

Wat werkt (niet) in het reken-wiskundeonderwijs?

Review van wetenschappelijk onderzoek naar de samenhang tussen het onderwijsleerproces en reken-wiskunde prestaties van leerlingen in het basisonderwijs

Welke factoren in het onderwijsleerproces hangen samen met de reken-wiskunde prestaties van leerlingen in het basisonderwijs? Deze vraag stond centraal in een review-studie die op verzoek van de Inspectie van het Onderwijs en NRO is uitgevoerd. De focus lag op factoren in het onderwijsleerproces die beïnvloedbaar zijn: knoppen waaraan gedraaid zou kunnen worden. Allereerst schetsen we een beeld van de reken-wiskunde prestaties van Nederlandse basisschoolleerlingen aan de hand van de meest recente peilingsonderzoeken in groep 6 en 8 en de eindtoetsen. Uit deze gegevens blijkt ook dat de verschillen in reken-wiskunde prestaties tussen leerlingen maar voor een klein deel, hooguit 10 procent, toe te schrijven zijn aan verschillen in het rekenonderwijs dat zij krijgen op school. Vervolgens zijn de resultaten uit internationaal en Nederlands onderzoek in kaart gebracht om de vraag te beantwoorden welke factoren van het onderwijsleerproces effectief zijn. Dit leverde een grote verzameling op van kenmerken van de les die positief samenhangen met reken-wiskunde prestaties, zoals een verscheidenheid aan interventies met specifieke instructie- en werkvormen, het toepassen van technologische en niet-technologisch hulpmiddelen en formatieve toetsing. Het is opvallend dat er geen robuuste onderzoeksresultaten zijn gevonden over de samenhang van leerstofaanbod of rekenmethode met reken-wiskunde prestaties. Verder vonden we een aantal kenmerken van de leerkracht, zoals reken-wiskunde kennis, en van de leerling, zoals rekenangst, die samenhangen met reken-wiskunde prestaties. Over de rol van kenmerken van de klas en van de school, zoals aanwezigheid van een rekencoördinator, zijn weinig onderzoeksresultaten gevonden. Alles bij elkaar genomen zijn er aanwijzingen dat de leerkracht de cruciale rol vervult in het vormgeven van goed reken-wiskunde onderwijs.

INLEIDING

Rekenen-wiskunde is een belangrijk leergebied op de basisschool. Er is dan ook veel aandacht voor de vraag 'wat werkt (niet) in het reken-wiskunde onderwijs?' Deze vraag leeft onder andere bij beleidsmakers en de Inspectie van het Onderwijs, ontwikkelaars van reken-wiskunde methoden, scholen, leerkrachten en onderzoekers. In het huidige artikel proberen we de resultaten van het vele onderzoek naar deze vraag in kaart te brengen. Dit is een bewerking van een uitgebreid rapport, dat op verzoek van de Inspectie en NRO is opgesteld (Hickendorff et al., 2017).

In het huidige artikel geven we eerst inzicht in hoe het ervoor staat met de reken-wiskunde prestaties van Nederlandse basisschoolleerlingen in groep 6 en 8, aan de hand van de resultaten van grootschalige peilingsonderzoeken en eindtoetsen. Daarbij gaan we ook in op de vraag hoe groot de verschillen tussen scholen zijn. Dit geeft inzicht in hoe groot de verschillen in effectiviteit van het rekenonderwijs momenteel zijn in Nederland.

Vervolgens gaan we in op de vraag wat wel en niet 'werkt' in het reken-wiskunde onderwijs.

Marian Hickendorff,
Terry Mostert, Hannah
van Dijk, Lisa Jansen,
Lisanne van der Zee,
Onderwijswetenschappen,
Universiteit Leiden
Marije Fagginger Auer,
Onderwijswetenschappen,
Universiteit Leiden en 10
voor de leraar, Den Haag¹

Hickendorff, M., Mostert,
T.M.M., Van Dijk, C.J.,
Jansen, L.L.M., Van der
Zee, L.L., & Fagginger
Auer, M.F. (2019). Wat
werkt (niet) in het
rekenwiskunde onderwijs?
Review van
wetenschappelijk
onderzoek naar de
samenhang tussen het
onderwijsleerproces en
rekenwiskunde prestaties
van leerlingen in het
basisonderwijs. *Volgens
Bartjens – ontwikkeling en
onderzoek*, 38(3), 41-49

Deze vraag hebben we geformuleerd als: Wat is de samenhang tussen beïnvloedbare factoren uit het onderwijsleerproces en de reken-wiskundeprestaties van leerlingen (aan het einde) van de basisschool?

Als definitie voor het onderwijsleerproces hanteerden we: alles wat er op en vanuit school gebeurt om een kind kennis en vaardigheden te laten verwerven, in dit geval specifiek op het gebied van het rekenen-wiskunde. Onder reken-wiskundeprestaties verstaan we in principe de leerlingprestaties op alle vier de domeinen van rekenen-wiskunde zoals die zijn beschreven in het referentiekader (Expertgroep doorlopende leerlijnen, 2008). Voor het gemak schrijven we in de rest van de review *rekenen*.

Deze vraag hebben we beantwoord aan de hand van een review van internationale en Nederlandse onderzoeksliteratuur. Als startpunt voor welke factoren binnen het onderwijsleerproces van invloed (kunnen) zijn op de rekenprestaties is gebruik gemaakt van schooleffectiviteitsliteratuur, en specifiek de review van Scheerens, Luyten, Steen, en Luyten-de Thouars (2007). In opdracht van de OECD en de Nederlandse en Duitse Ministeries van Onderwijs hebben zij een overzicht geconstrueerd van factoren die van invloed zijn op schooleffectiviteit. Dit overzicht hebben we aangevuld met factoren uit het praktische model van Marzano (2014) omdat dit model regelmatig in trainingen en op scholen wordt gebruikt. Ten slotte hebben we de factoren aangevuld aan de hand van het waarderingskader van de Inspectie van het Onderwijs (2008) zodat het aansluit bij de kennisbehoefte van de inspectie. Op deze wijze zijn de bevindingen uit de wetenschap, praktijk en de inspectie geïntegreerd binnen één overkoepelend schema. Factoren die daarin een rol spelen hebben we geordend aan de hand van op welk niveau ze een rol spelen. Hierbij wordt onderscheid gemaakt tussen factoren op het niveau van:

- de les, zoals een bepaalde instructie- of werkvorm;
- de leerkracht, zoals kennis, vaardigheden, en attituden;
- de klas, zoals een taakgerichte werksfeer;
- de school, zoals de rekenvisie en de aanwezigheid van een rekencoördinator;
- de leerling, zoals motivatie en rekenangst.

Uiteraard is dit een overmatige versimpeling van de werkelijkheid (zie bijvoorbeeld Biesta, 2007; Gravemeijer & Kirschner, 2007), maar het doel van deze indeling was enkel om enige structuur aan te brengen in de veelheid van factoren die een rol kunnen spelen. Verder hebben we ons alleen gericht op de beïnvloedbare factoren. Dit zijn 'knoppen waar je aan zou kunnen draaien'. Daarbij streefden we zo specifiek mogelijk de werkzame bestanddelen in kaart te brengen. Factoren die niet of nauwelijks beïnvloedbaar zijn zoals geslacht van de leerling en diversiteit van de klas hebben we buiten beschouwing gelaten, waarmee uiteraard niet de indruk gewekt moet worden dat die geen invloed hebben op de rekenprestaties. In de rest van het artikel beschrijven we de aanpak, bevindingen, conclusies en aanbevelingen van het onderzoek. Deze samenvatting is volledig, maar bevat noodzakelijkerwijs niet de details en nuances die eraan ten grondslag liggen, noch de literatuurverwijzingen naar de bronnen waarop de review is gebaseerd. Lezers die op bepaalde onderdelen meer willen weten verwijzen we dan ook naar het rapport van Hickendorff et al. (2017).

AANPAK

Onze aanpak bestond uit drie elementen. Het eerste element betreft nadere analyses op de meest recente grootschalige peilingsonderzoeken in Nederland, te weten het nationale onderzoek Periodieke Peiling van het Onderwijsniveau (PPON) bij leerlingen uit groep 8 uitgevoerd in 2011 (Scheltens, Hemker, & Vermeulen, 2013) en het internationale onderzoek *Trends in Mathematics and Science Study* (TIMSS) bij leerlingen uit groep 6 uitgevoerd in 2015 (Meelissen & Punter, 2016). Deze peilingen geven een beeld van de stand van zaken van de rekenprestaties van Nederlandse basisschoolleerlingen. Daarnaast zijn op de gegevens van beide peilingen nadere analyses uitgevoerd, die zich richtten op (a) de verschillen tussen scholen en (b) de samenhang tussen enerzijds rekenprestaties van leerlingen en anderzijds factoren uit het onderwijsleerproces, gemeten met leerkracht- en schoolvragenlijsten.

Het tweede en grootste element betreft een review van de wetenschappelijke literatuur naar de samenhang tussen factor(en) uit het onderwijsleerproces en de rekenprestaties van kinderen in de basisschoolleeftijd. Gezien de omvang van de onderzoeksvraag en de beperkte tijd hebben we de wetenschappelijke literatuur in twee categorieën verdeeld en voor beide een specifieke aanpak gedefinieerd. De eerste categorie was de internationale literatuur. Daar hebben we ons beperkt tot meta-analyses, waarin de resultaten van meerdere onderzoeken systematisch worden geïntegreerd, en reviews, waar het gaat om overzichtsartikelen over een bepaald onderwerp. Uiteraard is dat een flinke inperking van de verwerkte literatuur, maar tegelijkertijd leveren meta-analyses en reviews wel robuustere resultaten dan losse onderzoeken. De tweede categorie is literatuur waarin onderzoek wordt beschreven met Nederlandse basisschoolleerlingen. Omdat dit onderzoek in principe het meest relevant is voor de hoofdvraag en het aantal studies beperkt en daarmee behapbaar was, behandelden we hier wel individuele onderzoeken. Alle literatuur is via een systematische zoekstrategie verzameld, aangevuld met suggesties van wetenschappelijke experts.

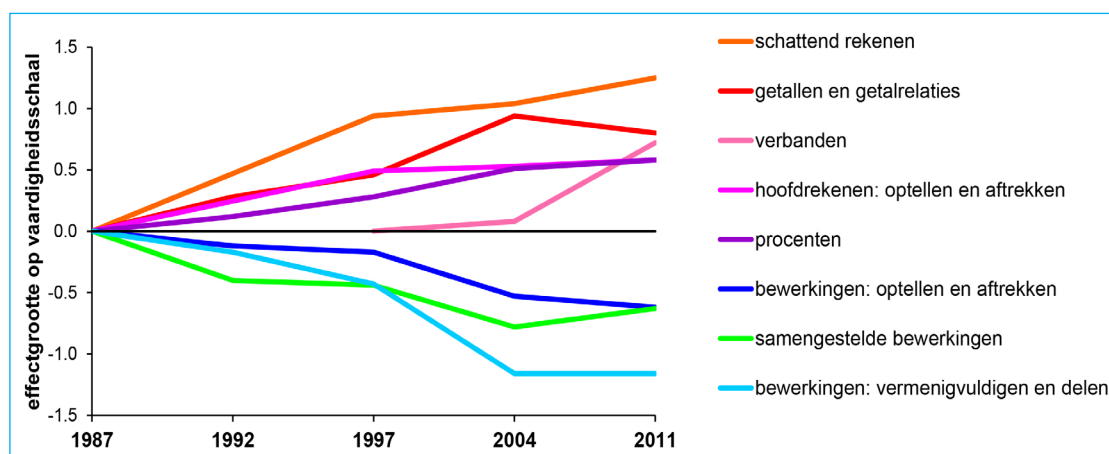
Het derde element betreft een veldraadpleging van praktijkexperts, bestaande uit leerkrachten, rekencoördinatoren, lerarenopleiders rekenen-wiskunde, methode-uitgevers, onderzoekers en beleidsmakers. Dit diende ter validatie én aanvulling van de bevindingen.

REKENPRESTATIES VAN NEDERLANDSE BASISCHOOLEERLINGEN

De meest grootschalige informatie die beschikbaar is over de huidige rekenprestaties van Nederlandse basisschoolleerlingen is afkomstig van de (inter)nationale peilingen. Er hebben in de afgelopen decennia twee typen grootschalige peilingen plaatsgevonden waarbij de rekenprestaties van Nederlandse basisschoolleerlingen in kaart werden gebracht: de Periodieke Peilingen van het Onderwijsniveau (PPON) in opdracht van het ministerie van OCW uitgevoerd door Cito, specifiek gericht op het Nederlandse basisonderwijs in groep 8; en de TIMSS in groep 6, die een internationale vergelijking mogelijk maken.

Periodieke Peiling van het Onderwijsniveau (PPON) in groep 8

In de jaren 1987, 1992, 1997, 2004 en 2011 zijn door Cito peilingen van het rekenniveau aan het einde van Nederlandse basisscholen verricht. In PPON-2011 is voor elk van de 22 gepeilde onderwerpen in kaart gebracht welke veranderingen in rekenprestaties zich hebben voorgedaan tussen 1987 en 2011 (Scheltens, Hemker, & Vermeulen, 2013) (afbeelding 1). Bij sommige onderwerpen zijn de veranderingen over deze periode gering, terwijl bij andere onderwerpen prestaties aanzienlijk zijn gedaald of gestegen. Na 2011 is nog geen nieuwe peiling uitgevoerd, maar Cito concludeert aan de hand van de resultaten van de jaarlijkse peilingen van het onderwijsniveau dat het overall rekenniveau grotendeels stabiel blijft (Hemker, 2018; Hemker & Van Weerden, 2015). Met betrekking tot absolute standaarden, de referentieniveaus, blijkt dat 93 procent van de groep-8-leerlingen in 2016-2017 het fundamentele niveau 1F haalde, waarmee wordt voldaan aan de ambitie dat minstens 90 procent van de leerlingen dit niveau zou moeten halen. Het percentage leerlingen dat het streefniveau 1S haalde blijft met 48 procent echter achter bij de ambitie van 65 procent (Hemker, 2018; Inspectie van het Onderwijs, 2018).



Afbeelding 1. Verandering van vaardigheidsniveaus voor de 8 van de 22 onderwerpen met grootste veranderingen (effectgrootte minstens 0,5) tussen PPON-1987 en PPON-2011 (Scheltens, Hemker, & Vermeulen, 2013)

Cito heeft voor de huidige review nadere analyses uitgevoerd op de gegevens van PPON-2011. Met multilevelanalyses is rekening gehouden met de niveaus waarop de gegevens zijn verzameld: groepen leerlingen zitten immers op dezelfde school. Deze analyses lieten zien dat 91-96 procent van de variantie in rekenprestaties toe te wijzen is aan verschillen tussen leerlingen die niet toe te schrijven zijn aan de school en slechts 4 tot 9 procent aan verschillen tussen scholen. Op het domein 'Verbanden' is het percentage variantie in rekenprestaties op schoolniveau het kleinst (4 procent), gevolgd door 'Meten en Meetkunde' (7 procent), 'Verhoudingen' (8 procent) en 'Getallen' (9 procent). De verschillen tussen scholen zijn dus op alle domeinen beperkt. Daarbij is het belangrijk in ogenschouw te nemen dat een aantal niet-beïnvloedbare factoren, zoals de sociaaleconomische status van de leerlingpopulatie, een deel van de verschillen verklaart. Zo had het gemiddelde opleidingsniveau van de ouders van de leerlingen op een school een significant effect op vrijwel alle rekenschalen en verklaarde daarmee een deel van de verschillen tussen scholen. Daardoor blijft er nog maar weinig variantie in rekenprestaties over die aan het rekenonderwijs zou kunnen worden toegeschreven.

Om meer inzicht te krijgen in welke factoren van het rekenonderwijs samenhangen met de rekenprestaties zijn verdere exploratieve analyses gedaan. De antwoorden van leerkrachten op een vragenlijst over het rekenonderwijsaanbod zijn daarin gerelateerd aan de rekenprestaties van leerlingen in verschillende rekendomeinen. Deze correlaties tussen leerkrachtenvariabelen en diverse rekenschalen waren slechts in enkele gevallen significant, namelijk bij drie tot vijf van de 48 leerkrachtvariabelen per schaal.

De gebruikte rekenmethode lijkt vooral samen te hangen met rekenprestaties bij 'Basisoperaties', waarbij de prestaties bij Wereld in Getallen hoger waren dan die bij Pluspunt, Rekenrijk en Alles Telt. Bij de andere onderwerpen maakte de gebruikte rekenmethode geen verschil². Rekenprestaties op diverse schalen waren beter bij oudere leerkrachten, leerkrachten die tevredener waren met de steun die leerlingen thuis ontvingen en leerkrachten die de rekenmethode die ze gebruikten beter bij zich vonden passen. Rekenprestaties op diverse schalen waren slechter bij leerkrachten die rapporteerden veelgemaakte fouten vaker klassikaal te bespreken en bij leerkrachten met meer leerlingen die externe steun ontvingen. Dat wil niet zeggen dat deze twee factoren daadwerkelijk een negatief effect op prestaties hebben, maar slechts dat ze ermee samengaan. Zo is het bijvoorbeeld niet meer dan logisch dat leerlingen die slechter presteren veelal meer externe hulp ontvangen.

Al met al hangen slechts enkele van de tientallen variabelen uit de leerkrachtvragenlijst op een beperkt aantal schalen samen met leerlingprestaties, waarbij het patroon over de verschillende schalen grotendeels vergelijkbaar is. De vragenlijst in de huidige vorm heeft daarmee slechts beperkte resultaten opgeleverd wat betreft de vraag welke factoren uit het onderwijsleerproces belangrijk zijn voor prestaties.

Trends in International Mathematics and Science Study (TIMSS) in groep 6

In de jaren 1995, 1999, 2003, 2007, 2011 en 2015 zijn in TIMSS de kennis en vaardigheid in rekenen en *science* getoetst van leerlingen in groep 6 van het Nederlandse basisonderwijs (Meelissen & Punter, 2016). In TIMSS-2015 scoorden Nederlandse leerlingen ver boven het algemene TIMSS-schaalgemiddelde voor rekenen. De score was wel significant lager dan die in 2011 en was de laagste die Nederland in de afgelopen twintig jaar heeft behaald, maar de verschillen zijn klein. De positie van Nederland ten opzichte van andere landen is in de afgelopen twintig jaar achteruit gegaan, zowel door Nederlands dalende score als door het vooruitgaan van veel andere landen.

Nederlandse leerlingen presteerden van de drie inhoudsdomeinen relatief het best op de domeinen 'Getallen' en 'Gegevensweergave', maar hun prestaties bij deze domeinen zijn in 2015 wel gedaald ten opzichte van 2011. De prestaties op het derde inhoudsdomein 'Geometrische vormen en meten' waren relatief het laagst, maar ongeveer hetzelfde bij de afgelopen drie peilingen. Dit onderwerp past volgens curriculumexperts en leerkrachten het minst bij het Nederlandse curriculum. Net als bijvoorbeeld leerlingen uit Scandinavische landen, Duitsland en Finland, presteerden Nederlandse leerlingen beter op redeneervragen dan op weet- en toepassingsvragen, waarbij de prestaties op weetvragen het laagst waren en ook gedaald in 2015 ten opzichte van 2011. In vergelijking met andere landen haalde een relatief groot gedeelte van de Nederlandse leerlingen het basisoniveau (99 procent in 2015), maar een relatief klein gedeelte het geavanceerde niveau (4 procent in 2015).

Ook op de gegevens van TIMSS-2015 zijn nadere analyses uitgevoerd, ditmaal door onderzoekers van Universiteit Twente in opdracht van het ministerie van OCW (Rebber, Van den Berg, & Meelissen, 2017). Net als in PPON-2011 bleek uit een multilevelanalyse dat een klein percentage van de variantie in rekenprestaties toe te schrijven was aan verschillen tussen scholen en klassen, rond de 10 procent. Ook in TIMSS blijken de verschillen tussen scholen in rekenprestaties dus beperkt, en interessant is dat dit nog sterker naar voren komt als we Nederland met andere landen vergelijken. In vier geselecteerde vergelijkingslanden - Vlaanderen, Engeland, Finland en Zweden - waren de verschillen tussen scholen en klassen (soms aanzienlijk) groter. Verder verklaarden in Nederland de leerlingkenmerken 'seks', 'taal thuis', namelijk leerling spreekt altijd of bijna altijd thuis de taal van de toets, en 'aantal boeken thuis', beschouwd als een indicator voor sociaal culturele status, samen 5 procent van de variantie op leerlingniveau en 3 procent op schoolniveau. Net als in PPON-2011 verklaarden een aantal niet-beïnvloedbare kenmerken dus een deel van de toch al beperkte verschillen tussen scholen. De mate waarin verschillen in rekenprestaties zijn toe te schrijven aan het rekenonderwijs lijkt dus gering.

Slechts één onderwijsleerprocesvariabele waarover leerkrachten zelf in een vragenlijst rapporteerden hing samen met de rekenprestaties van hun leerlingen: de waargenomen prestatiegerichtheid van het schoolklimaat. Opvallend was dat de mate waarin de inhoud van de toets aansloot bij het curriculum van groep 6 niet samenhang met de rekenprestaties.

SAMENHANG TUSSEN BEÏNVOEDBARE FACTOREN IN HET ONDERWIJSLEERPROCES EN REKENPRESTATIES

In de internationale onderzoeksliteratuur van na 2000 hebben we in totaal 33 meta-analyses en 19 reviews gevonden die de relatie tussen minstens één beïnvloedbare factor in het onderwijsleerproces en rekenprestaties of een breder domein van schoolprestaties waaronder in ieder geval rekenen hebben onderzocht bij leerlingen in de basisschoolleeftijd. Daarnaast hebben we apart een literatuurstudie gedaan naar onderzoek met Nederlandse basisschoolleerlingen. Daarvoor hebben we 26 (deel)studies bij Nederlandse basisschoolleerlingen geanalyseerd. Tot slot hebben we een veldraadpleging gehouden om onze bevindingen te valideren en aan te vullen. Samen met de bevindingen uit de nadere analyses

van de peilingsonderzoeken, leverde dit alles een grote verzameling van beïnvloedbare factoren op verschillende niveaus in het onderwijsleerproces die samenhangen met rekenprestaties.

De rekenles

Op het niveau van de les zijn de volgende beïnvloedbare factoren geïdentificeerd.

- Ten eerste blijkt uit de internationale en Nederlandse literatuur dat alle onderzochte interventies gericht op instructie- en werkvormen effectief zijn. Daarbij is het niet goed mogelijk om werkzame elementen te identificeren omdat de verschillende effectieve interventies soms tegenstelde uitgangspunten hebben, zoals interventies gericht op directe instructie of juist begeleidende instructie. Bovendien zijn er weinig studies die dit soort verschillende instructievormen rechtstreeks met elkaar vergelijken. Het voornaamste beeld dat rijst is dan ook dat 'iets manipuleren aan instructie- en/of werkvormen werkt'. Dit sluit ook aan bij de conclusie van het eerdere rapport van de Koninklijke Nederlandse Akademie voor Wetenschappen (2009) dat het (destijds) beschikbare Nederlandse onderzoek geen eenduidig beeld oplevert over de relatie tussen rekendidactiek en rekenvaardigheid, maar wel dat 'iets doen' aan de rekenles helpt.
- Ten tweede heeft het toepassen van technologische hulpmiddelen, zoals oefenprogramma's op de computer of tablet, en niet-technologise hulpmiddelen zoals werken met driedimensionale materialen een positieve relatie met rekenprestaties. Dit blijkt uit zowel internationaal als Nederlands onderzoek, zoals naar het gebruik van mini-games en de programma's Snappet en Rekeningtuin. Omdat het niet duidelijk is in hoeverre het toepassen van hulpmiddelen gepaard ging met extra aandacht of onderwijstijd voor rekenen is nader onderzoek gewenst naar wat precies het positieve effect teweeg brengt: het hulpmiddel of enkel 'iets toevoegen' aan de rekenles.
- Ten derde heeft formatieve toetsing, het gebruik van toetsgegevens om het leerproces te verbeteren, een positieve samenhang met rekenprestaties. Dit werkt zowel middels feedback aan de leerkracht, bijvoorbeeld via digitale leerlingvolgsystemen, als middels feedback aan de leerling per gemaakte opgave.
- Ten vierde bleek het differentiëren in niveaugroepen (homogeen of heterogeen) in internationaal onderzoek positief samen te hangen met rekenprestaties. Daarbij is het wel belangrijk om op te merken dat nader onderzoek gewenst is naar de specifieke en vergaande vorm van differentiëren die in Nederlandse rekenmethoden gebruikelijk is waarbij vaak gebruik wordt gemaakt van drie niveaugroepen (zie bijvoorbeeld Van Zanten, Van Graft, & Van Leeuwen, 2017).
- Ten slotte bleek uit internationaal onderzoek dat het maken van huiswerk positief gerelateerd is aan de rekenprestaties. Daarbij moet opgemerkt worden dat nog onduidelijk is hoe dit in de Nederlandse situatie zit.

Opvallend is dat er ook in de internationale literatuur geen robuuste onderzoeksresultaten zijn met betrekking tot leerstofaanbod of rekenmethode. Dit sluit aan bij de bevindingen van de nadere analyses van PPON-2011 en TIMSS-2015. Hoewel het belang van deze factoren voor de hand ligt, zie ook Hiebert en Grouws (2007) en Van Zanten en Van den Heuvel-Panhuizen (2014) lijkt het moeilijk deze gericht te onderzoeken. Mogelijk komt dit doordat de begrippen breed zijn, het leerstofaanbod sterk samenhangt met de wettelijke referentieniveaus en er daarom weinig variatie in aanbod bestaat, doordat de gebruikte rekenmethode samenhangt met andere school- en leerkrachtfactoren waardoor de effecten van rekenmethode niet zuiver te bepalen zijn, of doordat leerkrachten de rekenmethode verschillend gebruiken.

De leerkracht

Wat betreft de leerkracht hangen zowel meer algemene leerkrachtvaardigheden, namelijk effectief pedagogisch handelen en klassenmanagementvaardigheden, als rekenspecifieke kennis, namelijk vakinhoudelijke kennis en vakdidactische kennis (vaak aangeduid als PCK), positief samen met rekenprestaties. Omdat kennis van de leerkracht door de praktijkexperts de meest genoemde factor was, lijkt het zinnig om nader onderzoek te doen naar de specifieke bijdrage van verschillende componenten van vakinhoudelijke kennis en vakdidactische kennis. Uit de nadere analyses van PPON-2011 bleken verder de leeftijd van de leerkracht, de tevredenheid met de steun thuis, en het antwoord op de vraag hoe goed de rekenmethode bij de leerkracht past positief samen te hangen met rekenprestaties. De eerste twee factoren zijn niet of nauwelijks beïnvloedbaar. De derde factor sluit aan bij de hiervoor genoemde redenen waarom de rekenprestaties niet verschillen naar rekenmethode: mogelijk speelt het een belangrijke rol hoe goed de leerkracht de methode vindt passen. Nader onderzoek naar deze hypothese is gewenst.

Uit Nederlands onderzoek blijkt dat leerkrachtprofessionalisering op het gebied van vakkennis, differentiatievaardigheden en formatief toetsen, variërend van formatieve toetspraktijken in de klas tot het stimuleren van opbrengstgericht werken op het niveau van de school en de leerkracht, mogelijk effectief zijn. De resultaten zijn echter (nog) niet overtuigend. Intensieve en duurzame ondersteuning lijkt belangrijk. Over het effect van leerkrachtovertuigingen vonden we weinig onderzoeksliteratuur, behalve een negatieve invloed van het genderstereotype verwachtingen. Op basis van meer algemene schooleffectiviteitsliteratuur (Scheerens et al., 2007) en de mening van de praktijkexperts is het wel goed mogelijk

dat de prestatieverwachtingen van de leerkrachten en visie op (reken)didactiek samenhangen met de rekenprestaties van leerlingen. Deze hypothesen verdienen dus nader onderzoek.

De klas

Over de resultaten met betrekking tot kenmerken van de klas zijn geen relevante meta-analyses en reviews, onderzoek met Nederlandse leerlingen of relevante resultaten uit TIMSS of PPOON gevonden.

De school

Eén review ging in op evaluatie van het onderwijs door middel van summatieve toetsgegevens, middels accountability-programma's, die scholen verplichten tot afname van toetsen en rapportage van de resultaten, en schoolprestatiefeedback, waarbij in principe geen sprake is van sancties. Van beide vormen van evaluatie is het effect niet overtuigend aangetoond, en accountability-programma's leiden mogelijk tot negatieve bijeffecten. We hebben geen meta-analyses of reviews gevonden naar de effecten van andere kenmerken van de school zoals het rekenbeleid.

Onderzoek met Nederlandse leerlingen geeft aanwijzingen dat bij de implementatie van een programma de intensieve en duurzame ondersteuning vanuit de school (de directie) bijdraagt aan de leeropbrengsten. De resultaten van TIMSS-2015 leveren de hypothese dat prestatiegerichtheid van het schoolklimaat een positieve relatie met rekenprestaties heeft. Bovendien gaven de praktijkexperts aan dat het rekenbeleid op school een belangrijke factor is, bijvoorbeeld de aanwezigheid én beschikbare tijd en kwaliteit van een rekencoördinator, de hoeveelheid buitencurriculaire activiteiten op het gebied van rekenen, en de stimulering van samenwerking binnen teams. In Nederland is dit voor zover wij weten nog niet onderzocht, maar uit buitenlands onderzoek blijkt dat een rekencoach op school, die beschikbaar is voor leerkrachten en hen ondersteunt in hun eigen ontwikkeling binnen het rekendomein, zorgt voor betere prestaties bij leerlingen (zie bijvoorbeeld Taylor, 2017; Vale et al., 2010). Uit buitenlands onderzoek blijkt verder dat een instructionele leiderschapsstijl van het schoolhoofd een positieve invloed op de rekenprestaties kan hebben (Grissom, Loeb, & Master, 2013; Heck, 2000). Dit is een vorm van rekenbeleid waarin de verbetering van lesgeven door de leerkrachten en de bevordering van de leerlingprestaties centraal staan, door het definiëren van een missie, beheer van curricula, toezicht houden op lesgeven, volgen van de voortgang van leerlingen en het bevorderen van een positief leerklimaat.

De leerling

Onderzoek naar kenmerken van de leerling, ten slotte, liet zien dat de naïeve *beliefs* die leerlingen vaak hebben over rekenen, zoals dat rekenkennis bestaat uit losse stukjes informatie, externe attributie van resultaten en rekenangst negatief samenhangen met rekenprestaties. Rekenangst lijkt zowel oorzaak als gevolg van (lage) rekenprestaties te zijn. Over de andere overtuigingen – motivatie, *self-efficacy*, vertrouwen in competentie om taak uit te voeren, en attitude ten opzichte van rekenen – zijn geen meta-analyses of reviews gevonden. Onderzoek met Nederlandse basisschoolleerlingen liet wel zien dat zelfvertrouwen en motivatie positief samenhangen met rekenprestaties.

CONCLUSIES EN SLOTBESCHOUWINGEN

Allereerst bleek uit de meest recente peilingsonderzoeken in groep 6 en 8 en de eindtoetsen dat de rekenprestaties gemiddeld genomen stabiel zijn gebleven. Tegenover onderdelen waarop de prestaties achteruit zijn gegaan staan onderdelen waarop leerlingen het juist beter zijn gaan doen. Voldoende leerlingen halen het fundamentele niveau 1F, maar te weinig leerlingen halen het streefniveau 1S. In beide peilingen bleken de verschillen in rekenprestaties tussen Nederlandse basisschoolleerlingen maar voor een klein deel, hooguit 10 procent, toe te schrijven zijn aan verschillen tussen scholen. De verschillen tussen scholen zijn dus beperkt en ook kleiner dan in andere landen. Daarbij moeten we ook in ogenschouw nemen dat een aantal niet-beïnvloedbare factoren, zoals de sociaaleconomische status van de leerlingpopulatie bepaald door thuistaal en opleidingsniveau van de ouders, ook een deel van de toch al beperkte verschillen verklaren. De mate waarin scholen van elkaar verschillen in effectiviteit van het rekenonderwijs lijkt dus gering. Dat betekent niet dat scholen geen invloed *kunnen* hebben, maar laat vooral zien dat in de huidige situatie scholen erin slagen rekenonderwijs van een goede basiskwaliteit te leveren.

Vervolgens zijn de resultaten uit internationaal en Nederlands onderzoek in kaart gebracht om de vraag te beantwoorden welke factoren van het onderwijsleerproces effectief zijn. Dit leverde een grote verzameling van kenmerken van de les die positief samenhangen met reken-wiskundeprestaties, zoals een verscheidenheid aan interventies met specifieke instructie- en werkvormen, het toepassen van technologische en niet-technologische hulpmiddelen en formatieve toetsing. Veel discussies over het rekenonderwijs richten zich op welke instructievorm of didactiek het best 'werkt'. Het bestudeerde onderzoek laat echter zien dat tegengestelde instructievormen, zoals directe instructie en meer begeleidende instructie, beide werken, in de zin dat er betere prestaties worden behaald vergeleken met controle-groepen: iets veranderen aan de les werkt. Er is echter geen overtuigend bewijs is dat het ene beter werkt

dat het andere. Het is ook opvallend dat er geen robuuste onderzoeksresultaten zijn gevonden over de samenhang van leerstofaanbod of rekenmethode met reken-wiskunde-prestaties. Verder vonden we een aantal kenmerken van de leerkracht, zoals vakinhoudelijke kennis over rekenen, en van de leerling, zoals rekenangst, die samenhangen met reken-wiskunde-prestaties. Over de rol van kenmerken van de klas en van de school, zoals aanwezigheid van een rekencoördinator, zijn weinig onderzoeksresultaten gevonden, al geeft buitenlands onderzoek aanwijzingen dat een rekenbeleid gericht op verbetering van lesgeven en bevorderen van (reken)prestaties, bijvoorbeeld middels aanstellen van een rekencoach of -coördinator, rekenprestaties positief kunnen beïnvloeden.

BEPERKINGEN EN AANBEVELINGEN

De huidige review kent een aantal beperkingen. Een eerste categorie beperkingen is gekoppeld aan de analyse van (internationale) reviews en meta-analyses. Deze hebben het begrip 'rekenprestaties' meestal vrij smal gemeten, vaak beperkt tot het domein 'Getallen'. De bevindingen zijn dus niet automatisch te generaliseren naar de overige domeinen uit het referentiekader, en ook niet naar hogere-orde vaardigheden zoals probleemoplossen. Een tweede beperking is dat het feit dat er in onze zoektocht geen meta-analyses of reviews gevonden zijn over een bepaalde factor, zoals prestatieverwachtingen van de leerkracht, niet betekent dat er geen onderzoek bestaat. De breedte van het onderzoeksdomein 'factoren in het onderwijsleerproces' en de beschikbare tijd voor de review lieten het echter niet toe om individuele studies op te sporen. Het verdient aanbeveling dit wel te doen om een completer beeld te schetsen. Om die reden hebben we wel alle individuele studies met Nederlandse leerlingen opgenomen in het huidige review, maar die geven geen complete dekking van de factoren in het onderwijsleerproces. Ten derde is het opknippen van het onderwijsleerproces in losse factoren en apart de effecten onderzoeken een overmatige versimpeling van de complexiteit van het onderwijsleerproces. In werkelijkheid zullen verschillende factoren op elkaar inspelen en met elkaar samenhangen. Ten slotte beperkt de focus op empirische onderzoeken naar rekenprestaties de reikwijdte van de review. Zo blijft de literatuur over leertheorieën in het algemeen en reken-wiskundedidactiek in het bijzonder hoofdzakelijk buiten beschouwing.

Een andere categorie beperkingen betreft de nadere analyses van PPON-2011 en TIMSS-2015. Analyses van de samenhang tussen factoren in het onderwijsleerproces en de leeropbrengsten van leerlingen in een peilingsonderzoek zijn per definitie correlatieel. Dat betekent dat wel kan worden vastgesteld of een bepaalde factor zoals de gebruikte rekenmethode met rekenprestaties samenhangt, maar niet of de factor de uitkomsten beïnvloedt, in de zin dat er sprake is van een causaal verband. Als er een samenhang tussen een 'knop waaraan te draaien valt' en rekenprestaties gevonden wordt betekent dit nog niet dat het 'draaien aan die knop' ook echt resulteert in betere rekenprestaties. Men moet dus heel voorzichtig zijn met de conclusies en praktijkbevelingen die op basis van dit soort correlatieve onderzoeken gedaan worden.

Om robuustere conclusies te trekken zijn meer gecontroleerde onderzoeken nodig. Experimenten of quasi-experimenten waarbij een bepaalde (les)factor systematisch gevarieerd wordt zijn methodologisch sterk maar in de onderwijspraktijk niet altijd praktisch haalbaar. Bovendien kunnen vraagtekens gezet worden bij de vertaalslag naar de grootschalige onderwijspraktijk (zie bijvoorbeeld Biesta, 2007). Een alternatieve manier om meer gecontroleerde onderzoeken te doen, zonder iets te manipuleren aan de lespraktijk, is het volgen van leerlingen. Middels een dergelijke longitudinale opzet kan onderzocht worden hoeveel het gevolgde onderwijs toevoegt aan de rekenprestaties van leerlingen, en of dat samenhangt met bepaalde kenmerken van het onderwijsleerproces, zoals het niveau van vakinhoudelijke en/of vakdidactische kennis van de leerkracht die ze op dat moment hebben of specifieke instructie en/of werkvormen. Een tweede alternatieve manier om robuustere conclusies te trekken is door nascholing van leerkrachten (teams) op specifieke vaardigheden, om vervolgens te analyseren in hoeverre dat invloed heeft (a) op de leerkrachtvaardigheden en -gedrag en (b) op de leerlingvorderingen. Vervolgens kan ook nog onderzocht worden of de mate van leerkrachtvaardigheden/-gedrag samenhangt met de leervorderingen (een dosis-effect analyse). Essentieel bij een dergelijke opzet is een voor- en natoets en een representatieve controlegroep. Dergelijke onderzoeksprogramma's zijn recent uitgevoerd met betrekking tot differentiëren (Van de Weijer-Bergsma et al., 2016) en formatief toetsen (Doolaard, 2013a, 2013b; Veldhuis & van den Heuvel-Panhuizen, 2015; Verbeek, Ledoux, & Glaudé, 2012; Visscher, 2015). Ook heeft vragenlijstsonderzoek naar instructiepraktijken, zoals in beide peilingen is gedaan, een aantal beperkingen (Porter, 2002). Vragenlijsten zijn beperkt tot wat de onderzoekers a priori hebben besloten te bevragen, zelfrapportage heeft een aantal inherente beperkingen zoals sociale wenselijkheid en verkeerde inschatting van de feitelijke praktijk en ze zijn beperkt in de mate van complexiteit van instructiepraktijken die gemeten kan worden. Porter (2002) beargumenteert dat vragenlijsten over de inhoud van instructie valide en daarmee de beste optie kunnen zijn, maar niet als het gaat om de kwaliteit van de instructie. Leslogboeken of lesobservaties kunnen een meer valide alternatief zijn (Heidelberg, Van den Bergh, Kuhlemeier, & Hoeks, 1993) maar zijn lastiger op grote schaal zoals in een peilingsonderzoek te implementeren. Voor het volgende peilingsonderzoek rekenen-wiskunde einde (speciaal) basisonder-

wijs, dat in 2019 zal plaatsvinden³, doen wij in het originele rapport (Hickendorff et al., 2017) aanbevelingen welke factoren van het onderwijsleerproces meegenomen moeten worden en op welke manier zij gemeten kunnen worden.

Tot slot zien we een discrepantie in de focus van de onderzoeksliteratuur en door praktijkexperts genoemde thema's. Op grond daarvan bevelen wij nader onderzoek op in ieder geval de volgende thema's aan:

- De invloed van de vakdidactische kennis van de leerkracht op de kwaliteit van de rekenlessen en de rekenprestaties van leerlingen.
- Het effect van nascholing van leerkrachten in hun vakinhoudelijke en/of vakdidactische kennis op het gebied van rekenen-wiskunde.
- De rol van de rekencoördinator en rekenbeleid/-visie van de school meer in het algemeen.
- De verschillen tussen rekenmethoden in onder andere leerstofaanbod en niveaudifferentiatie, en het gebruik van rekenmethoden in rekenles.

Literatuur

- Biesta, G. (2007). Why "what works" won't work: evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57(1), 1–22. <http://doi.org/10.1111/j.1741-5446.2006.00241.x>
- Doolaard, S. (2013a). *Effecten van het trainings- en begeleidingstraject "Streef: Geruik maken van opbrengsten"*. Groningen: GION/RUG.
- Doolaard, S. (2013b). *Het streven naar kwaliteit in scholen voor primair onderwijs*. Groningen: GION/RUG.
- Expertgroep doorlopende leerlijnen. (2008). *Over de drempels met rekenen. Consolideren, onderhouden, gebruiken en verdiepen. Onderdeel van de eindrapportage van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen*. Enschede: Expertgroep doorlopende leerlijnen taal en rekenen.
- Gravemeijer, K., & Kirschner, P. A. (2007). Naar meer evidence-based onderwijs? *Pedagogische Studiën*, 2007(84), 463–472.
- Grissom, J. A., Loeb, S., & Master, B. (2013). Effective Instructional Time Use for School Leaders. *Educational Researcher*, 42(8), 433–444. <http://doi.org/10.3102/0013189X13510020>
- Heck, R. H. (2000). Examining the Impact of School Quality on School Outcomes and Improvement: A Value-Added Approach. *Educational Administration Quarterly*, 36(4), 513–552. <http://doi.org/10.1177/00131610021969092>
- Heidelberg, J., Van den Bergh, H., Kuhlemeier, H., & Hoeks, J. (1993). Tijdsbesteding van docenten in het voortgezet onderwijs op verschillende manieren gemeten: een validatiestudie. *Tijdschrift Voor Onderwijsresearch*, 18, 306–316.
- Hemker, B. (2018). *Werkwijze Meting Taal en Rekenen 2017*. Arnhem: Cito. Retrieved from http://www.cito.nl/onderzoek-en-wetenschap/deelname_nat_onderzoek/ppon/jaarlijks_peilingsonderzoek
- Hemker, B., & Van Weerden, J. (2015). *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 van het basisonderwijs in 2014*. Arnhem: Cito. Retrieved from http://www.cito.nl/onderzoek-en-wetenschap/deelname_nat_onderzoek/ppon/jaarlijks_peilingsonderzoek/jaarlijks_peilingsonderzoek_zevende_meting
- Hickendorff, M., Mostert, T. M. M., Van Dijk, C. J., Jansen, L. L. M., Van der Zee, L. L., & Auer, M. F. F. (2017). *Rekenen op de basisschool. Review van de samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en de rekenwiskunde-prestaties van basisschoolleerlingen*. Universiteit Leiden. Retrieved from <https://www.nro.nl/wp-content/uploads/2017/12/Rekenen-op-de-basisschool-review-405-17-920.pdf>
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second Handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age Publishing.
- Inspectie van het Onderwijs. (2008). Basisvaardigheden rekenen- wiskunde in het basisonderwijs. Retrieved from <http://www.onderwijsinspectie.nl/binaries/content/assets/publicaties/2008/Basisvaardigheden+rekenen-wiskunde+in+het+basisonderwijs.pdf>
- Inspectie van het Onderwijs. (2018). *Onderwijsverslag: De Staat van het Onderwijs 2016/2017*. Utrecht: Inspectie van het Onderwijs.
- Koninklijke Nederlandse Akademie voor Wetenschappen. (2009). *Rekenonderwijs op de basisschool. Analyse en sleutels tot verbetering*. Amsterdam: KNAW.
- Marzano, R. J. (2014). *Wat werkt op school*. Rotterdam: Bazalt Educatieve Uitgaven.
- Meelissen, M., & Punter, A. (2016). *Twintig jaar TIMSS. Ontwikkelingen in leerlingprestaties in de exacte vakken in het basisonderwijs 1995-2015*. Enschede: Universiteit Twente.
- Porter, A. C. (2002). Measuring the Content of Instruction: Uses in Research and Practice. *Educational Researcher*, 31(7), 3–14. <http://doi.org/10.3102/0013189X031007003>
- Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). *Review and meta-analysis of school and teaching effectiveness*. Enschede: Universiteit Twente.
- Scheltens, F., Hemker, B., & Vermeulen, J. (2013). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 5*. Arnhem: Cito.

- Taylor, J. E. (2017). *The Effectiveness of Instructional Mathematics Coaching: A Study on How the Implementation of Instructional Mathematics Coaches in Elementary Schools Impacts Student Achievement and Promotes Teacher Self-Efficacy*. Joliet, Illinois: ProQuest Dissertations Publishing.
- Vale, C., Davies, A., Weaven, M., Hooley, N., Davidson, K., & Loton, D. (2010). Leadership to improve mathematics outcomes in low SES schools and school networks. *Mathematics Teacher Education and Development*, 12(2), 47–71.
- Van de Weijer-Bergsma, E., Van Luit, H., Prast, E., Kroesbergen, E., Kaskens, J., Compagnie-Rietberg, C., ... Logtenberg, H. (2016). *Differentiëren in het rekenonderwijs: Hoe doe je dat in de praktijk?* Doetinchem: Graviant scientific & educational books.
- Van Zanten, M., & van den Heuvel-Panhuizen, M. (2014). Freedom of Design: The Multiple Faces of Subtraction in Dutch Primary School Textbooks. In Y. Li & G. Lappan (Eds.), *Mathematics Curriculum in School Education* (pp. 231–259). Dordrecht: Springer. <http://doi.org/10.1007/978-94-007-7560-2>
- van Zanten, M., van Graft, M., & van Leeuwen, B. (2017). *Leerplankundige verkenning van TIMSS-trends*. Enschede: SLO, nationaal expertisecentrum leerplanontwikkeling.
- Veldhuis, M., & van den Heuvel-Panhuizen, M. (2015). Supporting primary school teachers to improve their assessment practice in mathematics: Effects on students' learning. *Paper Submitted for Publication*.
- Verbeek, F., Ledoux, G., & Glaudé, M. (2012). *Op weg naar opbrengstgericht leiderschap: evaluatie van het project "Versterken kwaliteit bestuur en management."* Amsterdam: Kohnstamm Instituut.
- Visscher, A. J. (2015). *Over de zin van opbrengstgericht(er) werken in het onderwijs*. Groningen: GION.

Which factors in the educational process are related to mathematics achievement of primary school students? On request of the Inspectorate of Education and NRO (Netherlands Initiative for Education Research) a review study addressing this question was conducted. The focus was on factors that can be influenced.

First we present an overview of the mathematics achievement results of Dutch primary school children using the most recent large-scale educational assessments in fourth and sixth grade and the End of Primary School Tests. These show that the differences in mathematics achievement between students can be explained for only a small part, no more than 10 per cent, by differences in the mathematics education these students received.

Next, results from international and Dutch research have been reviewed to answer the question which factors in the educational process are effective. This yielded a large number of lesson characteristics that were positively related to mathematics achievement, such as a diversity of instructional interventions, using technological and non-technological aids, and formative assessment. It is striking that no robust results were found regarding the effects of opportunity-to-learn conceptualized as mathematics textbook. Furthermore, we identified teacher characteristics, such as mathematical content knowledge, and student characteristics, such as mathematics anxiety, that are related to mathematics achievement. Few results were found regarding the role of classroom characteristics and school characteristics. All in all, the teacher seems to have a crucial role in implementing good mathematics education.

Noten

- 1 Dit onderzoek is (mede) gefinancierd door het Nationaal Regieorgaan Onderwijsonderzoek, onder projectnummer 405-17-920.
- 2 In het verslag van PPO-2011 werden meer significante effecten gerapporteerd (Scheltens et al., 2013), maar dit zou kunnen komen doordat daar (ten onrechte) geen rekening werd gehouden met de multilevelstructuur van de data.
- 3 Zie <https://www.onderwijsinspectie.nl/onderwerpen/peil-onderwijs> en <https://www.nro.nl/nro-projecten-vinden/?projectid=405-18-921-peiling-rekenen-wiskunde-einde-speciaal-basisonderwijs>.