complex en zo variabel dat voorspelbaarheid altijd problematisch blijft en dat dus, zelfs als men er in geslaagd is om die complexiteit en variatie experimenteel of statistisch te beheersen, de kracht van wat men gevonden heeft met betrekking tot één (onderzocht) aspect of onderdeel zal het altijd weer moeten afleggen tegen de werking van een veelheid van niet-bestudeerde andere factoren. De gevonden verbanden of causaliteiten zullen met andere woorden altijd opnieuw ‘gerelativeerd’ of ‘gesitueerd’ moeten worden. Cobb en Jackson (2008) verwoorden dit tweede probleem als volgt:

The key point to note for our purposes is that knowledge claims associated with experimental studies reflect a particular conception of the individual. The knowledge claims refer to an abstract, collective individual or statistical aggregate that is constructed by combining measures of psychological attributes of the participating students (e.g., measures of mathematics achievement). This statistically constructed individual is abstract in the sense that it does not correspond to any particular student. This methodological approach of investigating the performance of the abstract, collective subject rests on two underlying assumptions. First, students possess larger or smaller measurable amounts of discrete, isolatable psychological attributes (Danziger, 1990). This assumption makes it legitimate to combine measures of student performance. Second, the learning environments in which students acquire these attributes are composed of independent features that the investigator can manipulate and control directly. The implicit ontology is that of designed environmental settings made up of separate independent variables and students composed of collections of dependent psychological attributes. Together, these two theoretical suppositions ground experimental studies that seek to discern causal relationships between the manipulation of instructional conditions and the subsequent performance of the collective, abstract individual. (p.574)

De vraag is dus of het wel mogelijk en wenselijk is dat *al* het wetenschappelijk onderzoek van het wiskunde-leren en -onderwijzen dat verricht wordt, vanuit een dergelijk wetenschappelijk gezichtspunt vertrekt.

En ten derde is er nog de kwestie van de waardegeladenheid van het onderzoek van wiskundeleren en -onderwijzen: wat één persoon als een overtuigend bewijs voor de efficiëntie van een bepaalde onderwijskundige of vakdidactische visie, aanpak of methode zal beschouwen (bijvoorbeeld: de staartdeling snel en foutloos kunnen uitvoeren bij een gegeven kale som; de berekeningswijze voor het delen van een breuk door een breuk kennen en

vlot kunnen uitvoeren; ...), zal een andere als minder belangrijk of zelfs compleet irrelevant worden beschouwd omdat hij nu eenmaal, vanuit zijn of haar opvattingen over wiskunde en/of zijn of haar mens- en maatschappijvisie, andere doelen veel waardevoller vindt (Salomon & Ben-Zvi, 2006).

Opvallend in dit verband is bijvoorbeeld dat de effect-maten die 'traditionalisten' hanteren vaak anders c.q. beperkter zijn dan die van de 'reformisten', ook al beseffen laatstgenoemden dat sommige van de variabelen waar zij ook - of misschien zelfs bovenal - in geïnteresseerd zijn zich (veel) moeilijker lenen tot een objectieve, betrouwbare en valide meting.

Volgens Kilpatrick (1981, p.22) moeten we daarom durven erkennen dat onderzoek van wiskundelers en -onderwijzen hoe dan ook een zekere *reasonable ineffectiveness* heeft, en allicht ook altijd zal blijven hebben. Enigszins provocerend verwijst Kilpatrick naar Kristol (1973, p.62) die over de relatieve ineffectiviteit van het onderzoek rond allereerste complexe psychologische en maatschappelijke problemen het volgende zegt:

Everywhere we hear the refrain: 'We can go to the moon, can't we?' Well, why can't we do something equally marvelous about the ghettos or education or whatever? The answer is, of course, that going to the moon is easy whereas improving our system of education is hard. The one is nothing but a technological problem, the other is everything but a technological problem. Doing something about education means doing something about people - teachers, students, parents, politicians - and people are just not that manipulable.

Tegen deze achtergrond is het niet verrassend dat heel wat *mathematics educators* in de Verenigde Staten ontgoocheld en gegriefd gereageerd hebben op het enge methodologische uitgangspunt van het NMAP en een lans gebroken hebben voor de mede-erkenning van meer 'contextgevoelige' pogingen om waardevolle en bruikbare wetenschappelijke kennis te genereren.<sup>14</sup>

Ter afsluiting laat ik nog even drie gerenommeerde tenoren in dit debat nog een laatste keer hun fundamentele bezorgdheid uiten over het 'methodologisch credo' van het NMAP en de impact daarvan op de praktijk en het beleid van het wiskundeonderwijs in de Verenigde Staten.

Only certain types of research, meeting rigid methodological criteria, were considered to meet the requirement of 'best available scientific evidence'. If the goal is to improve mathematics education, there are many important questions - perhaps the most important that are not amenable to such forms of enquiry. (Greer, 2008, p.367)

An authentic report to the president and the American people should have reflected this dynamic condition.<sup>15</sup> Instead, the Panel chose to pronounce a narrow methodological doctrine that has adherents in the scientific community but is by no means settled ground. (Greeno & Collins, 2008, p.619)

Unfortunately, the Panel took an overly narrow view of what counts as scientific evidence, thereby failing to capitalize on much of what is known about mathematics learning and teaching. As a consequence, the Panel's report is less effective than it would otherwise have been in supporting policy makers and teachers to make substantial improvements. (Cobb & Jackson, 2008, p.573)

### *Onduidelijkheid over de evidentie waarop de conclusies en aanbevelingen zijn gebaseerd*

Zoals eerder gezegd, was in de door het NMAP te volgen procedure niet alleen voorzien dat elke werkgroep de hiervoor beschreven algemene methodologische eisen en categorisering kon verfijnen en aanpassen voor hun specifiek deelgebied, maar ook dat het panel ook over bepaalde prangende actuele kwesties een oordeel kon uitspreken dat niet *evidence-based* was doch louter op *collective professional judgment*. Maar volgens Confrey et al. (2008) is het vaak moeilijk, zo niet onmogelijk, om in NMAP-eindrapport onderscheid te maken tussen conclusies en aanbevelingen die gebaseerd zijn op die strenge methodologische criteria en andere die eerder de (persoonlijke) ervaringen, opinies en voorkeuren van de commissieleden weerspiegelen:

In the report - the document that policy makers will see - it is impossible to distinguish among summary statements, opinions, and evidence-backed statements (or to know that different standards underlie the various evidence-based statements, making some of them more trustworthy than others). Even the serious reader is confused upon seeing the following statement: A small number of questions have been deemed to have such currency as to require comment from the Panel, even if the scientific evidence was not sufficient to justify research-based findings. In those instances, the Panel has spoken on the basis of collective professional judgment, but it has also endeavored to minimize the number and scope of such comments. (NMAP, 2008, p.12) Although such statements are not unusual in policy documents, collective judgment must be carefully distinguished from evidence-based claims. (p.635)

Ook Thompson (2008, p.583), die de werkzaamheden en aanbevelingen van de werkgroep rond *curricular content* speciaal onder de loep heeft genomen, ergert zich aan die onduidelijkheid over de evidentie waarop de conclusies en aanbevelingen zijn gebaseerd. Hoewel het NMAP claimt het aantal commentaren en aanbevelingen die gebaseerd zijn op *collective professional judgment* tot een minimum te hebben beperkt, komt hij na grondig speurwerk tot de opvallende conclusie dat alle aanbevelingen die door de werkgroep rond 'curricular content' zijn gedaan en in het NMAP eindrapport zijn overgenomen tot stand zijn gekomen op basis van *professional judgment* (terwijl in het hele hoofdstuk over *curricular content* de uitdrukking *professional judgment* slechts één keer voorkomt). Daarbij komt nog, aldus Thompson (2008), dat het panel:

Consciously chose to ignore the professional judgment of more reform-minded experts who had perspectives that differed from the Panel's. (p.584)

En hij vervolgt (Thompson, 2008):

The task group's heavy reliance on professional judgment to make its recommendations calls into question its members' qualifications for exercising such judgment. For example, what biases did they have? How familiar were they with, and how well did they understand, the basic research literature on mathematics teaching and learning? The Panel's deep reliance on professional judgment in formulating its curricular content recommendations, without a transparent effort to represent competing viewpoints, leaves the Panel open to the charge that it might be composed of the most politically connected members of the community but not the most knowledgeable regarding mathematics learning, teaching, and curricula. (p.584)

De hele werkwijze heeft volgens Thompson (2008) tot de volgende *ironic result* geleid:

As a result of focusing only on investigations of efficacy or effect that passed through its self-defined evidence and quality standards, the task group does not cite a single result from basic research in mathematics education. At the same time, it found that studies of effect and efficacy of instructional or curricular approaches that meet its standards are rare. The net effect, then, is that chapter 4 of the Panel's report is neither informed by basic research in mathematics education nor grounded in high quality studies of efficacy and effect. Put another way, the Panel's content recommendations neither convey a reason for believing that these recommendations will work (because they are uninformed by any mechanisms from basic research by which they might be expected to work) nor inspire confidence that they will be effective (because of the paucity of experimental comparisons). (p.583)

Anders gezegd, zonder dat dit door het panel duidelijk erkend wordt, zijn vele van haar conclusies juist niet gebaseerd op resultaten van onderzoek dat beantwoordt aan de door de commissie zelf opgelegde standaarden van betrouwbaarheid en validiteit en is het vaak moeilijk, zonet onmogelijk, te achterhalen welke dat wel en niet zijn.

### *Een simplistisch model van 'Research and Development' (R&D)*

Een laatste belangrijk terugkerend punt van kritiek op het 'final report' van het NMAP betreft het simplistisch R&D model waar (impliciet) van uitgegaan wordt. Dit model, aldus Greeno en Collins (2008):

Assumes that the contributions of research to educational improvement are findings showing that an instructional treatment with some specified feature resulted in better average performance than an instructional treatment that lacked that feature. (p.620)

Het soort van (noodzakelijk geachte) *interface* tussen het onderzoek en de praktijk van het (wiskunde)onderwijs waar hierbij van uitgegaan wordt, is niet meer dan:

A set of recommendations that specify features of curricular content, teaching and teacher education, instructional approaches, practices, materials, and assessment. (Greeno & Collins, 2008, p.620).

Volgens hen heeft een dergelijk model ernstige tekorten:

It seems straightforward, but it neglects a fundamental feature of educational reality. A finding that something produced improvement on the average leaves open the question of what, in some specific situations, produced larger-than-average effects and what, in other situations, produced smaller-than-average or no or negative effects.

In her AERA presidential address, Ann Brown (1994) argued (as did John Seely Brown, 2002, for nonschool settings) that learning programs by the same name always vary. So whether a school or district should adopt a program cannot simply depend on that program's having had, on average, a good effect. It also needs to be conditioned on whether the resources the program needs to succeed are available in that setting. (p.621)

En, niet zo verwonderlijk voor onderzoekers die tot het situationisme paradigma behoren, vervolgen zij (p.621):

Efforts to aid educational practice need to recognize that it is deeply situated in settings of local resources and constraints. We believe a strategy of developing exemplary models and providing resources for their spread is much more likely to result in improvement of education than general recommendations are (...).

In dit verband verwijzen ze naar een alternatief R&D model - het zogenoemde *problem-solving research and development*-model, ontwikkeld door Resnicks team, waarin de relatie tussen onderzoek en onderwijs helemaal anders wordt opgevat:

The Panel report presupposes a traditional pipeline model of the flow of information from research to improvement of practice. According to this familiar, although overly simple, story, the complete pipeline has fundamental research at one end, which passes results to design and development, where instructional materials and resources such as curriculum or interactive software are constructed. These are evaluated in applied research, where the Panel's report picks up the story. Practitioners then receive lists of resources and programs that have been field tested and shown to have beneficial effects in the conditions of the evaluative research. A system of dissemination is needed to ensure widespread use of programs and resources that have been shown to be effective. In contrast, projects organized as problem-solving research and development have very different relations between researchers and professional educators from those of the traditional model. The goals of the project include both improving educational effectiveness and contributing to the advancement of fundamental understanding of educational processes. This combination of effort toward both practical

and theoretical scientific progress was discussed and advocated by Stokes (1997), who referred to work involving such combined goals as occurring in 'Pasteur's quadrant'.<sup>16</sup> A team of workers, including researchers, professional educators, and designers and developers, takes on a shared commitment to work on a practical problem of improving some aspect or aspects of educational effectiveness, to develop resources that can contribute to the achievement of that improvement, and to study processes of change in educational practice and processes of student learning in order to contribute to fundamental understanding in educational and learning sciences. Although members of the team are all committed to achieving the multiple goals of their joint effort, they retain different professional identities and primary responsibilities and accountabilities for accomplishing those several goals' (Greeno & Collins, 2008, p.621).<sup>17</sup>

In een andere bijdrage in het ER themanummer, dat helemaal aan deze kwestie is gewijd, onderschrijft Sloane (2008, p.625) Greeno en Collins' kritiek op het R&D model dat (impliciet) door het NMAP gehanteerd wordt en stelt hij vervolgens een model voor van de:

... phases of research required to support high-quality causal inference in mathematics education: 'Phase I studies are used to establish feasibility; Phase II studies are used to demonstrate initial efficacy and to demonstrate improvement over historical norms; and Phase III studies are used to confirm efficacy by use of comparative randomized trials. Finally, Phase IV studies require follow-up studies in the real world of scaling to ensure effectiveness. (American Statistical Association, 2007)

Verderop in zijn bijdrage werkt hij dit model overigens nog verder uit in een meer verfijnd continuüm bestaande uit tien opeenvolgende en in elkaar vloeiende stappen, waarbij hij opmerkt dat:

... without such an organizing structure it is unlikely that the Panel's goals will be met. (Sloane, 2008, p.627).

De tien stappen zijn: *basic research, hypothesis development and measurement, pilot-applied research, prototyping A, prototyping B, efficacy trials, effectiveness trials, implementation trials, sustainability research, and scaling studies*. Voor elke stap beschrijft hij in detail het soort onderzoeksvragen dat daarin aan de orde is, de meest aangewezen onderzoeksmethoden en -technieken, de eisen omtrent de grootte en samenstelling van de proefgroep en (statistische) analyse. Het zou me veel te ver voeren dit model in detail te bespreken, maar enkele punten ervan breng ik toch even onder de aandacht. Ten eerste vestigt Sloane (2008, p.625) er uitdrukkelijk de aandacht op dat:

For the four phases of the RCT model to be viable, a prephase of basic research must already exist.

Hij erkent dus het grote belang van basisonderzoek op het

terrein van (wiskunde) leren en -onderwijzen, dat nog aan de eerste fase van zijn model voorafgaat.

Ten tweede maakt hij onderscheid tussen *efficacy trials* en *effectiveness trials*:

*Efficacy trials* assess the value or worth of an educational treatment or program. To do so, these trials must provide tests of whether the treatment, instructional strategy, or curricular program does more harm than good when delivered under *optimal* conditions. In medical research, optimal conditions would include the highest quality production of the drug (or drugs) under study. Only later are such drugs produced under normal manufacturing conditions. In mathematics education, efficacy might include a selection of the most qualified teachers. (For example, those teachers originally trained by the curriculum innovator are assigned at random to treatment and control groups, ensuring, as best as one can, that the treatment has the highest possibility of being implemented with fidelity.) By way of contrast, *effectiveness trials* provide tests of whether the formally tested and efficacious treatment does more harm than good when it is delivered under real-world conditions. (Sloane, 2008, p.625).

Ten derde beklemtoont Sloane (2008) dat deze verschillende fasen op een geordende wijze op elkaar moeten volgen en in elkaar moeten uitvloeien om te komen tot wetenschappelijk verantwoorde uitspraken over de richting waarin het wiskundeonderwijs moet evolueren. Met het doen van aanbevelingen dat er zowel nood is aan kleinschalige als grootschalige (quasi)experimentele studies, zoals in het NMAP-rapport gebeurt, komen we niet veel vooruit, aldus Sloane (2008, p.626):

What becomes clear when we look carefully at these research phases is that (...) the phases are fundamentally nested in nature. That is, research at one level builds on that at another level or phase. This is a critical insight, as it suggests the possibility that the basic redress called for by the NMAP to add experimental studies, small and large, along some continuum of research is an overly simple response to the need to organize mathematics education research in support of improved learning on the part of students and teachers.

Een gelijkaardig voorstel om te komen tot wetenschappelijk verantwoorde (wiskunde)curricula, is enkele jaren geleden trouwens ook al uitgewerkt door Schoenfeld (2006; zie ook Burkhardt & Schoenfeld, 2003), eveneens in reactie op een - in zijn ogen - naïef beeld van hoe onderzoek het beleid en de praktijk van het onderwijs dient te beïnvloeden waar de voorstanders van het *evidence-based principle* in de Verenigde Staten van uit blijken te gaan. Omwille van de gelijkenis met Sloanes analyse en voorstel, ga ik er evenwel niet verder op in. Wie enigszins vertrouwd is met het onderzoeksbeleid in de Lage Landen zal het er allicht mee eens zijn dat de door Sloane en Schoenfeld voorgestelde R&D-modellen omwille van de beperkte beschikbare middelen voor onderwijsonderzoek moeilijk haalbaar zijn, al kunnen er

misschien toch zinvolle stappen in die richting worden gezet. Zo kunnen, mits voldoende zorgvuldig tewerkgegaan wordt bij de dataselectie en -analyse en mits de nodige terughoudendheid in acht wordt genomen bij de interpretatie van de uitkomsten, de resultaten van periodieke peilingen zoals PPOON, en wellicht ook PISA en TIMSS, helpen om beter zicht te krijgen op de samenhang tussen reken-wiskundemethoden (en/of andere relatief makkelijk te achterhalen kenmerken van wiskundeonderwijs) en opbrengsten van wiskundeonderwijs.

Een ander mogelijke piste bestaat erin om longitudinaal schooleffectiviteitsonderzoek, waarin grote cohorten van leerlingen gedurende een lange periode bij bepaalde onderdelen van hun schoolloopbaan gevolgd worden, te combineren met zorgvuldig gekozen en goed geplande en voorbereide interventies (bijvoorbeeld een of meerdere nieuwe aanpakken van wiskundeonderwijs) in een deel van de gevolgde scholen.

Ten slotte signaleer ik dat er onder onderzoekers van (wiskunde)onderwijs geen eensgezindheid bestaat over de plaats, functie of statuut van *case studies*, *design experiments* en andere niet-(strikt-)experimentele methoden in een R&D-programma en (dus) over het statuut van de wetenschappelijke en beleidsrelevante kennis die ze kunnen genereren. Terwijl volgens de een deze onderzoeksmethoden enkel waarde hebben in de initiële, verkennende fasen van het onderzoek en hoe dan ook altijd gevolgd moeten worden door meer systematisch, meer gecontroleerd en grootschaliger (quasi-)experimenteel onderzoek om tot gefundeerde en algemene uitspraken over (wiskunde)leren en -onderwijzen te kunnen komen (zie bijvoorbeeld Philips & Dolle, 2006), verdedigen anderen de stelling dat zij volkomen evenwaardige methoden zijn die net zo goed wetenschappelijke kennis kunnen genereren.

Zo argumenteert Boaler (2008):

There is a common belief that qualitative studies of a particular case are not generalizable because researchers have not looked across multiple settings. I disagree - qualitative case studies *can* provide highly generalizable findings, not by showing something repeated across cases but by providing the depth of observation and analysis that enables readers to understand a connection or phenomenon clearly and judge its applicability to other cases. (p.592)

Greeno en Collins (2008) zijn dezelfde overtuiging toegedaan:

Although the cases that are studied in analytical research of this kind<sup>18</sup> can include comparisons (as these examples did), the important findings of the research are not of the form Treatment A was better than Treatment B. Instead, design-based research and case studies provide empirical grounding that can expand the space of possibilities for education resources and practices. The conclusion is not just that something worked or worked on average better than something else, but that a particular kind of learning outcome is possible, along with detailed information about how that

learning can be brought about. Through these studies, and through detailed analyses of patterns of interaction in classrooms, the field is developing a more comprehensive theoretical understanding of learning and expansion of the space for designing and utilizing improved learning resources and practices. By ignoring this research, justifying its selectivity with its narrow criterion of evidentiary quality, the Panel's report neglected the most important issues that need to be addressed in order to improve mathematics education significantly. (p.620)

Cobb en Jackson (2008) aanvaarden evenmin dat er voor *design experiments* enkel een rol weggelegd is in de initiële, exploratieve fase van een onderzoekscyclus:

It is sometimes assumed that the sole purpose of design research studies is to refine an intervention that then can be evaluated experimentally. We have sought to clarify that although design research studies can serve this purpose, they and studies conducted using other methodologies can also produce forms of knowledge that have value in their own right and that can inform experimental studies and complement the forms of knowledge they produce. (p.576)

In hun argumentatie verwijzen deze auteurs geregeld naar Maxwell (2004), die onderscheid maakt tussen twee verschillende visies op causaliteit:

Maxwell (2004) would say that experimental psychologists tend to hold a *regularity* view of causality. From this perspective, causality cannot be directly observed but can only be measured in the regularity of relationships between events. This necessitates the determination of a systematic relationship between inputs (e.g., an experimental condition or a curricular treatment) and outputs (e.g., a subject's performance in an experiment or the learning outcomes for a group of students). Mathematics education researchers typically rely on a different but compatible aspect of scientific explanation, which Maxwell calls *process causality*. This is based on a conceptual analysis of processes by which some events influence others. Whereas regularity causation is most useful in establishing that something happens, process causality addresses why or how it happened. (Lobato, 2008, p.595)

---

## 4 Tot slot

Uit deze uitvoerige bespreking van de achtergrond, inhoud en receptie van het *final report* van het NMAP is aan de ene kant gebleken dat, op rechtstreeks initiatief van de president van de Verenigde Staten, een panel bestaande uit een twintigtal onderzoekers na anderhalf jaar intensief zoek-, lees- en schrijfwerk een rapport heeft geproduceerd bestaande uit 45 *research-based* conclusies en aanbevelingen voor het elementair wiskundeonderwijs. Veel van deze conclusies en aanbevelingen vormen een redelijk goede (maar al bij al niet bijster veel zeggende) samenvatting van wat er uit het onderzoek af

te leiden valt omtrent vakinhouden die onvoldoende beheerst worden en (dus) meer aandacht verdienen, didactische aanpakken die meer of minder effectief blijken te zijn, aandachtspunten voor de lerarenopleiding, voor ‘assessment’, onderzoek, enzovoort.

Maar aan de andere kant is uit het voogaande ook gebleken dat er vanuit diverse hoeken - en met name vanuit kringen van *reform-based math educators*, wetenschappers die tot andere theoretische strekkingen behoren dan de informatieverwerkingsbenadering, kwalitatieve onderzoekers en *critical math theorists* - stevige kritiek geuit is op de samenstelling van de commissie, op een aantal belangrijke inhoudelijke tekorten en eenzijdigheden van het rapport, op het radicale methodologische uitgangspunt van het panel, alsmede op de naïviteit van het (impliciet) gehanteerde R&D model. Dat de vermelde kritieken geformuleerd zijn door internationaal gerenommeerde wetenschappers uit de *instructional science*, *instructional psychology* en *mathematics education* en onder andere gepubliceerd zijn in een van de toptijdschriften van het ruime vakgebied (‘Educational Researcher’), verklaart mede waarom ik het de moeite vond om zo uitvoerig stil te staan bij de kritieken en argumenten van deze critici.<sup>19</sup>

Ik ben deze bijdrage begonnen met een verwijzing naar de *math war* die thans in Nederland volop aan de gang is en die mogelijk zal leiden tot de oprichting van een panel in de schoot van de KNAW met een opdracht gelijk aan die van het NMAP in de Verenigde Staten.<sup>20</sup> Hoewel het een grove vergissing zou zijn om het Nederlandse en Amerikaanse (onderzoek van het) wiskundeonderwijs over één kam te scheren, zal dat KNAW-panel, als het er komt, met gelijkaardige problemen, spanningen en uitdagingen geconfronteerd worden als het NMAP en kan het daarom nuttig zijn lessen te trekken uit de wijze waarop het NMAP is te werk gegaan bij het samenstellen van de commissie, bij het aflijnen van haar opdracht, bij het bepalen en funderen van haar inhoudelijke en methodologische uitgangspunten, bij het formuleren van haar eindrapport, maar ook uit de (kritische) wijze waarop het rapport in bepaalde kringen is ontvangen. Enkele belangrijke aandachtspunten wil ik tot slot van deze bijdrage nog even herhalen.

Ten eerste, als het doel is het wiskundeonderwijs op peil te houden of te verbeteren, moeten we erkennen dat er vele vragen en problemen zijn die niet zomaar op te lossen zijn met de *evidence-based approach* waar in het *final report* van het NMAP zo resoluut en radicaal voor gekozen wordt. Als er in Nederland een gelijkaardig panel komt, dan valt te hopen dat het er niet van uitgaat dat het maken van een synthese van de beschikbare empirische evidentie van wat goed wiskundeonderwijs is, een klus is die snel en eenvoudig kan worden geklaard. Waarmee ik niet heb gezegd dat deze commissie de lat voor wat ze als goed en bruikbaar onderzoek erkent, niet

hoog zou moeten leggen. Maar bij het bepalen van de standaarden van onderzoekskwaliteit zal men moeten aanvaarden dat een bepaalde maatstaf van kwaliteit van empirisch onderzoek die in andere gebieden zo vanzelfsprekend en algemeen aanvaard (b)lijkt te zijn, niet zonder meer transfereerbaar is naar het vakgebied van ‘mathematics education’, omwille van de complexiteit, oncontroleerbaarheid en onvoorspelbaarheid van het studie-object en omwille van het onvermijdelijk waardegeladen karakter van de vraagstellingen. Een duidelijke explicitering en verantwoording van de selectiecriteria is hoe dan ook aangewezen.

Ten tweede mag de bovenstaande analyse niet (verkeerd) begrepen worden als een pleidooi om resultaten van *case studies* en *design experiments* zonder meer te gebruiken als doorslaggevend bewijsmateriaal voor de waarde van een bepaalde (vakdidactische) methode of aanpak. Zelf ben ik de mening toegedaan dat resultaten van geslaagde *design experiments* niet volstaan als empirische verantwoording voor het doorvoeren van ingrijpende vernieuwingen op het vlak van doelen, inhouden en aanpakken van wiskundeonderwijs, zelfs niet wanneer meerdere van dergelijke ontwikkelingsonderzoeken globaal genomen in dezelfde (positieve) richting zouden wijzen. Aanvullende evidentie afkomstig van meer ‘afstandelijke’, grootschalige en gecontroleerde effectmetingen lijkt me hoogst aangewezen - ook met het oog op theorievorming trouwens (zie ook De Corte & Verschaffel, 2002; Phillips & Dolle, 2006). Waarmee niet gezegd is dat de uitkomsten en opbrengsten ervan niet mede gebruikt kunnen worden in een krachtig pleidooi voor de optimalisering van een bepaald aspect of onderdeel van het wiskundeonderwijs in één of andere richting.

Ten derde heb ik in deze bijdrage enkel ‘over een muurtje gekeken’, namelijk naar de Verenigde Staten. Maar ook in andere landen zijn analoge ontwikkelingen en discussies aan de gang. Interessant hierbij is dat in sommige van die andere landen beleidsgerichte rapporten verschijnen die een heel andere richting uitgaan dan dat van het NMAP in de Verenigde Staten. Zo is er het rapport van het ‘Centre for Research in Pedagogy and Practice’ in Singapore (Walshaw, 2008), waarin een versie van *evidence-based educational policy formation* gehanteerd wordt die niet gedomineerd wordt door de logica van *randomized trials* en *high-stake tests* en die wel ruimte laat voor triangulatie van kwantitatieve én kwalitatieve methoden. Een ander voorbeeld is het antwoord dat het Ministerie van Onderwijs in New Zealand gegeven heeft op de *evidence based policy development*, via een rapport (Anthony & Walshaw, 2006) waarin uitvoerig en overtuigend betoogd wordt dat complexiteit hét (systeem)kenmerk is van de onderwijspraktijk en dat dit noopt tot een pluriformiteit inzake methoden voor dataverzameling en data-analyse.

De volgende citaten uit dit rapport zijn illustratief voor de

duidelijk verschillende methodologische teneur van deze rapporten met dat van het 'math panel' in de Verenigde Staten.

The multilayered data base provided by the 'best evidence survey' did not offer linear causal explanations simply because the sheer complexity of the teaching learning relation precludes the possibility of identifying clear-cut cause effect relationships' (Sfard, 2005, p.407). Instead, the intent was to seek out those pedagogical practices that produced particular student outcomes in particular settings and to particular educational ends, rather than to make generalised pronouncements of 'effective' or 'quality' practice'. (Walshaw, 2007, p.21)

Core academic student outcomes were expanded from mastery of skills and concepts to exemplify mathematical proficiency that traced out the dispositions and habits of mind underlying what mathematicians do in their work. (...) Proficiency outcomes include the ability to think creatively, critically and logically; to structure and organise; and to process information. Outcomes also encompass an enjoyment of intellectual challenge and extend to the development of skills to interpret and critically evaluate mathematical information in a variety of contexts, and to solve problems that contribute to an understanding of the world. In addition to these academic outcomes, a range of social outcomes was identified (...). (Walshaw, 2007, p.20)

Ten slotte stip ik aan dat een duidelijke en overtuigende standpuntbepaling van wat men als 'empirische evidentie' erkent, slechts één van de zes eisen is waaraan een commissie zoals het NMAP (en dus ook de Nederlandse KNAW-commissie, als zij er komt) volgens Confrey et al. (2008) zou moeten voldoen, om tot geloofwaardige conclusies en aanbevelingen te kunnen komen die kans maken om een reële bijdrage te leveren aan de verbetering van het reken-wiskundeonderwijs.

Het hele lijstje ziet er als volgt uit:

1. Appropriate composition of an expert panel representing the relevant fields.
2. Articulation of methodological standards that are appropriate for the questions under investigation and are applied consistently and fairly for the duration of the panel's work.
3. Solicitation of appropriate testimony and documented responsiveness to it.
4. Provision of adequate time and opportunity for deliberation, and support to enable the panel to complete its charge.
5. Adequate independent external review for quality, impartiality, and objectivity.
6. Adequate internal review for internal consistency among preliminary reports and final products in terms of accuracy, conciseness, readability, and intellectual integrity. (p.631)

In hun bijdrage tot het eerdergenoemde ER-thema-nummer argumenteren Confrey et al. (2008, p.631) dat de

eerste, de tweede en de zesde eis niet voldaan zijn en dat het NMAP-rapport daarom een *detrimental effects on policy-related activity in the field* heeft gehad (Confrey et al., 2008, p.631). Als de KNAW-commissie komt, dan wens ik haar veel inspiratie, breedtenkendheid, wijsheid en moed toe om de vele valkuilen en hindernissen die ze op haar weg zal tegenkomen, het hoofd te bieden. Hopelijk kan deze bijdrage daar een bescheiden hulp toe bieden.

## Noten

- 1 Deze bijdrage is deels gebaseerd op een plenaire lezing gehouden op de ELWIEr Conferentie 'Over muurtjes heen kijken', 28 november 2008, Jaarbeurs, Utrecht.
- 2 Met dank aan Wim Van Dooren en drie anonieme beoordelaars voor de waardevolle commentaren op een eerdere versie van deze tekst.
- 3 Dit instituut is in de loop van de afgelopen decennia meermaals van naam veranderd.
- 4 Zie bijvoorbeeld de website van de 'Mathematically Correct' beweging in de Verenigde Staten.: <http://mathematicallycorrect.com/>, die begint met de stelling: ' $2+2=4$ , *There is a mathematically correct solution*' en wordt als volgt omschreven: 'This web site is devoted to the concerns raised by parents and scientists about the invasion of our schools by the 'New-New Math and the need to restore basic skills to math education'.
- 5 Op het moment van de lezing (november 2008) was het nog onduidelijk of die commissie er zou komen. Intussen is beslist dat zij er komt, en is zij trouwens al met haar werkzaamheden begonnen.
- 6 Een van de anonieme beoordelaars van dit manuscript, die overigens adviseerde om dit manuscript niet te publiceren, schreef bijvoorbeeld in zijn/haar review: 'Het is treurig dat een club van wijze mannen (en politici) gevoelig is voor dit klimaat en meent het 'maatschappelijk debat' naar zich toe te moeten trekken door een commissie in het leven te roepen die als een soort opperrechter een onmogelijk Salomonsoordeel moet vellen over de kwaliteit van het Nederlandse wiskundeonderwijs. Dit ondemocratisch initiatief heeft met de arrogantie van de macht dan met manmoedigheid te maken die de auteur de commissie toeschrijft.'
- 7 Zie: <http://nces.ed.gov/timss/htm>
- 8 Zie: <http://www.pisa.oecd.org/html>
- 9 Zie: <http://www.ed.gov/about/bdscomm/list/mathpanel/index.html>
- 10 Ik heb het woord 'domeinoverstijgende' hier toegevoegd omdat er binnen 'mathematics education' als wetenschappelijke discipline ook nog andere theoretische kaders bestaan en gebruikt worden, zoals het model van 'realistic mathematics education' (afgekort: RME).
- 11 Zie: ([http://en.wikipedia.org/wiki/Evidence-based\\_medicine](http://en.wikipedia.org/wiki/Evidence-based_medicine))
- 12 Het *design experiment* (of ontwikkelingsonderzoek) is een methode die zich vooral richt op kleinschalig kwalitatief onderzoek in enkele klassen of scholen - vaak in een nauw samenwerkingsverband tussen onderzoekers, ontwikkelaars en leerkrachten - met als typisch kenmerk het cyclisch ontwikkelen, uitproberen en bijstellen van nieuwe vormen van wiskundeonderwijs (Gravemeijer, 1994). Sceptici ten aanzien van deze methode vinden dat zij geen voldoende

- bewijs oplevert voor de effectiviteit van nieuwe onderwijsvormen en curricula, omdat er geen vooraf vaststaande, welomlijnde ‘interventie’ is, omdat het aantal klassen waarmee gewerkt wordt meestal (zeer) klein is, omdat er vrijwel nooit gewerkt wordt met (adequate) controle-groepen, en omdat de mogelijkheden om te generaliseren daarom veel te beperkt zijn.
- 13 Putnam (2003, p.161) geeft bijvoorbeeld de volgende samenvatting: ‘Students in these new curricula generally perform as well as other students on traditional measures of mathematical achievement, including computational skill, and generally do better on formal and informal assessments of conceptual understanding and ability to use mathematics to solve problems. These chapters demonstrate that ‘reform-based mathematics curricula can work’.
  - 14 Reeds in 2003 heeft het Research Advisory Committee van het National Council of Teachers of Mathematics in de U.S. trouwens hetzelfde standpunt ingenomen: ‘(f)or many reasons, it is a mistake to take the controlled experiment as the archetype for research that can improve school mathematics. Educational research is not simply a matter of comparing methods in the way one might compare medical treatments.’
  - 15 Greeno verwijst hiermee naar het gegeven dat er binnen ons wetenschapsgebied niet één algemeen erkend paradigma bestaat.
  - 16 Stokes’ eerste quadrant is wat we traditioneel verstaan onder *basic research* (fundamenteel onderzoek). Stokes noemt dit quadrant ook Bohrs quadrant, verwijzend naar Bohrs vroege onderzoek naar de atomaire structuur. Het tweede quadrant, Edisons quadrant, genaamd, omvat wat traditioneel *applied research* (toegepast onderzoek) genoemd wordt: onderzoek dat bovenal tot doel heeft een praktisch probleem op te lossen en niet zozeer gericht is op het vinden van een omvattend theoretisch verklaring-principe. Het derde quadrant, ook Pasteurs quadrant genaamd, slaat op onderzoek dat op beide terreinen tegelijk sterk staat. Immers, weinigen hebben meer bijgedragen tot ons fundamenteel inzicht in microbiologische processen dan Louis Pasteur, maar Pasteurs werk was evenzeer gericht op het oplossen van reële gezondheidsproblemen.
  - 17 Voor een heldere uiteenzetting over verschillende opvattingen en aanpakken van de relatie tussen onderwijsonderzoek, onderverwijfsvernieuwing en -innovatie, waarin de hierboven beschreven onderscheidingen terugkomen, verwijzen we naar Nelissen (2003).
  - 18 Greeno en Collins verwijzen hiermee naar *case studies*, *design experiments*, en dergelijke.
  - 19 Anderen hebben dit rapport van het NMAP eerder enthousiast onthaald en verwijzen er graag naar ter ondersteuning van hun kritiek op constructivistisch en/of realistisch geïnspireerd wiskundeonderwijs (zie bijvoorbeeld Feys, 2008).
  - 20 Zoals in voetnoot 1 vermeld, is de oprichting van dit panel intussen een feit.
- se students in Pangarau/Mathematics in centers and schools*. Wellington: Learning Media.
- Becker, J.P. & C. Selter (1996). Elementary school practices. In: A. Bishop, K. Clements, C. Keitel & C. Laborde (eds.). *International handbook of mathematics education. Part I*. Dordrecht: Kluwer, 511-564.
- Benbow, C.P. & L.R. Faulkner (2008). Rejoinder to the critiques of the National Mathematics Advisory Panel Final Report. *Educational Researcher*, 37(9), 645-648.
- Boaler, J. (2008). When politics took the place of inquiry: A response to the National Mathematics Advisory Panel’s review of instructional practices. *Educational Researcher*, 37(9), 588-594.
- Burkhardt, H. & A.H. Schoenfeld (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, 32(9), 3-14.
- Chappell, M. (2003). Keeping mathematics front and center: Reaction to middle-grades curriculum projects research. In: S. Senk & D. Thompson (eds.). *Standards-oriented school mathematics curricula: What does research say about student outcomes*. Mahwah, NJ: Erlbaum, 285-294.
- Cobb, P. & K. Jackson (2008). The consequences of experimentalism in formulating recommendations for policy and practice in mathematics education. *Educational Researcher*, 37(9), 573-581.
- Confrey, J., A.P. Maloney & K.H. Nguyen (2008). Breaching the conditions for success for a National Advisory Panel. *Educational Researcher*, 37(9), 631-637.
- Craats, J. van de (2007). Waarom Daan en Sanne niet kunnen rekenen. *Nieuw Archief voor Wiskunde*. 58(2), 132-136.
- Danziger, K. (1990). *Constructing the subject*. New York: Cambridge University Press.
- De Corte, E., B. Greer & L. Verschaffel (1996). Learning and teaching mathematics. In: D. Berliner & R. Calfee (eds.). *Handbook of educational psychology*. New York: Macmillan, 491-549.
- De Corte, E. & L. Verschaffel (2002). High-powered learning communities: Design experiments as a lever to bridge the theory-practice canyon. *Prospects*, 32, 517-531.
- Feys, R. (ed.). (2008). *Onderwijskrant*. (Themanummer: Mad Math en Math War), nr. 146.
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht, The Netherlands: Reidel.
- Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht, The Netherlands: Kluwer.
- Geary, D.C., A.W. Boykin, S. Embretson, V. Reyna, R. Siegler, D.B. Berch et al. (2008). *Report of the Task Group on Learning Processes*. Retrieved October 27, 2008, from <http://www.ed.gov/about/bdscomm/list/mathpanel/report/learning-processes.pdf>
- Good, T. (2008, April 7). Teaching mathematics in Grades 3-5 classrooms. Teachers College Record. Retrieved November 20, 2008, from <http://www.tcrecord.org/Content.asp?contentid=15196>
- Gravemeijer, K. (1994). *Developing realistic mathematics education*. Utrecht, The Netherlands: Freudenthal Institute, Utrecht University.
- Greeno, J.G. & A. Collins. (2008). Commentary on the Final Report of the National Mathematics Advisory Panel. *Educational Researcher*, 37(9), 618-623.
- Greeno, J.G., A.M. Collins & L.B. Resnick (1996). Cognition and learning. In: D.C. Berliner & R. C. Calfee (eds.). *Hand-*

## Literatuur

- American Statistical Association (2007). *Using statistics effectively in mathematics education research*. Alexandria: VA. Author.
- Anthony, G. & M. Walshaw (2006). *The best evidence synthesis of pedagogical approaches that facilitate learning for diver-*



- book of educational psychology. New York, NY: Macmillan, 15-46.
- Greer, B. (ed.) (2008). Critical notice on the National Mathematics Panel advisory report. *The Montana Mathematics Enthusiast*, 5(2-3), 365-428.
- Greer, B. (ed.) (2008). Reaction to the Final Report of the National Mathematics Advisory Panel. *The Montana Mathematics Enthusiast*, 5(2&3), 365-370.
- Heuvel-Panhuizen, M. van den (2005). Can scientific research answer the 'what' question of mathematics education? *Cambridge Journal of Education*, 35(1), 35-53.
- Hiebert, J., T.P. Carpenter, E. Fennema, K. Fuson, P. Human & H. Murray (1996). Problem solving as a basis for reform in curriculum and instruction. The case of mathematics. *Educational Researcher*, 25, 12-21.
- Hiebert, J., T.P. Carpenter, E. Fennema, K.C. Fuson, D. Wearne & H. Murray (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.
- Kelly, A.E. (2008). Reflections on the National Mathematics Advisory Panel Final Report. *Educational Researcher*, 37(9), 561-564.
- Kelly, A.E. (ed.) (2008). Special Issue on *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. *Educational Researcher*, 37(9), 561-648.
- Kilpatrick, J. (1981). The reasonable ineffectiveness of research in mathematics education. *For the Learning of Mathematics*, 2(2), 22-29.
- Kristol, I. (1973). Some second thoughts. *New York Times*, 8 January, 55(62).
- Lester, F.K. (ed.). *Second handbook of research on mathematics teaching and learning*. Greenwich, CT: information Age Publishing.
- Lobato, J. (2008). On learning processes and the National Mathematics Advisory Panel Report. *Educational Researcher*, 37(9), 595-601.
- Maxwell, J.A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, 33(2), 3-11.
- Moldavan, C. (2008). Ruminations on the final report of the National Mathematics Panel. *Mathematics Educator*, 18(1), 2-7.
- National Council of Teachers of Mathematics. (2008). *NCTM response to National Mathematics Advisory Panel report*. Retrieved November 11, 2009, from <http://www.nctm.org/NMPresponse.aspx>
- National Council of Teachers of Mathematics Research Advisory Committee (2003). Educational Research in the No Child Left Behind Environment. *Journal for Research in Mathematics Education*, 34 (3), 185-190.
- Nelissen, J. (2003). Opmattingen over innovatie en implementatie. *Tijdschrift voor Nascholing en Onderzoek van het Reken/Wiskundeonderwijs*, 21(4), 14-21.
- Phillips, D.C. & J.R. Dolle (2006). From Plato to Brown and beyond: Theory, practice and the promise of design experiments. In: L. Verschaffel, F. Dochy, M. Boekaerts & S. Vosniadou (eds.). *Instructional psychology: Past, present and future trends. Sixteen essays in honour of Erik De Corte*. Oxford: Elsevier, 277-392.
- Putnam, R. (2003). Commentary on four elementary mathematics curricula. In: S. Senk & D. Thompson (eds.). *Standards-oriented school mathematics curricula: What does research say about student outcomes*. Mahwah, NJ: Erlbaum, 161-178.
- Ralston, A. (2008). A nation still at risk. *MAA Focus*, 28(6), 18-19.
- Salomon, G. & D. Ben-Zvi (2008). The difficult marriage between education and technology: Is the marriage doomed? In: L. Verschaffel, F. Dochy, M. Boekaerts & S. Vosniadou (eds.). (2006). *Instructional psychology: Past, present and future trends. Sixteen essays in honour of Erik De Corte*. Oxford: Elsevier, 209-222.
- Schoenfeld, A. H. (2006). Method. In: F. K. Lester (ed.). *Second handbook of research on mathematics teaching and learning*. Greenwich, CT: information Age Publishing, 69-110.
- Schoenfeld, A.H. (2007). Problem solving in the United States, 1970-2008. Research and theory, practice and politics. *ZDM - International Journal on Mathematics Education*, 39(5-6), 537-552.
- Senk, S. & D. Thompson, D. (eds.) (2003). *Standards-oriented school mathematics curricula: What does research say about student outcomes*. Mahwah, NJ: Erlbaum.
- Sloane, F.C. (2008). Randomized trials in mathematics education: Recalibrating the proposed high watermark. *Educational Researcher*, 37(9), 624-630
- Spillane, J.P. (2008). Policy, politics, and the National Mathematics Advisory Panel Report: Topology, functions, and limits. *Educational Researcher*, 37(9), 638-644.
- Stokes, D.E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.
- Swafford, J. (2003). Reaction to high school curriculum projects research. In: S. Senk & D. Thompson (eds.). *Standards-oriented school mathematics curricula: What does research say about student outcomes*. Mahwah, NJ: Erlbaum, 457-468.
- Thompson, P.W. (2008). On professional judgment and the National Mathematics Advisory Panel Report: Curricular content. *Educational Researcher*, 37(9), 582-587.
- Treffers, A. (1987). *Three dimensions. A model of goal and theory description in mathematics education. The Wiskobas project*. Dordrecht, The Netherlands: Reidel.
- U.S. Department of Education (2002). *Strategic Plan 2002-2007*. Retrieved September 12, 2003 from <http://www.ed.gov/pubs/strat-plan2002-2007/index.html>
- U.S. Department of Education (2008). *Foundations for Success. The final report of the National Mathematics Advisory Panel*. Retrieved January 17 2009 from <http://www.ed.gov/about/bdscomm/list/mathpanel/index.html>.
- U.S. Department of Education. (2008). *Biographies of panel members—National Mathematics Advisory Panel*. Retrieved July 16, 2008, from <http://www.ed.gov/about/bdscomm/list/mathpanel/bios/index.html#panel>.
- Verschaffel, L. (2002). *25 jaar ontwikkelingen in het Nederlandse wiskundeonderwijs op de basisschool vanaf de zijlijn bekeken*. Uitgenodigde lezing gehouden in het kader van het achtste Symposium van de Historische Kring Reken- en Wiskunde Onderwijs over 'De roerige jaren zestig. Van Moderne Wiskunde naar Realistisch Wiskundeonderwijs', 25 mei 2002, Utrecht.

- Verschaffel, L. (2008). *Wat is er van realistisch wiskundeonderwijs gerealiseerd?* Uitgenodigde lezing gehouden op het symposium over 'Realistisch wiskundeonderwijs, hoe realistisch is dat' naar aanleiding van het afscheid van Prof. Dr. Koeno Gravemijer, Utrecht, Universiteit Utrecht.
- Verschaffel, L. & E. De Corte (1996). Number and arithmetic. In: A. Bishop, K. Clements, C. Keitel, & C. Laborde (eds.). *International handbook of mathematics education. Part I*. Dordrecht: Kluwer, 99-138.
- Verschaffel, L., B. Greer & E. De Corte (2007). Whole number concepts and operations. In: F. K. Lester (ed.). *Second handbook of research on mathematics teaching and learning*. Greenwich, CT: Information Age Publishing, 557-628.
- Walshaw, M. (2007). Best evidence research redefined in mathematics education. In: E. de Freitas & K. Nolan (eds.). *Critical insights and in(ter)ventions into mathematics education*. Springer: New York, 13-25.

---

*In this article we provide an extensive overview of the background, the content and the reception of the 'Final Report of the National Mathematics Advisory Panel in the U.S.'. This panel has tried to resolve the ongoing 'math war' between 'traditionalists' and 'reformists'. This review focuses on critical reactions from various groups of scientists - reform-based math educators, scientists who follow different theoretical approaches than the information-processing one, qualitative researchers, and 'critical math theorists' - on (1) the composition of the panel, (2) important shortcomings and imbalances in the observations and recommendations being made, (3) the panel's radical methodological assumption about what is to be considered as 'good (empirical) evidence', and (4) the panelists' simplistic model of R&D. Although it would be a big mistake to consider the mathematics education landscape in the U.S. and the Netherlands as identical, there are some interesting analogies between the 'math war' in the U.S. and the current commotion and discussion about the quality and future of (elementary) mathematics education in the Netherlands.*