Rekenen op de basisschool

Review van de samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en de rekenwiskundeprestaties van basisschoolleerlingen

M. Hickendorff, T.M.M. Mostert, C.J. van Dijk, L.L.M. Jansen, L.L. van der Zee, & M.F. Fagginger Auer

Universiteit Leiden, Instituut Pedagogische Wetenschappen, afdeling Onderwijswetenschappen

Voorwoord

Voor u ligt een rapport dat is opgesteld in opdracht van het Nationaal Regieorgaan Onderwijs-onderzoek op verzoek van de Inspectie van het Onderwijs, in aanloop naar de nieuwe ronde van Peil.Onderwijs rekenen-wiskunde. Peilingen van het onderwijsniveau brengen de leeropbrengsten in kaart, zoals rekenwiskundeprestaties van leerlingen aan het einde van het basisonderwijs. In de voorgaande peilingen rekenen-wiskunde, de zogenaamde PPONs (Periodieke Peilingen van het Onderwijsniveau), is het onderwijsaanbod in kaart gebracht. De samenhang tussen dit onderwijsaanbod en de leeropbrengsten is echter nauwelijks onderzocht. Voor de nieuwe ronde van Peil.Onderwijs heeft de Inspectie van het Onderwijs als doel gesteld om meer inzicht te krijgen in de relatie tussen de leeropbrengsten en 'beïnvloedbare factoren van het onderwijsleerproces'. De voor u liggende review is geïnitieerd om richting geven aan dat proces, door in kaart te brengen wat er vanuit wetenschappelijk onderzoek bekend is over deze relatie. Op die manier kunnen aanbevelingen worden gedaan over welke factoren in het nieuwe peilingsonderzoek onderzocht moeten worden, en hoe dat het beste kan worden gedaan. Bovendien geeft het een beeld van naar welke gebieden er nog weinig onderzoek is gedaan, zodat ook aanbevelingen kunnen worden gedaan voor onderzoeksthema's binnen het rekenwiskundeonderwijs die het peilingsonderzoek overstijgen.

Deze review is uitgevoerd door een team van onderzoekers verbonden aan de Universiteit Leiden. Hoofdonderzoekers zijn dr. Marian Hickendorff (projectleider), universitair docent Onderwijswetenschappen van het Instituut Pedagogische Wetenschappen, en dr. Marije Fagginger Auer, werkzaam als psychometrisch deskundige en projectleider bij de Vereniging Hogescholen op het project *10voordeleraar*, en voor de huidige review aangesteld als onderzoeker bij Onderwijswetenschappen van het Instituut Pedagogische Wetenschappen. Dr. Kees van Putten, universitair docent Methoden en Statistiek van het Instituut Psychologie, had een adviserende rol, met name in het opstarten van het onderzoek. De rest van het onderzoeksteam bestond uit de onderzoeksmedewerkers Hannah van Dijk, Lisa Jansen, Terry Mostert en Lisanne van der Zee.

Eén onderdeel van de huidige review was een nadere analyse op bestaande onderzoeksgegevens van de meest recente peiling rekenen-wiskunde, PPON-2011. Speciale dank gaat uit naar Bas Hemker en Hans Kuhlemeier voor het uitvoeren en rapporteren van deze analyses. Verder gaat onze dank uit naar Marleen van der Lubbe, Heleen Vinckemöller en Bauke Milo van de Inspectie van het Onderwijs en Anneke Noteboom van de SLO voor het meelezen en meedenken bij de review en het mee-organiseren van de veldraadpleging. Tot slot gaat onze dank uit naar alle experts, zowel wetenschappelijk als meer praktijkgericht, die we hebben geraadpleegd. Het project was een unieke gelegenheid om mensen die op vele verschillende manieren bij het rekenwiskundeonderwijs betrokken zijn te spreken en ook met elkaar in gesprek te laten treden. Wij hopen dat deze koers van samenwerking tussen onderzoek, beleid en praktijk voortzetting krijgt.

Het is een omvangrijk rapport geworden en een leeswijzer is dan ook op zijn plaats. De uitgebreide samenvatting beschrijft de achtergrond, aanpak, bevindingen, conclusies en aanbevelingen. 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 eigenlijke rapport:

- de inleiding in Hoofdstuk 1;
- de methode van literatuuronderzoek in Hoofdstuk 2;
- de resultaten van de review van de internationale literatuur in Hoofdstuk 3;
- de resultaten van de onderzoeken met Nederlandse basisschoolleerlingen in Hoofdstuk 4;
- de resultaten van nadere analyses op de meest recente peilingsonderzoeken PPON-2011 en TIMSS-2015 in Hoofdstuk 5;
- de resultaten van de veldraadpleging in Hoofdstuk 6;
- de conclusies en aanbevelingen in Hoofdstuk 7.

Wij wensen u veel leesplezier!

Inhoudsopgave

Voorwoord 3
Inhoudsopgave 5
Samenvatting 7
I. Achtergrond en doelstellingen 7
II. Aanpak 9
III. Resultaten: meta-analyses en reviews 10
IV. Resultaten: onderzoeken met Nederlandse basisschoolleerlingen 16
V. Resultaten: nadere analyses peilingen 19
VI. Resultaten: veldraadpleging 21
VII. Conclusies en aanbevelingen 22
1 Inleiding 26
2 Onderzoeksmethode 29
2.1 Literatuuronderzoek 29
2.2 Expertraadpleging 32
3 Resultaten internationale meta-analyses en reviews 33
3.1 Factoren op het niveau van de les 33
3.2 Factoren op het niveau van de leerkracht 61
3.3 Factoren op het niveau van de klas 66
3.4 Factoren op het niveau van de school 66
3.5 Factoren op het niveau van de leerling 67
4 Onderzoeken met Nederlandse basisschoolleerlingen 70
4.1 Reken(verbeter)programma's 70
4.2 Rekenspellen en oefenplatforms op de computer 71
4.3 Trainingen die aansturen op het noteren van uitwerkingsstappen 76
4.4 Professionaliseringsactiviteiten 77
4.5 Schoolvisie: Daltononderwijs 82
4.6 Leerlingovertuigingen 82
5 Resultaten nadere analyses recente peilingen 84
5.1 Stand van zaken rekenniveau Nederlandse basisschoolleerlingen. 84
5.2 Nadere analyses PPON-2011 86
5.3 Nadere analyses TIMSS-2015 90
5.4 Conclusies nadere analyses PPON-2015 en TIMSS-2011 91
6 Resultaten veldraadpleging 92

7 Conclusies en aanbevelingen 95
7.1 Samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en rekenprestaties. .95
7.2 Aanbevelingen voor toekomstig peilingsonderzoek 97
7.3 Beperkingen 100
7.4 Aanbevelingen voor overig nader onderzoek. 101
Referenties 102
Bijlage: Leerkrachtvariabelen in nadere analyses PPON-2011 109

Samenvatting

I. Achtergrond en doelstellingen

Peilingen van het onderwijsniveau rapporteren over leeropbrengsten, zoals rekenwiskundeprestaties. Een doel van de nieuwe ronde van Peil.Onderwijs[1] is meer inzicht te krijgen in de relatie tussen deze opbrengsten en beïnvloedbare factoren van het onderwijsleerproces. De huidige review is geïnitieerd om richting geven aan dat proces.

Het voornaamste doel van de huidige review is een overzicht te geven van de wetenschappelijke bevindingen over de samenhang tussen beïnvloedbare factoren uit het onderwijsleerproces en de rekenwiskundeprestaties van leerlingen (aan het einde) van de basisschool. Onder *rekenwiskundeprestaties* 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 rekenprestaties. Het *onderwijsleerproces* hebben we als volgt gedefinieerd: 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.

Om een indicatie te krijgen van factoren binnen het onderwijsleerproces die 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) waarin zij in opdracht van de OECD en de Nederlandse en Duitse Ministeries van Onderwijs een overzicht hebben geconstrueerd van factoren die van invloed zijn op schooleffectiviteit. Dit 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 voor 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 factorenschema (zie Figuur S1).

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), van de leerkracht (zoals kennis, vaardigheden, en attituden), van de klas (zoals een taakgerichte werksfeer), van de school (zoals de rekenvisie en de aanwezigheid van een rekencoördinator), en van de leerling (zoals motivatie en rekenangst). Uiteraard is dit een overmatige versimpeling van de werkelijkheid, 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 ('knoppen waar je aan zou kunnen draaien'), waarbij we streefden 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.

Naast het hoofddoel zijn er nog twee aanvullende doelen. Het eerste aanvullende doel is het formuleren van hypothesen met betrekking tot kenmerken van het onderwijsleerproces die gepeild zouden moeten worden in toekomstig nationaal peilingsonderzoek. Het tweede aanvullende doel is het identificeren van blinde vlekken: wat zijn aspecten binnen dit domein die nog niet zijn onderzocht, maar waarvoor dat wel gewenst is?

[1] Zie ook op de site van de Inspectie van het Onderwijs (https://www.onderwijsinspectie.nl/onderwerpen/peil-onderwijs) en de programmapagina van NRO (https://www.nro.nl/onderzoeksprojecten/peil-onderwijs/).

Figuur S1: Schema van beïnvloedbare factoren in het onderwijsleerproces die mogelijk samenhangen met de rekenprestaties van leerlingen.

niveau	categorie factor(en)		factor(en)
les	*opportunity to learn*	leerstofaanbod	leerstofinhoud
		tijd	instructietijd, oefentijd, *time-on-task*
	instructie- en werkvormen		directe instructie, begeleide/constructivistsche instructie, probleemgestuurd leren, onderzoekend leren, samenwerkend leren, voorbeeldgestuurd leren, zelfregulerend leren
	hulpmiddelen	technologisch	computerondersteuning, *games*
		niet-technologisch	concreet materiaal zoals een rekenrek
	toetsing	summatief	eindtoetsen
		formatief	feedback, monitoren eigen prestaties
	overige leskenmerken		rekenmethode, differentiatie, vakintegratie, huiswerk
leerkracht	kennis en vaardigheden		vakkennis, pedagogische vakkennis, klassenmanagement, helderheid uitleg
	professionele ontwikkeling		professionalisering, professionele leeromgevingen
	overtuigingen		prestatieverwachtingen, visie op (reken)didactiek, professionele motivatie, rekenangst
klas	sfeer		werkklimaat, veiligheid, relaties, ondersteuning, regels, prestatiegerichtheid
school	rekenbeleid		rekencoördinator, belang van rekenen binnen de school
	evaluatie van het (reken)leerproces		monitoren van prestaties, inzicht (reken)behoeftes leerlingen, opbrengstgericht werken
	klimaat		structuur
leerling	overtuigingen		motivatie, *self-efficacy*, attitude, rekenangst

II. Aanpak

Onze aanpak bestond uit drie elementen. Het eerste en grootste element betreft een review van de wetenschappelijke literatuur naar de relatie 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 is de internationale literatuur, waar we ons hebben beperkt tot meta-analyses en reviews. Uiteraard is dat een flinke beperking in de verwerkte literatuur, maar tegelijkertijd leveren meta-analyses en reviews wel robuustere resultaten dan losse onderzoeken. De tweede categorie is onderzoek met Nederlandse basisschoolleerlingen. Omdat dit onderzoek in principe het meest relevant is voor de hoofdvraag en het aantal studies beperkt en daarmee behapbaar was, behandelen we hier wel individuele onderzoeken. De literatuur is via een systematische zoekstrategie verzameld, aangevuld met suggesties van wetenschappelijke experts.

Het tweede element bestaat uit 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 acht, uitgevoerd in 2011 (Scheltens, Hemker, & Vermeulen, 2013) en het internationale onderzoek Trends in Mathematics and Science Study (TIMSS) bij leerlingen uit groep zes, uitgevoerd in 2015 (Meelissen & Punter, 2016). Voor PPON is bij de onderzoekers van Cito de opdracht uitgezet nadere analyses uit te voeren naar de samenhang tussen enerzijds het onderwijsaanbod zoals gemeten via een vragenlijst aan de leerkrachten, en anderzijds de prestaties van leerlingen op verschillende (sub)schalen van rekenen-wiskunde. Voor TIMSS hebben onderzoekers van de Universiteit Twente, die verantwoordelijk waren voor het Nederlandse aandeel in TIMSS-2015, nadere analyses uitgevoerd in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap (Rebber, Van den Berg, & Meelissen, 2017). Voor de huidige review waren twee onderzoeksvragen relevant omdat die over factor(en) van het onderwijsleerproces gingen. Deze vragen betreffen de samenhang van leerprestaties met (a) de percepties van leerkrachten over hun didactische vaardigheden en het leerklimaat op school en (b) het uitgevoerde curriculum (percentage onderwezen onderwerpen).

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

III. Resultaten: meta-analyses en reviews

In totaal hebben we 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. In het vervolg bespreken we de effecten van de factoren op de verschillende niveaus uit Figuur S1: kenmerken van de les, leerkracht, klas, school en leerling. Voor elk niveau presenteren we een schematisch overzicht van de onderzochte factoren. De verschillende kleuren geven aan of de factor een positieve (groen), negatieve (rood) of geen (grijs) samenhang had met rekenprestaties. Als de factor wit is, is er geen meta-analyse of review gevonden met de huidige zoekmethode. Tussen haakjes staan de aantallen meta-analyses of reviews die per categorie factoren zijn besproken.

a. Factoren op het niveau van de les

Figuur S2 is een schematische weergave van de effecten van de kenmerken van de les. We bespreken achtereenvolgens de lesinhoud, lesvorm, en inhoud- en vormoverstijgende factoren.

Lesinhoud

Bij de inhoud van de les staat het 'wat' centraal (in tegenstelling tot het 'hoe', dat bij instructie- en werkvormen wordt besproken). Het is aannemelijk dat het *leerstofaanbod* samenhangt met rekenprestaties, maar behalve de bevinding dat de mate van aansluiting tussen inhoud van het curriculum en inhoud van de toets positief samenhangt met schoolprestaties in het algemeen, blijkt deze hypothese moeilijk te onderbouwen met onderzoeksresultaten specifiek voor rekenen[2]. Ook is het waarschijnlijk dat de *tijd* die besteed wordt aan instructie, oefenen, of meer algemeen *time-on-task* samenhangt met rekenprestaties, maar hier zijn geen onderzoeksresultaten specifiek voor rekenen gevonden.

Lesvorm

Wat betreft de vorm van de les onderzochten we ten eerste verschillende *instructie- en werkvormen*. Omdat werkvormen die in de les gebruikt worden vaak nauw aansluiten bij de instructievorm, en in de bestudeerde literatuur het onderscheid vaak niet eenduidig is, bespreken we deze samen. In 15 verschillende meta-analyses zijn de effecten van interventies met bepaalde instructie- en/of werkvormen onderzocht. Instructievormen zijn grofweg te ordenen in meer directe, expliciete instructie door de leerkracht en meer begeleid(end)e constructivistische instructie waarbij de leerling centraal staat. Alle verschillende interventies hadden een positief effect op de rekenprestaties (van voor- naar natoets en/of vergeleken met een controlegroep), maar de effectiviteit van de verschillende instructievormen verschilde onderling niet. Een effectieve instructie helpt dus, waarbij het minder van belang lijkt welk type instructie dit is. Een uitzondering daarbij is dat er in twee meta-analyses expliciete instructie en ontdekkend leren (met of zonder begeleiding) rechtstreeks met elkaar zijn vergeleken. Leerlingen presteerden bij ontdekkend leren zónder begeleiding slechter dan bij expliciete instructie, maar bij ontdekkend leren mét begeleiding beter dan gemiddeld over de andere instructievormen samengenomen. Enige vorm van begeleiding door de leerkracht lijkt dus nuttig.

[2] De nadere analyses op data van TIMSS-2015 (Hoofdstuk 5.3) zijn hiervoor geschikter.

Figuur S2: Resultaten factoren op het niveau van de les.

	categorie factor(en)		factor(en)
lesinhoud (wat?)	*opportunity to learn* (1)	leerstofaanbod (1)	▪ **aansluiting inhoud curriculum met inhoud toetsen**
			▪ *leerstofinhoud*
		tijd (0)	▪ *instructietijd* ▪ *oefentijd* ▪ *'time-on-task'*

lesvorm (hoe?)	instructie- en werkvormen (15)		▪ **interventie met bepaalde instructie-/werkvorm:**
			  ▪ **expliciete instructie, directe instructie**
			  ▪ **meer begeleid(end)e instructie**
			  ▪ **moderne instructievormen (leerling centraal)**
			  ▪ **innovatieve instructievormen**
			  ▪ **extra oefenopgaven**
			  ▪ **leerlingen zelf vraagstukken laten (her)vormen**
			  ▪ **'Cover-Copy-Compare' methode**
			  ▪ **samenwerkend leren (groepswerk, 'peer-tutoring')**
			  ▪ **aanmoedigen tot 'self explanation'**
			  ▪ **stimuleren zelfregulerend leren**
			▪ **expliciete instructie vs. ontdekkend leren**

			▪ directe instructie vs. begeleide/constructivistische instructie ▪ verschillende rekendidactieken (in NL onderzoek)
			▪ *probleemgestuurd leren* ▪ *voorbeeldgestuurd leren*
	hulpmiddelen (12)	technologisch (8)	▪ **computertechnologie, educatieve applicaties** ▪ **exploratieve digitale leeromgevingen** ▪ **gebruik van mobiele apparaten in de klas**
			▪ computerondersteunde leeromgevingen ▪ video-games
		niet-technologisch (4)	▪ **gebruik drie-dimensionale modellen** ▪ **gebruik concept maps**
inhoud- en vorm-overstijgend	toetsing (3)	summatief (0)	▪ *eindtoetsen*
		formatief (3)	▪ **verschillende vormen van formatieve toetsing** ▪ **gebruik (digitale) leerlingvolgsystemen door leerkracht** ▪ **feedback-per-opgave aan leerling**
			▪ *monitoren eigen prestaties door leerlingen*
	overige factoren (6)		▪ **differentiatie: werken in groepen** ▪ **huiswerk**
			▪ vakintegratie rekenen en 'science' ▪ rekenmethode

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties.
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties.
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

Verder is het stimuleren van samenwerkend leren effectief, zowel wanneer een groep heterogene leerlingen samenwerkt aan een opdracht als één leerling de taak van tutor op zich neemt. Ook het aansporen tot *self explanation* (het door de leerling laten verklaren van een concept of nieuwe informatie) is effectief, met name als dat gebeurt via *scaffolding* (training vooraf, of voorstructurering van de verklaringen). Het stimuleren van zelfregulerend leren waarbij leerlingen actief betrokken zijn bij hun eigen leerproces is ook effectief. Ten slotte zijn twee effectieve rekenspecifieke instructiemethoden gevonden: (1) het leerlingen zelf laten (her)formuleren van rekenvraagstukken, en (2) de zogenaamde *Cover-Copy-Compare*-methode waarbij leerlingen een uitgewerkt rekenvraagstuk moeten bekijken en bedekken, dan weer opschrijven en ten slotte vergelijken met het origineel. We hebben geen relevante meta-analyses of reviews gevonden naar de effectiviteit van probleemgestuurd leren of voorbeeldgestuurd leren. Dit alles geeft het beeld dat *een* vorm van interventie in instructie- en/of werkvormen een positief effect heeft op de rekenprestaties. Omdat de 'effectieve' instructie- en/of werkvormen heel divers zijn qua onderliggende principes, maar nauwelijks rechtstreeks met elkaar vergeleken worden, is het moeilijk om vast te stellen wat de meest werkzame bestanddelen zijn, en rijst vooral het beeld dat iets veranderen aan de instructie- of werkvormen helpt.

Een tweede factor van de vorm van de les is het *gebruik van hulpmiddelen*. Dit kunnen technologische of niet-technologische hulpmiddelen zijn. Technologische hulpmiddelen, zoals computerondersteunde instructie, exploratieve digitale leeromgevingen en het gebruik van mobiele apparatuur in de rekenles, hebben in het algemeen allemaal een positief effect op de rekenprestaties. Er zijn aanwijzingen dat met name een kort gebruik een positief effect heeft. Wanneer computerprogramma's ook (adaptieve) instructie aanbieden lijkt dit het meest effectief als daarnaast ook leerkrachtondersteuning plaatsvindt. Over het gebruik van video-games in de rekenles bestaat nog geen consensus. Niet-technologische hulpmiddelen kunnen een groot positief effect hebben. Meer specifiek lijkt het gebruik van driedimensionale materialen een effectief hulpmiddel, vooral als dit met veel instructiebegeleiding gebeurt. Het gebruik van *concept maps* (diagram van verschillende concepten en hun onderlinge relaties) had ook een positief effect op de leerprestaties, waarbij het effect groter was als leerlingen de *concepts maps* zelf construeerden in plaats van enkel bestudeerden.

Inhoud- en vormoverstijgende kenmerken van de les

Als laatste bespreken we een aantal kenmerken van de les die zowel het 'wat' als het 'hoe' raken. De eerste daarvan is het *toepassen van toetsing*. Over het effect van summatieve toetsing in de les (aan het einde van het leerproces, meestal via gestandaardiseerde toetsen) zijn geen meta-analyses of reviews gevonden. Formatieve toetsing, waarbij op de één of andere manier toetsgegevens worden gebruikt om het leerproces te verbeteren, hangt positief samen met leerprestaties in het algemeen, en rekenprestaties specifiek. Twee ontvangers van de feedback kunnen worden geïdentificeerd: de leerkracht en de leerling. Het gebruik van de feedback uit digitale toets- of leerlingvolgsystemen door leerkrachten hangt positief samen met de rekenprestaties van hun leerlingen. Dit is vooral zo wanneer leerkrachten zich hierin professionaliseren, ze frequent feedback krijgen en een koppeling met de benodigde instructie wordt gemaakt. Als leerlingen in digitale toets- of leeromgeving per gemaakte opgave feedback ontvangen heeft dit een positief effect op hun rekenprestaties. Dit is vooral effectief als het uitgebreide feedback met uitleg is, in plaats van alleen feedback of het antwoord goed of fout is en wat het goede antwoord had moeten zijn. Over het effect van het monitoren van de eigen prestaties door leerlingen zijn geen relevante meta-analyses of reviews gevonden.

Andere inhoud- en vormoverstijgende factoren zijn differentiatie, huiswerk, vakintegratie en rekenmethode. Wat betreft *differentiatie* is het werken met niveaugroepen onderzocht. Dit kan een positief effect hebben, vooral wanneer het werken met groepen onderdeel is van een breder programma waarin ook veel aandacht is voor aangepaste instructie en waarin leerkrachten getraind worden. Er blijken geen verschillen tussen het werken met homogene en heterogene niveaugroepen. *Huiswerk* heeft een kleine positieve samenhang met (reken)prestaties. *Vakintegratie* door middel van

het integreren van rekenen en *science* had geen effect op de rekenprestaties. Ten slotte blijkt dat er tussen *rekenmethoden* geen grote verschillen bestaan in rekenprestaties[3].

b. Factoren op het niveau van de leerkracht

Figuur S3 is een schematische weergave van de effecten van de kenmerken van de leerkracht. *Kennis en vaardigheden* van de leerkracht hangen positief samen met de rekenprestaties van de leerlingen. De wat meer algemene leerkrachtvaardigheden die betrekking hebben op het creëren van een productieve leeromgeving in de klas (effectief pedagogisch handelen, klassenmanagement-vaardigheden) hebben een positieve samenhang met rekenprestaties. Daarnaast hebben zowel de vakinhoudelijke kennis over rekenen als de pedagogische vakinhoudelijke kennis (het totaal van vakinhoudelijke en vakdidactische kennis over bijvoorbeeld leerstrategieën en misconcepties van leerlingen bij rekenen) een positieve relatie met de rekenprestaties van leerlingen. Over de leerkrachtvaardigheid om helder uitleg te geven is geen meta-analyse of review gevonden. Het belang van vakinhoudelijke kennis wordt nog ondersteund door het positieve effect van *professionaliseringsactiviteiten* op dit gebied. Daarnaast is er een positief verband tussen het deelnemen aan professionele leergemeenschappen, samenwerkingsverbanden tussen vakgenoten gericht op het delen van informatie en kennis, en de rekenprestaties van leerlingen. Ten slotte bleek dat een specifieke *leerkrachtovertuiging*, (geïnduceerde) stereotype verwachtingen ten nadele van meisjes, ook resulteerde in lagere rekenprestaties van meisjes. Over de overige leerkrachtovertuigingen - prestatieverwachtingen, visie op (reken)didactiek, professionele motivatie en rekenangst van de leerkracht - zijn geen relevante meta-analyses of reviews gevonden.

Figuur S3: Resultaten factoren op het niveau van de leerkracht.

niveau	categorie factor(en)	factor(en)
leerkracht	kennis en vaardigheden (3)	▪ **vakinhoudelijke kennis** ▪ **pedagogische vakinhoudelijke kennis** ▪ **klassenmanagementvaardigheden** ▪ **pedagogisch handelen (creëren productieve leeromgeving)**
		▪ *helderheid uitleg*
	professionele ontwikkeling (2)	▪ **professionalisering (vakinhoudelijke kennis)** ▪ **professionele leergemeenschappen**
	overtuigingen (1)	▪ stereotypering (naar geslacht leerling)
		▪ *prestatieverwachtingen* ▪ *visie op (reken)didactiek* ▪ *professionele motivatie* ▪ *rekenangst*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties.
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties.
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

c. Factoren op het niveau van de klas

Figuur S4 is een schematische weergave van de effecten van de kenmerken van de klas. We hebben geen meta-analyses of reviews gevonden naar de effecten van kenmerken van de klas, zoals werkklimaat, veiligheid, relaties, ondersteuning, regels en prestatiegerichtheid, op de rekenprestaties

[3] Voor de Nederlandse situatie analyseren we verschillen in rekenprestaties tussen rekenmethoden ook nog in de nadere analyses op PPON-2011.

van leerlingen. Een mogelijke verklaring hiervoor is dat dit wellicht niet los te zien is van de leerkracht, die immers een cruciale rol speelt bij de sfeer in de klas. Dat blijkt bijvoorbeeld ook uit de bevinding (bij Leerkrachtfactoren) dat klassenmanagementvaardigheden van leerkrachten invloed hebben op de rekenprestaties.

Figuur S4: Resultaten factoren op het niveau van de klas.

niveau	categorie factor(en)	factor(en)
klas	sfeer (0)	▪ *werkklimaat* ▪ *veiligheid* ▪ *relaties* ▪ *ondersteuning* ▪ *regels* ▪ *prestatiegerichtheid*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties.
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties.
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

d. Factoren op het niveau van de school

Figuur S5 is een schematische weergave van de gevonden effecten van de kenmerken van de school. Eén review ging in op evaluatie van het onderwijs door middel van 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 en schoolklimaat.

Figuur S5: Resultaten factoren op het niveau van de school.

niveau	categorie factor(en)	factor(en)
school	rekenbeleid (0)	▪ *rekencoördinator* ▪ *belang van rekenen binnen de school*
	evaluatie van het (reken)leerproces (1)	▪ *monitoren van prestaties* ▪ *'accountability' –programma's* ▪ *schoolprestatiefeedback* ▪ *inzicht (reken)behoeftes leerlingen* ▪ *opbrengstgericht werken*
	klimaat (0)	▪ *structuur*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

e. Factoren op het niveau van de leerling

Figuur S6 is een schematische weergave van de gevonden effecten van de kenmerken van de leerling, waarbij we enkel leerlingovertuigingen onderzochten. De naïeve *beliefs* die leerlingen vaak hebben over rekenen (zoals dat rekenkennis bestaat uit losse stukjes informatie), externe attributie van resultaten en rekenangst hangen negatief samen met rekenprestaties. Rekenangst lijkt zowel oorzaak

als gevolg van rekenprestaties te zijn. Over de andere overtuigingen – motivatie, *self-efficacy* (vertrouwen in eigen competentie om taak uit te voeren) en attitude ten opzichte van rekenen – zijn geen meta-analyses of reviews gevonden.

Figuur S6: Resultaten factoren op het niveau van de leerling.

niveau	categorie factor(en)	factor(en)
leerling	overtuigingen (6)	▪ rekenangst ▪ externe (vs interne) attributie resultaten ▪ naïeve 'beliefs' over rekenen
		▪ *motivatie* ▪ *self-efficacy* ▪ *attitude*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

IV. Resultaten: onderzoeken met Nederlandse basisschoolleerlingen

Onderzoek met Nederlandse basisschoolleerlingen bespreken we als aparte categorie. Omdat de meeste studies resultaten met betrekking tot meerdere factoren uit het schema van Figuur S1 bevatten is het niet goed mogelijk om de resultaten per factor te bespreken. In plaats daarvan hebben we de studies meer thematisch geordend. Meer informatie over de individuele studies, inclusief de referenties, staat in Hoofdstuk 4.

a. Reken(verbeter)programma's

Twee studies onderzochten de effectiviteit van twee verschillende reken(verbeter)programma's. In de ene studie ging het om een experimenteel programma *Kwaliteitsversterking rekenen en wiskunde* gericht op adaptieve instructie, met elementen van formatieve toetsing, specifieke instructiekenmerken zoals kwaliteit van de instructie, en actief leren. Dit programma was geïmplementeerd via een geplande veranderingsstrategie met een ruime implementatietijd, intensief en duurzaam ondersteund vanuit de directie en externe experts. In de grootschalige effectiviteitsstudie uitgevoerd in groep 3 bleek het programma een positief effect op de rekenprestaties van leerlingen te hebben.

In de andere studie werd de methodiek *Zo leer je kinderen rekenen* onderzocht, gekenmerkt door systematische opbouw, herhaling bij het oefenen, aanbieden van één strategie en groepsgewijze directe instructie. Het zes weken lang dagelijks 75 minuten oefenen volgens deze methodiek had een positief effect op de automatisering van de basisbewerkingen van kinderen uit de bovenbouw van het regulier en speciaal basisonderwijs van wie de meesten een achterstand hadden in de basisvaardigheden.

Samenvattend kan de inzet van doelgerichte reken(verbeter)programma's effectief zijn, als de implementatie heel sterk wordt ondersteund (*Kwaliteitsversterking rekenen en wiskunde*) en/of het programma intensief wordt ingezet (*Zo leer je kinderen rekenen*). Omdat de programma's sterk verschilden qua doelstellingen en uitgangspunten (zoals actief leren in *Kwaliteitsversterking rekenen en wiskunde*, versus groepsgewijze directe instructie in *Zo leer je kinderen rekenen*), en ook qua methodologie van het onderzoek, is het niet eenvoudig om algemene werkzame elementen te identificeren.

b. Rekenspellen en oefenplatforms op de computer

Er zijn verschillende onderzoeken gevonden waarin het spelen van reken(oefen)spellen op de computer is onderzocht. Het spel *Hit the target* had als doel kinderen spelenderwijs te leren dat variabelen met elkaar kunnen samenhangen. Het spelen van dit spel had een positief effect op de prestaties van groep-8-leerlingen op niet-routinematige, puzzelachtige rekenopgaven. Een andere set studies onderzocht de effectiviteit van mini-games gericht op vermenigvuldigen en delen (zowel het opdoen van nieuwe kennis via exploratie als het automatiseren van bestaande rekenfeiten en -strategieën). Het thuis spelen van deze spelletjes, gecombineerd met een nabespreking op school, had een positieve bijdrage aan de vermenigvuldig- en deelvaardigheden van leerlingen uit groep 4 en 5. Een populair rekenoefenprogramma is *Snappet*, waarbij leerlingen rekenopgaven op een tablet oefenen, directe feedback krijgen over de juistheid van hun antwoord en leerlingen en leerkrachten inzicht krijgen in de voortgang. Het oefenen met *Snappet* heeft positieve effecten op de rekenprestaties van leerlingen uit groep 5 (bij groep 4, 6 en 7 zijn de resultaten minder eenduidig). Een ander rekenoefenprogramma is *Rekentuin*: een adaptief online rekenspel met directe feedback over de juistheid van het antwoord, met nadruk op het oefenen van de bewerkingen. Leerlingen uit groep 5-8 die Rekentuin speelden gingen vooruit in hun prestaties op automatisering van de basisbewerkingen, maar er was geen controlegroep om dit tegen af te zetten. Het maakte niet uit hoe hoog de succeskans was bij de opgaven die leerlingen maakten. Ten slotte had het spelen van een

spel op de tablet gericht op het optellen en aftrekken tot 20 weinig effect op hun prestaties van leerlingen uit groep 3.

Samenvattend kunnen zowel speciaal ontworpen, specifieke, computerspellen als algemenere oefenplatforms een positief effect hebben op de rekenprestaties. Hierbij moet wel aangetekend worden dat dit vrijwel altijd een toevoeging was aan het reguliere curriculum, dus het is niet mogelijk om vast te stellen of het beter werkt dan een andere toevoeging.

c. Trainingen die aansturen op het noteren van uitwerkingsstappen

In twee studies is het effect van het stimuleren van het noteren van rekenstappen onderzocht bij leerlingen uit groep 7-8, bij deelopgaven of samengestelde bewerkingsopgaven. Deze studies gaven wel aanwijzingen dat vooral op een school waar men niet gewend is leerlingen te stimuleren hun rekenstappen te noteren winst geboekt kan worden door aandacht te geven aan het noteren van stappen. Er werd echter geen overtuigend bewijs gevonden dat het expliciet stimuleren de rekenstappen op te schrijven of systematische feedback op het gestructureerd noteren van de rekenstappen te geven, meer effect had dan enkel het oefenen met de opgaven.

d. Professionaliseringsactiviteiten

Professionalisering gericht op differentiatie

Een grootschalig nascholingstraject voor leerkrachten gericht op differentiëren heeft nog geen eenduidige effecten op de differentiatievaardigheden van de leerkracht of de rekenresultaten van de leerling laten zien.

Professionalisering gericht op formatief toetsen

Er zijn verschillende studies naar de effectiviteit van trainingsprogramma's gericht op formatief toetsen uitgevoerd. Eén studie liet zien dat het volgen van (drie) professionaliseringsworkshops gericht op formatieve toetspraktijken in de dagelijkse lespraktijk een positief effect had op de rekenprestaties van de leerlingen uit groep 5. Een set van andere studies onderzocht de effecten van interventies gericht op opbrengstgericht werken door leerkrachten. De grootschalige projecten Opbrengstgericht Leiderschap (gericht op het bevorderen van opbrengstgericht werken door schoolbesturen) en de twee Streef-projecten (het eerste gericht op het verhogen van de kennis van leerkrachten over de referentieniveaus en het tweede gericht op het gebruikmaken van Cito-toetsresultaten) lieten weliswaar effecten zien op school- en leerkrachtniveau, maar niet op de leerlingprestaties. Het Focus-project ten slotte bestond uit vier deelstudies naar de effectiviteit van de Focus-interventie gericht op opbrengstgericht werken. Drie van de vier studies (twee in groep 3-8, één in groep 6) lieten een positief effect op de rekenprestaties zien, waarbij teamtrainingen en aanvullende begeleiding en feedback in de klas lijken te helpen. De effecten waren sterker voor leerlingen van lagere sociaaleconomische klasse, en op scholen met relatief veel leerlingen van lagere sociaaleconomische klasse.

Professionalisering gericht op instructievaardigheden leerkracht

In een pilot-onderzoek zijn de effecten van het inzetten van het professionaliseringsinstrument *Lesson Study* bij rekenlessen op de basisschool onderzocht. Het doel van *Lesson Study* is het verbeteren van de instructievaardigheid van de leerkracht, door voorbereiding en nabespreking in teams af te wisselen met uitvoeringsmomenten met observaties. Hoewel de resultaten van de pilot veelbelovend zijn, laten de verzamelde gegevens en analyses geen harde uitspraken over de effectiviteit toe.

Samenvatting

De studies naar professionaliseringsactiviteiten voor leerkrachten gericht op differentiatievaardigheden en instructievaardigheden laten geen overtuigend beeld van de effecten op rekenprestaties zien. Professionaliseringsactiviteiten op het gebied van formatief toetsen kunnen effectief zijn voor de

rekenprestaties van leerlingen, waarbij de intensiviteit van de training en het soort begeleiding een rol spelen.

e. Schoolvisie

Op basis van gegevens van het COOL[5-18] project zijn de verschillen tussen Daltonscholen en traditionele scholen op onder andere rekenprestaties onderzocht. Gecorrigeerd voor verschillen in achtergrondkenmerken van de leerlingpopulaties bleken er nauwelijks verschillen te zijn.

f. Leerlingovertuigingen

Op basis van gegevens van het PRIMA-cohortonderzoek is onder andere de relatie tussen psychosociaal functioneren (motivatie, zelfvertrouwen en welbevinden) van basisschoolleerlingen uit groep 2, 4, 6 en 8 en hun rekenprestaties onderzocht. Zelfvertrouwen en vooral motivatie hadden een positieve samenhang met rekenprestaties.

In twee andere studies stond rekenangst bij jongens en meisjes centraal. Jongens en meisjes verschilden niet van elkaar in rekenangst en ervaren rekencompetentie, maar meisjes hadden wel hogere toetsangst. In de eerste studie leek er voor meisjes een negatieve relatie te zijn tussen rekenangst en rekenprestaties, maar in de tweede studie, waarbij gecontroleerd werd voor ervaren rekencompetentie, bleek er geen relatie te zijn tussen rekenangst en rekenprestaties.

Samenvattend lijken de leerlingovertuigingen zelfvertrouwen en motivatie samen te hangen met rekenprestaties. De negatieve samenhang tussen rekenangst en rekenprestaties loopt mogelijk via de ervaren rekencompetentie.

V. Resultaten: nadere analyses peilingen

a. Nadere analyses PPON-2011 (groep 8)

Onderzoekers van Cito hebben voor de huidige review nadere analyses uitgevoerd op de gegevens van PPON-2011, met als doel gegevens van het onderwijsaanbod (op basis van de leerkrachtvragenlijst) te relateren aan de rekenprestaties van 2532 leerlingen (108 scholen) in verschillende rekenwiskundedomeinen. Als eerste bleek dat slechts een beperkte hoeveelheid van de variantie in rekenprestaties op schoolniveau was, tussen de 4 en 9 procent. Daarmee is het deel van de rekenprestaties dat potentieel te verklaren is door school- en leerkrachtvariabelen klein. Bovendien waren de correlaties tussen de 48 onderscheiden school- en leerkrachtenvariabelen en de verschillende schaalscores slechts in enkele gevallen significant. Rekenmethode lijkt vooral samen te hangen met rekenprestaties bij Basisoperaties, en niet zozeer bij andere onderwerpen[4]. De sociaaleconomische status van de leerlingpopulatie van scholen had bij alle schalen (op één na) een significant effect en is dus een belangrijke variabele om voor te corrigeren.

Rekenprestaties op diverse schalen waren beter bij oudere leerkrachten, leerkrachten die tevredener waren met de steun die leerlingen thuis ontvangen en leerkrachten die de rekenmethode die ze gebruiken 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 ontvangen. Dat wil niet zeggen dat deze twee factoren daadwerkelijk een negatief effect op prestaties hebben, maar slechts dat ze ermee samengaan (zo is meer externe hulp voor leerlingen die slechter presteren niet meer dan logisch).

Al met al hangen slechts enkele van de tientallen variabelen uit de leerkrachtvragenlijst bij een beperkt aantal inhoudsdomeinen samen met rekenprestaties. De gebruikte rekenmethode hing nauwelijks samen met de rekenprestaties (enkel bij de schaal Basisoperaties.

b. Nadere analyses TIMSS-2015 (groep 6)

Onderzoekers van Universiteit Twente hebben in opdracht van het Ministerie van OCW nadere analyses uitgevoerd op de data van TIMSS-2015 (Rebber et al., 2017). Hierbij werden onder andere de rekenprestaties van leerlingen gerelateerd aan twee factoren: de percepties van leerkrachten over hun didactische vaardigheden en het leerklimaat op hun school; en het uitgevoerde curriculum (percentage onderwezen onderwerpen in de toetsopgaven). In TIMSS-2015 lag het percentage van de variantie op schoolniveau rond de 10 procent (internationaal gezien relatief laag, maar vergelijkbaar met PPON-2011). Van de leerkrachtpercepties hing alleen de perceptie van de leerkracht van de prestatiegerichtheid van het schoolklimaat positief samen met de rekenprestaties, de andere factoren (percepties van werkdruk, vertrouwen in didactisch handelen, waargenomen mate van toerusting per inhoudsdomein en frequentie van bijscholing) niet. Verder was de mate waarin elk inhoudsdomein in de toets is behandeld in het curriculum niet gerelateerd aan de rekenprestaties in het algemeen en specifiek per domein.

c. Conclusies nadere analyses peilingen

Zowel in PPON-2011 (groep 8) als TIMSS-2015 (groep 6) ligt van de verschillen in rekenprestaties tussen leerlingen slechts een klein deel (4-10%) op het niveau van de school. Dit is dus ook de maximale hoeveelheid te verklaren variantie door beïnvloedbare en niet-beïnvloedbare kenmerken van het onderwijsleerproces. Rekenmethode (PPON) en behandelde leerstofgebieden (TIMSS) hingen niet of nauwelijks samen met de rekenprestaties. De factoren die wel positief met rekenprestaties

[4] In het verslag van PPON-2011 werden meer significante rekenmethode-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.

samenhingen geven tezamen niet een heel coherent of overtuigend beeld. De in deze peilingen gehanteerde leerkrachtvragenlijsten hebben slechts beperkt informatie opgeleverd over de relatie tussen factoren van het onderwijsleerproces en rekenprestaties van leerlingen.

VI. Resultaten: veldraadpleging

De praktijkexperts noemden als belangrijkste beïnvloedbare factoren van het onderwijsleerproces kenmerken van de leerkracht, en dan vaak de kennis van de leerkracht, gevolgd door (formatieve) toetsing en rekenbeleid van de school. Voor formatieve toetsing en (pedagogisch) vakinhoudelijke kennis van de leerkracht sluiten de resultaten aan bij die uit de internationale en Nederlandse onderzoeken, die laten zien dat deze factoren positief samenhangen met de rekenprestaties. Aangezien voor rekenbeleid van de school geen onderzoeksresultaten beschikbaar waren, verdient het aanbeveling deze blinde vlek in de komende peiling mee te nemen én er nader gericht onderzoek naar te doen.

VII. Conclusies en aanbevelingen

a. Samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en rekenprestaties

Op basis van de analyse van internationale onderzoeksliteratuur (33 meta-analyses en 19 reviews) en onderzoek bij Nederlandse basisschoolleerlingen (26 (deel)studies en nadere analyses op de meest recente peilingen TIMSS-2015 en PPON-2011) blijkt dat er een grote verzameling van beïnvloedbare factoren in het onderwijsleerproces samenhangt met rekenprestaties.

Op het niveau van de *les* zijn dat:

- een verscheidenheid aan interventies gericht op instructie- en werkvormen (waarbij er geen eenduidige werkzame elementen te identificeren zijn behalve dan dat 'iets manipuleren aan instructie- /werkvormen werkt');
- het toepassen van technologische en niet-technologische hulpmiddelen;
- het formatieve gebruik van toetsgegevens om het leerproces te verbeteren;
- het differentiëren in niveaugroepen;
- het maken van huiswerk.

Opvallend is dat er geen robuuste onderzoeksresultaten zijn met betrekking tot leerstofaanbod of rekenmethode: noch in de internationale literatuur, noch in de nadere analyses van PPON-2011 en TIMSS-2015. Hoewel het belang van deze factoren voor de hand ligt (zie ook Hiebert & Grouws, 2007; Van Zanten & van den Heuvel-Panhuizen, 2014) lijkt het moeilijk deze te onderzoeken. Mogelijk komt dit doordat de begrippen heel 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.

Wat betreft de *leerkracht* hangen zowel meer algemene leerkrachtvaardigheden (effectief pedagogisch handelen, klassenmanagementvaardigheden) als rekenspecifieke kennis (vakinhoudelijke kennis en pedagogisch vakinhoudelijke kennis) samen met rekenprestaties. Omdat (pedagogische) vakinhoudelijke 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 (pedagogisch) vakinhoudelijke kennis. Professionaliseringsactiviteiten op het gebied van vakinhhoudelijke kennis, differentiatievaardigheden en formatief toetsen (variërend van formatieve toetspraktijken in de klas tot het stimuleren van opbrengstgericht werken) zijn mogelijk effectief. Over het effect van leerkrachtovertuigingen vonden we weinig literatuur, behalve een negatieve invloed van het induceren van stereotype verwachtingen. Op basis van meer algemene literatuur (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.

Over de kenmerken van de *klas* werden geen relevante meta-analyses en reviews gevonden. Dat geldt ook voor onderzoek met Nederlandse leerlingen of relevante resultaten uit TIMSS of PPON. Omdat klassenkenmerken waarschijnlijk niet los gezien kunnen worden van de leerkracht denken wij niet dat het nodig is hier specifieke hypothesen voor te formuleren.

Over de kenmerken van de *school* is op basis van de internationale literatuur enkel geconstateerd dat evaluatie van het onderwijs in accountability-programma's en in schoolprestatiefeedback geen overtuigende effecten heeft en mogelijk ook negatieve bijeffecten. Onderzoek met Nederlandse leerlingen geeft aanwijzingen dat bij de implementatie van een programma de intensieve en duurzame ondersteuning door de school (vanuit de directie) belangrijk is. 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 (zie ook positieve effecten van professionele leeromgevingen). Deze aspecten van de school verdienen dus nader onderzoek.

Onderzoek naar kenmerken van de *leerling*, ten slotte, liet zien dat de naïeve *beliefs* die leerlingen vaak hebben over rekenen, externe attributie van resultaten en rekenangst negatief samenhangen met rekenprestaties. Rekenangst lijkt zowel oorzaak als gevolg van rekenprestaties te zijn. Over de andere overtuigingen – motivatie, *self-efficacy* 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.

b. Aanbevelingen voor toekomstig peilingsonderzoek

Voordat we ingaan op aanbevelingen over welke beïnvloedbare factoren van het onderwijsleerproces meegenomen zouden moeten worden in peilingsonderzoek om de samenhang met rekenprestaties van leerlingen te analyseren, is het raadzaam eerst nog stil te staan bij de bevinding dat in de meest recente peilingen slechts 4-10 procent van de variantie in rekenprestaties op het niveau van de school lag. Dit is daarmee ook het maximum wat in het komende peilingsonderzoek mogelijk verklaard kan worden door beïnvloedbare factoren in het onderwijsleerproces. Omdat het verder waarschijnlijk is dat een aantal niet-beïnvloedbare factoren, zoals de sociaaleconomische status van de leerlingpopulatie, ook een deel van de schoolvariantie zullen verklaren, moeten we bescheiden zijn in de verwachtingen wat het 'meenemen van factoren in het onderwijsleerproces' zal opleveren.

Daarnaast is duidelijk geworden dat de nadere analyses van zowel PPON-2011 als TIMSS-2015 slechts beperkte informatie hebben opgeleverd over de relatie tussen factoren van het onderwijsleerproces (zoals gemeten middels vragenlijsten) en rekenprestaties van leerlingen. Gegeven de beperkingen van vragenlijstonderzoek naar instructie in het algemeen en de gehanteerde vragenlijsten van PPON-2011 in het bijzonder, raden wij aan om in een volgende peilingsonderzoek (a) meer metingen te doen dichter bij het onderwijsleerproces, bijvoorbeeld middels lesobservaties en/of leslogboeken, met name voor de kenmerken van de les, (b) wanneer vragenlijsten worden gebruikt gerichtere, meer onderscheidende vragen te stellen en (c) gegevens bij alle actoren in het onderwijsleerproces te verzamelen, dat wil zeggen naast van leerkrachten ook van leerlingen en schooldirectie.

Gegevens over de kenmerken van de *les* kunnen het best verzameld worden door lesobservaties of leslogboeken, met name als het vragen over het 'hoe' betreft (zoals de mate waarin het directe-instructiemodel gevolgd wordt). Vragen over het 'wat' kunnen ook met gerichte vragen in een vragenlijst voor de leerkracht beantwoord worden. Wij bevelen aan in ieder geval de volgende kenmerken te onderzoeken:

- de rekenmethode en het gebruik daarvan;
- de onderwijstijd;
- *time on task;*
- het hoe en wat van het gebruik van technologische en niet-technologische hulpmiddelen;
- het volgen van het directe-instructiemodel;
- de mate van ondersteuning bij meer zelfstandige leeractiviteiten als ontdekkend leren of computerprogramma's;
- het gebruik van verschillende vormen van formatieve toetsing;
- de manier van differentiatie;
- de hoeveelheid huiswerk.

Gegevens over de kenmerken van de *leerkracht* kunnen het best verzameld worden door gerichte vragen in een vragenlijst voor de leerkracht, liefst aangevuld met kennistoetsen die de leerkrachten zelf maken. Kanttekening is dat op deze manier enkel kenmerken van de leerkracht van groep 8 meegenomen worden, terwijl de leerlingen al 6-8 jaar rekenonderwijs hebben gekregen. Relevante aspecten zijn in ieder geval:

- vakinhoudelijke kennis rekenen-wiskunde;
- vakdidactische kennis rekenen-wiskunde;
- professionaliseringsactiviteiten op het gebied van rekenen-wiskunde;
- deelname aan leergemeenschappen/teamoverleg rondom rekenonderwijs;
- prestatieverwachtingen;
- mate waarin de rekenmethode bij de leerkracht past.

Gegevens over de kenmerken van de *school* kunnen het best verzameld worden door gerichte vragen in een vragenlijst aan de directeur, eventueel aangevuld met interviews en/of een analyse van de schoolgids. Relevante aspecten zijn in ieder geval:

- rekenvisie en rekenbeleid;
- belang van rekenen;
- mate van opbrengstgericht werken;
- prestatiegerichtheid van de school.

Gegevens over de kenmerken van de *leerling* kunnen het best verzameld worden door gerichte vragen in een vragenlijst aan de leerling. Daarbij is het aan te bevelen op zoek te gaan naar gevalideerde meetinstrumenten, en/of de vragenlijst van TIMSS te gebruiken. Dat laatste heeft als voordeel dat de resultaten van TIMSS (groep 6) en Peil (groep 8) met elkaar te vergelijken zijn.

Relevante aspecten zijn in ieder geval:

- *beliefs* over rekenen;
- motivatie voor rekenen;
- mate van rekenangst;
- attributie van rekenprestaties;
- attitude ten opzichte van rekenen;
- plezier in rekenen;
- zelfvertrouwen in rekenen;
- *self-efficacy* met betrekking tot rekenen.

c. Beperkingen

De huidige review heeft een aantal beperkingen. Met betrekking tot de analyse van (internationale) reviews en meta-analyses betreft dit de smalheid van de maat van rekenprestaties, het ontbreken van een bespreking van individuele studies als er geen meta-analyses of reviews gevonden waren, en de versimpeling van het onderwijsleerproces tot losse factoren. Met betrekking tot de nadere analyses van de meest recente peilingen is het erg belangrijk te realiseren dat samenhang geen causaal verband indiceert: dat 'een knop waaraan te draaien valt' samenhangt met rekenprestaties, betekent nog niet dat 'draaien aan die knop' resulteert in hogere rekenprestaties. Men moet dus heel voorzichtig moet zijn met praktijkaanbevelingen. Ten slotte beperkt de focus op empirische onderzoeken de reikwijdte van het onderzoek doordat niet wordt ingegaan op belangrijke theorieën over leren in het algemeen en rekenwiskundedidactiek in het bijzonder.

d. Aanbevelingen voor nader onderzoek.

Zoals al eerder genoemd beperkt de correlationele aard van peilingsonderzoeken de interpretatie van verbanden tussen factoren in het onderwijsleerproces en rekenprestaties. Om robuustere conclusies te trekken zijn meer gecontroleerde onderzoeken nodig. Eén manier om dit 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 manier om robuustere conclusies te trekken is door nascholing van leerkrachten(teams) op specifieke vaardigheden in te zetten, 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 een natoets en een representatieve controlegroep. Dergelijke onderzoeksprogramma's zijn recent uitgevoerd met betrekking tot formatief toetsen en differentiëren.

Wij bevelen nader onderzoek aan op de volgende thema's: de pedagogisch vakinhoudelijke kennis van de leerkracht, de rol van de rekencoördinator en rekenbeleid/-visie van de school, en de rekenmethoden (inhoud en gebruik door leerkrachten).

1 Inleiding

Peilingen van het onderwijsniveau rapporteren over leeropbrengsten, zoals rekenwiskundeprestaties. Een doel van de nieuwe ronde van Peil.Onderwijs is meer inzicht te krijgen in de relatie tussen deze opbrengsten en (beïnvloedbare) factoren van het onderwijsleerproces. De huidige review is geïnitieerd om richting geven aan dat proces. Het hoofddoel van de huidige review is een weergave van de laatste stand van zaken betreffende de samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en rekenwiskundeprestaties van kinderen aan het (einde van het) basisonderwijs.

Onder rekenwiskundeprestaties verstaan we in principe de leerlingprestaties op alle vier de domeinen van rekenen-wiskunde zoals die zijn beschreven in de referentieniveaus (Expertgroep doorlopende leerlijnen, 2008). Voor het gemak schrijven we in de rest van de review rekenprestaties. Het onderwijsleerproces hebben we als volgt gedefinieerd: 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. Om een indicatie te krijgen van factoren binnen het onderwijsleerproces die van invloed (kunnen) zijn op de rekenprestaties is gebruik gemaakt van schooleffectiviteitsliteratuur, en specifiek de review van (Scheerens et al., 2007) waarin zij in opdracht van de OECD en het Nederlandse en Duitse ministerie van Onderwijs een overzicht hebben geconstrueerd van factoren die van invloed zijn op schooleffectiviteit. Dit 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 (zie bijvoorbeeld Inspectie voor 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 factorenschema, dat staat weergegeven in Figuur 1.1.

Factoren die daarin een rol spelen hebben we geordend aan de hand van het niveau waarop ze spelen. Hierbij wordt onderscheid gemaakt tussen factoren op het niveau van de les (zoals een bepaalde instructie- of werkvorm), van de leerkracht (zoals kennis, vaardigheden, en attituden), van de klas (zoals een taakgerichte werksfeer), van de school (zoals de rekenvisie en de aanwezigheid van een rekencoördinator), en van de leerling (zoals motivatie en rekenangst). Uiteraard is dit een overmatige versimpeling van de werkelijkheid, 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 enkel gericht op de beïnvloedbare factoren ('knoppen waar je aan zou kunnen draaien'), waarbij we streefden 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[5] hebben we buiten beschouwing gelaten, waarmee uiteraard niet de indruk gewekt moet worden dat die niet van invloed zijn op de rekenprestaties.

Naast het hoofddoel, het in kaart brengen van de samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en rekenprestaties van basisschoolleerlingen, zijn er nog twee aanvullende doelen. Het eerste aanvullende doel is het formuleren van hypothesen met betrekking tot kenmerken van het onderwijsleerproces die gepeild zouden moeten worden in toekomstig nationaal peilingsonderzoek, en aanbevelingen hoe deze gepeild kunnen worden. Het tweede aanvullende doel

[5] Eén leerlingkenmerk verdient nadere toelichting: werkgeheugen/executieve functies. Het is overtuigend aangetoond dat dit kenmerk samenhangt met de rekenprestaties van leerlingen (Friso-Van Den Bos, Van Der Ven, Kroesbergen, & Van Luit, 2013) maar de mate van beïnvloedbaarheid staat ter discussie. Hoewel er werkgeheugentrainingen bestaan is het effect op het werkgeheugen gering en beperkt zich dit tot enkel de geoefende (werkgeheugen)taak en is er geen overtuigend bewijs dat er transfer van het effect naar schoolse vaardigheden plaatsvindt (zie bijvoorbeeld Schwaighofer, Fischer, & Bühner, 2015). Daarom hebben we er voor gekozen werkgeheugen/executieve functies bij de niet (of nauwelijks) beïnvloedbare factoren te plaatsen.

is het identificeren van blinde vlekken: wat zijn aspecten binnen dit domein die nog niet zijn onderzocht, maar waarvoor dat wel gewenst is?

Deze doelen worden nagestreefd middels een narratieve review van de wetenschappelijke literatuur, aangevuld met nadere analyses op de meest recente peilingsonderzoeken in Nederland (PPON-2011 bij leerlingen in groep 8, en TIMSS-2015 bij leerlingen in groep 6) en een veldraadpleging van praktijkexperts.

Figuur 1.1: Schema van factoren in het onderwijsleerproces die mogelijk samenhangen met rekenprestaties.

niveau	categorie factor(en)		factor(en)
les	*opportunity to learn*	leerstofaanbod	leerstofinhoud
		tijd	instructietijd, oefentijd, *time-on-task*
	instructie- en werkvormen		directe instructie, begeleide/constructivistsche instructie, probleemgestuurd leren, onderzoekend leren, samenwerkend leren, voorbeeldgestuurd leren, zelfregulerend leren
	hulpmiddelen	technologisch	computerondersteuning, *games*
		niet-technologisch	concreet materiaal zoals een rekenrek
	toetsing	summatief	eindtoetsen
		formatief	feedback, monitoren eigen prestaties
	overige leskenmerken		rekenmethode, differentiatie, vakintegratie, huiswerk
leerkracht	kennis en vaardigheden		vakkennis, pedagogische vakkennis, klassenmanagement, helderheid uitleg
	professionele ontwikkeling		professionalisering, professionele leeromgevingen
	overtuigingen		prestatieverwachtingen, visie op (reken)didactiek, professionele motivatie, rekenangst
klas	sfeer		werkklimaat, veiligheid, relaties, ondersteuning, regels, prestatiegerichtheid
school	rekenbeleid		rekencoördinator, belang van rekenen binnen de school
	evaluatie van het (reken)leerproces		monitoren van prestaties, inzicht (reken)behoeftes leerlingen, opbrengstgericht werken
	klimaat		structuur
leerling	overtuigingen		motivatie, *self-efficacy*, attitude, rekenangst
niet-beïnvloedbare factoren niet of nauwelijks beïnvloedbare kenmerken van de school, de klas, de leerkracht en de leerling, die mogelijk direct of indirect via beïnvloedbare factoren, van invloed zijn op de rekenprestaties			
niveau	**voorbeelden (niet uitputtend)**		
leerkracht	geslacht, leeftijd		
klas	diversiteit klas		
school	schoolgrootte, diversiteit populatie; buurt (zie ook SES leerling)		
leerling	leerproblemen		persoonlijkheid
	geslacht		werkgeheugen/executief functioneren.
	intelligentie		sociaaleconomische status (SES)
	leeftijd		migratieachtergrond

2 Onderzoeksmethode

2.1 Literatuuronderzoek

Om inzicht te krijgen in de huidige stand van zaken binnen wetenschappelijk onderzoek naar de samenhang tussen beïnvloedbare factoren uit het onderwijsleerproces en de rekenwiskundeprestaties van leerlingen (aan het einde) van de basisschool is een literatuuronderzoek uitgevoerd. De resultaten daarvan dragen bij aan het formuleren van hypothesen met betrekking tot kenmerken van het onderwijsleerproces die gepeild zouden moeten worden. Daarnaast geeft het inzicht in welke factoren nog niet zijn onderzocht, maar waarvoor dat wel gewenst is ('blinde vlekken').

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 is de internationale literatuur, waar we ons hebben beperkt tot meta-analyses en reviews waarin de resultaten van meerdere studies geïntegreerd worden. Uiteraard is dat een flinke beperking in de literatuur die meegenomen is, maar tegelijkertijd leveren meta-analyses en reviews wel robuustere resultaten dan individuele onderzoeken. De tweede categorie is onderzoek met Nederlandse basisschoolleerlingen. Omdat dit onderzoek in principe het meest relevant is voor de hoofdvraag en het aantal studies beperkt en daarmee behapbaar was, behandelen we hier wel alle losse onderzoeken. Zoals hieronder in meer detail wordt toegelicht is alle literatuur via een systematische zoekstrategie verzameld, aangevuld met suggesties van wetenschappelijke experts.

2.1.1 Internationale literatuur

Meta-analyses

Meta-analyses zijn kwantitatieve samenvattingen van individuele empirische studies. Middels meta-analyses is het mogelijk om de kwantitatieve resultaten van onderzoeken te vergelijken en te integreren (Lipsey & Wilson, 2001). Om een goede vergelijking te kunnen maken is het belangrijk dat in de studies hetzelfde construct wordt onderzocht en dat de studies ongeveer hetzelfde design hebben, zodat geen appels met peren worden vergeleken. De effecten uit eerdere studies worden samengevat in één maat voor het effect: de effectgrootte (*effect size*, ES). De effectgrootten die vaak worden gebruikt in meta-analyses zijn Cohen's *d* en Hedge's *g* . Deze effectgrootten geven het verschil in gemiddelden tussen de experimentele en controlegroepen weer, waarbij dat verschil is gestandaardiseerd (dat wil zeggen, niet afhankelijk van de specifieke schaal van het meetinstrument of de variantie in de steekproef), en wordt daarom het gestandaardiseerde verschil genoemd. De waarden van de effectgrootte worden in de huidige review geïnterpreteerd aan de hand van de richtlijnen in Tabel 2.1 (Cohen, 1988).

Tabel 2.1: Richtlijnen voor interpretatie van effectgrootten

absolute waarde ES	interpretatie effect
< 0.20	verwaarloosbaar – klein
0.20 – 0.50	klein – matig
0.50 – 0.80	matig – groot
> 0.80	groot

Naast dat het middels meta-analyses mogelijk is om een betrouwbaar beeld te krijgen van het mogelijke effect van een bepaalde factor op basis van een verzameling onderzoeken, hebben meta-analyses nog een ander voordeel. Meta-analyses bieden namelijk ook de mogelijkheid om

effectgrootten met elkaar te vergelijken en zo te onderzoeken waardoor er verschillen tussen de effectgrootten bestaan. De statistische significantie van deze verschillen wordt bepaald aan de hand van de *Q*- of I^2-waarde. In een onderzoek naar het effect van het gebruik van technologische applicaties op rekenprestaties bijvoorbeeld (Cheung & Slavin, 2013), wordt er middels de Q-waarde onderzocht of er een verschil is tussen effectgrootten van studies waarbij technologische applicaties kort (minder dan dertig minuten per week) worden ingezet en de effectgrootten van studies waarbij technologische applicaties lang (meer dan 75 minuten per week) worden ingezet. Zo'n variabele die invloed kan hebben op de effectgrootten, zoals de intensiteit of duur van de interventie in het bovenstaande voorbeeld, wordt ook wel een *moderator* genoemd.

Omdat moderatoranalyse de studies uit de meta-analyses in categorieën verdeelt aan de hand van kenmerken van de studie, kan het voorkomen dat een categorie weinig studies bevat. De resultaten van moderatoronderzoek dienen daarom met voorzichtigheid te worden geïnterpreteerd, met name wanneer de totale meta-analyse weinig studies bevat, maar veel categorieën onderscheidt. Daarnaast is het belangrijk op te merken dat als een moderator samenhangt met de effectgrootten dit enkel een (correlationele) samenhang is en geen causaal verband. Het betreft immers kenmerken van verschillende studies en geen systematisch gemanipuleerde variabelen. Omdat er meestal niet zoveel studies in een meta-analyse zitten speelt ook het risico van *confounding* van moderatoren. Als bijvoorbeeld alle studies met een korte interventie bij rekenen zijn uitgevoerd en alle studies met een lange interventie bij lezen, is het onmogelijk de effecten van interventieduur (kort of lang) en vakgebied (taal of rekenen) uit elkaar te halen.

Ten slotte is publicatie-*bias* nog een relevant punt. Omdat bekend is dat studies zonder significante resultaten moeilijker gepubliceerd te krijgen zijn, is het waarschijnlijk dat gepubliceerde studies een vertekende selectie van alle uitgevoerde studies is. De effect sizes in een meta-analyse zijn daarom mogelijk groter dan zij in werelijkheid zijn. Wanneer in een meta-analyse publicatie-*bias* een rol leek te spelen volgens de auteurs hebben we dit vermeld.

De meta-analyses die zijn gevonden in de geraadpleegde databases, zijn systematisch in kaart gebracht aan de hand van het analysekader:

- Titel, auteurs en jaartal van publicatie;
- Populatie;
- Land(en);
- Inclusiecriteria van studies;
- Aantal studies (in meta-analyse totaal, en specifiek voor het PO), en indien dit aantal niet goed uit de publicatie te halen is in plaats van het aantal studies het aantal effectgrootten;
- Aantal leerlingen;
- Domein (enkel rekenen, of schoolprestaties meer algemeen);
- Factor(en) van het onderwijsleerproces;
- Uitkomstmaat/-maten;
- Uitkomsten: per uitkomstmaat de effectgrootte en het aantal studies;

In de review presenteren we de geincludeerde meta-analyses in overzichtstabellen met de gegevens zoals vermeld in Tabel 2.2.

Tabel 2.2: Gegevens in de overzichtstabellen over de geïncludeerde meta-analyses

kolom	toelichting
studie	Auteurs en jaartal van publicatie en het onderwerp van de meta-analyse.
aantal studies	Aantal studies in meta-analyse totaal, en specifiek voor het PO. ▪ Indien het aantal studies niet goed uit de publicatie te halen was rapporteren we in plaats daarvan het aantal effectgrootten ▪ Indien het aantal studies in het PO niet goed uit de publicatie te halen rapporteren we '? in het PO'.
uitkomstmaat	De uitkomstmaat of uitkomstmaten waarover wordt gerapporteerd, waar mogelijk toegespitst op rekenen.
uitgesplitst naar	Als de moderator-analyse voor een specifieke moderator significant was rapporteren we de effectgroottes voor de verschillende categorieën van de moderatorvariabele (zoals duur van de interventie). Als voor een bepaald niveau of categorie van de moderatorvariabele (bijvoorbeeld 'minder dan 30 minuten per week') de gemiddelde effectgrootte niet significant (*ns*) was hebben we dat er uitdrukkelijk bij vermeld in de overzichtstabel.

Reviews

Naast meta-analyses zijn ook enkele reviews geraadpleegd. Reviews zijn overzichtsartikelen waarin literatuur over een bepaald onderwerp wordt samengevat. Reviews hebben verschillende soorten functies zoals de lezer op de hoogte brengen van gepubliceerde artikelen uit één of meerdere onderzoeksgebieden, beschrijvingen van wat al onderzocht is en waar nog blinde vlekken zijn en wat nog onderzocht dient te worden, zoals de huidige review (Woodward, 1977). Deze blinde vlekken kunnen betrekking hebben op zowel inhoudelijke als methodologische kwesties. Reviews bevatten vaak geen effectgrootten, maar geven wel een geïntegreerd overzicht van de meest recente bevindingen. In de huidige review beschrijven we de meest belangrijke conclusies uit de reviews bij de betreffende factoren. Het analysekader voor reviews was zoveel mogelijk analoog aan dat voor de meta-analyses.

Zoekstrategie

Na de totstandkoming van het factorenschema op basis van de schooleffectiviteitsliteratuur, is de zoektocht naar wetenschappelijke onderzoeksliteratuur in gang gezet, met als doel om systematisch in kaart te brengen welke meta-analyses en reviews op dit moment bestaan. Daarom is onderzocht op welke manier meta-analyses binnen rekenonderzoek vaak worden opgebouwd en welke databases en zoektermen worden gebruikt. Om systematisch een balans op te maken van de verschillende meta-analyses en reviews zijn meerdere databases geraadpleegd: *Web of Science*, *Scopus* en *Ebscohost* (waaronder *Eric* en *Psycinfo*). Daarnaast is ook een specifieke database voor onderzoek naar rekenen aangesproken, *Math Educ*.

De zoektermen waarmee in deze databases is gezocht naar meta-analyses en reviews waren *mathematics/math/maths/arithmetic* en *meta-analysis/review*, waarbij gebruik werd gemaakt van alle mogelijke combinaties. Naast het gebruik van deze zoektermen, zijn er verschillende criteria opgesteld voor de selectie van de artikelen:

- De artikelen zijn gepubliceerd in 2000 of later, zodat de resultaten zijn gebaseerd op de meest recente onderzoeksbevindingen.
- Het onderzoek vond plaats in een Westers of vergelijkbaar onderwijssysteem, zodat een vergelijking met het Nederlandse onderwijssysteem mogelijk is.
- Eén van de uitkomstmaten of de enige uitkomstmaat is rekenprestaties.

- De studies bevatten proefpersonen van de basisschool (*K*-6)[6].
- Het artikel diende geschreven te zijn in het Nederlands of Engels

Omdat meta-analyses en reviews in zijn algemeenheid kwaliteitseisen stellen aan de studies die geïncludeerd worden, zijn er geen kwaliteitseisen gesteld aan de meta-analyses en reviews.

Meta-analyses en reviews werden geëxcludeerd als zij enkel ingingen op leerlingen met speciale leerbehoeften (*Special Educational Needs, at-risk students*) of geen beïnvloedbare factor van het onderwijsleerproces betroffen. Zo is onderzoek naar ouderbetrokkenheid bijvoorbeeld buiten beschouwing gelaten, omdat dit niet vanuit school plaatsvindt en daarmee niet binnen onze definitie van onderwijsleerproces valt.

Vanwege de beperkte tijd van het project is ervoor gekozen om artikelen uit de eerste honderd resultaten, geordend op relevantie, te selecteren. Niet elke database bleek geordend te zijn op relevantie, zoals *Math Educ* (geordend op jaartal). Bij *Math Educ* is er daarom voor gekozen om alle 263 meta-analyses te screenen. Aangezien onder *Ebscohost* zowel *Eric* als *Psycinfo* valt, zijn de eerste 150 resultaten bekeken. De zoekstrategie heeft geleid tot een selectie van 33 meta-analyses en 19 reviews die voldeden aan de inclusiecriteria, dat wil zeggen dat de samenhang tussen ten minste één beïnvloedbare factor van het onderwijsleerproces en rekenprestaties is onderzocht, met ten minste ook resultaten bij leerlingen uit de basisschoolleeftijd.

2.1.2 Onderzoeksmethode review onderzoeken met Nederlandse basisschoolleerlingen.

Voor artikelen over Nederlands onderzoek zijn de databases van *Pedagogische Studiën* en *Volgens Bartjens Ontwikkeling en Onderzoek*, voorheen de Panama-Post, doorzocht. Hierbij zijn de zoektermen 'rekenen', 'rekenonderwijs', 'rekenprestaties' en 'rekenopgave' gebruikt. De onderzoeken die gericht waren op het basisonderwijs en een relatie onderzochten tussen minstens één beïnvloedbare factor in het onderwijsleerproces enerzijds en rekenprestaties anderzijds zijn meegenomen in de review. Dit leverde in totaal 6 bruikbare artikelen. Aangevuld met de suggesties van de experts (zie hieronder) zijn in totaal 26 (deel)studies met Nederlandse basisschoolleerlingen geïncludeerd.

2.2 Expertraadpleging

Om alle relevante internationale en Nederlandse literatuur te vinden werden daarnaast ook experts op het gebied van rekenonderzoek uit Nederland en Vlaanderen geraadpleegd. Door middel van een email werden zij op de hoogte gebracht van de inhoud en het doel van deze review en werd hen gevraagd welke literatuur hiervoor zeker geraadpleegd zou moeten worden. Ook werd expliciet gevraagd of zij relevant onderzoek met Nederlandse basisschoolleerlingen kenden. Op deze manier werden 41 experts benaderd. Zij waren verbonden aan verschillende instellingen (of zelfstandig): Cito, Erasmus Universiteit, Hogeschool van Arnhem en Nijmegen, Hogeschool Inholland, KU Leuven, Open Universiteit, Radboud Universiteit Nijmegen, Rijksuniversiteit Groningen, SLO, TU Eindhoven, Universiteit Amsterdam, Universiteit Gent, Universiteit Maastricht, Universiteit Twente, Freudenthal Instituut en Educatie & Pedagogiek van Universiteit Utrecht en Vrije Universiteit Amsterdam. Van de benaderde experts stuurden 24 een reactie met suggesties (in sommige gevallen ook namens collega's van dezelfde onderzoeksgroep). De door hen aangedragen literatuur werd volgens de bovenstaande criteria voor internationale en Nederlandse literatuur gescreend op bruikbaarheid voor deze review, en wanneer geschikt geïncludeerd in het onderzoek.

Naast het verzoek om literatuursuggesties werd de experts ook gevraagd feedback te gegeven op het factorenschema. Hieruit voortgekomen aanvullingen werden opgenomen in het schema.

[6] In veel meta-analyses werden de resultaten van leerlingen uit K-12 onderzocht. In het Amerikaanse onderwijssysteem zijn dit leerlingen van de kleuterklas (Kindergarten) tot en met klas 12 (het zesde jaar van het voortgezet onderwijs). Groep 8 in Nederland staat gelijk aan klas (grade) 6.

3 Resultaten internationale meta-analyses en reviews

3.1 Factoren op het niveau van de les

De kenmerken van de les die als startpunt zijn gehanteerd staan in Figuur 3.1. In de volgende paragrafen bespreken we de gevonden meta-analyses en reviews.

Figuur 3.1. Startpunt van de factoren op het niveau van de les.

niveau	categorie factor(en)		factor(en)
les	*opportunity to learn*	leerstofaanbod	leerstofinhoud
		tijd	instructietijd, oefentijd, *time-on-task*
	instructie- en werkvormen		directe instructie, begeleide/constructivistsche instructie, probleemgestuurd leren, onderzoekend leren, samenwerkend leren, voorbeeldgestuurd leren, zelfregulerend leren
	hulpmiddelen	technologisch	computerondersteuning, *games*
		niet-technologisch	concreet materiaal zoals een rekenrek
	toetsing	summatief	eindtoetsen
		formatief	feedback, monitoren eigen prestaties
	overige leskenmerken		rekenmethode, differentiatie, vakintegratie, huiswerk

3.1.1 *Opportunity to learn*

De rekenprestaties van leerlingen uit het basisonderwijs worden niet alleen beïnvloed door hoé het rekenen wordt onderwezen, zoals beschreven wordt in de paragraaf 3.1.2 over instructie- en werkvormen, maar ook door wát precies wordt onderwezen; de inhoud van de les. De inhoud van de les wordt vaak beschreven aan de hand van de begrippen als leerstofaanbod en *Opportunity to learn* (OTL). OTL wordt door sommigen gezien als de aansluiting van de inhoud van het aangeboden onderwijs met de inhoud van examens en toetsen (Scheerens, 2016), maar er bestaat variatie in de definities. Hiebert en Grouws (2007) definiëren OTL namelijk als de omstandigheden die leerlingen in staat stellen om tijd te besteden aan onderwijstaken. OTL varieert volgens hen omdat het beïnvloed kan worden door leerkrachtfactoren zoals de nadruk die leerkrachten op onderwerpen leggen, de verwachting en tijd die zij voor onderwijstaken opstellen en het soort vragen dat ze aan de leerlingen stellen. Doordat Hiebert en Grouws (2007) onder OTL ook andere lesfactoren verstaan zoals de helderheid van de instructie, de organisatie van de les en het klassenmanagement, lijkt dit weinig te verschillen van onze definitie van het *onderwijsleerproces*.

Omdat OTL zoals beschreven door Hiebert en Grouws (2007) meer omvat dan de inhoud of lesstofaanbod, past de definitie van Scheerens (2016) voor OTL mogelijk beter bij de beschrijving van OTL als factor. Scheerens (2016) beschrijft dat de aansluiting tussen het aangeboden onderwijs en de inhoud van de toetsen een klein effect heeft op algemene schoolprestaties, maar de huidige literatuuranalyse leverde geen meta-analyse op over het effect van OTL op rekenprestaties specifiek. Een denkbare verklaring voor de afwezigheid van meta-analyses zijn de verschillen in definities van OTL, wat het uitvoeren van een meta-analyse bemoeilijkt. OTL wordt daarnaast ook voor een groot deel bepaald door de nationale referentiekaders voor rekenen, waardoor de invloed moeilijk te bepalen is bij vergelijkend onderzoek binnen één land. Om goed zicht te krijgen op de invloed van OTL zijn internationale peilingsonderzoeken daarom zeer waardevol. De beschrijvingen van de analyses uit de TIMSS geven daarom meer inzicht in de invloed van leerstofaanbod op de rekenprestaties van leerlingen.

3.1.2 Instructie- en werkvormen

Het effect van instructie- en werkvormen is de laatste jaren uitgebreid onderzocht. Het zijn twee aparte concepten die zowel afzonderlijk als samen een effect kunnen hebben op het onderwijsleerproces. Werkvormen die de leerlingen gebruiken in de les sluiten vaak nauw aan op de instructievorm die de leerkracht gebruikt. In de literatuur is er daarom niet altijd een duidelijk onderscheid te maken tussen instructie en werkvorm. De effecten van instructie- en werkvormen zijn daarom niet altijd te scheiden en worden beide toegelicht.

Voordat we ingaan op verschillende meta-analyses en reviews naar instructievorm, ontdekkend leren, *problem-posing* (het door de leerling laten vormen van problemen en vraagstukken), coöperatief leren en *self explanation* (het door de leerling laten verklaren van een concept of nieuwe informatie), bespreken we eerst het rapport van de KNAW (2009) over rekenonderwijs op de basisschool. Een onderdeel daarvan was een review van het beschikbare onderzoek naar de relatie tussen rekendidactiek en rekenvaardigheid, en dan specifiek bij Nederlandse kinderen op de basisschool. De resultaten van in totaal 28 studies waarin een vergelijking tussen instructieprogramma's of curricula worden gemaakt worden besproken en geïntegreerd.

Het bereik van dit soort vergelijkend wetenschappelijk onderzoek blijkt smal, en de resultaten geven dan ook geen eenduidig beeld. De conclusie die de commissie trekt is dan ook dat "*het bestudeerde materiaal geen algemeen wetenschappelijk gefundeerde uitspraken over de relatie tussen rekendidactiek en rekenvaardigheid rechtvaardigt*" (p. 74-75). Verder merken zij op dat er binnen een rekendidactiek grotere verschillen bestaan dan tussen rekendidactieken (wat zou wijzen op een grote rol voor de implementatie van de didactiek). Meer onderwijstijd voor rekenen lijkt tot betere resultaten te leiden, net als experimentele programma's die in kleine groepjes worden geïmplementeerd. Ten slotte lijken zwakkere rekenaars meer gebaat bij een sturende rol van de leerkracht vergeleken met een vrijere vorm van instructie.

Instructievorm

Afgelopen jaren vond een verschuiving plaats van de traditionele klassikale instructie naar een uitgebreid palet aan instructievormen waaruit de leerkracht kan kiezen. Er zijn vier meta-analyses gevonden naar het effect van instructievorm op rekenprestaties (zie Tabel 3.1).

Burns, Codding, Boice en Lukito (2010) hebben een meta-analyse uitgevoerd naar de effecten van verschillende instructievormen voor basisschoolleerlingen met verschillende rekenniveaus. De instructievormen zijn onderzocht door middel van vergelijkingen met alternatieve instructies of door middel van *multiple baseline* designs. Interventies met de nadruk op expliciete instructie en het laten zien hoe een rekenprobleem kon worden opgelost, hadden een groot positief effect op de rekenprestaties van leerlingen die nog geen of weinig kennis hadden en een klein tot matig positief effect op de rekenprestaties van leerlingen die al redelijk bekwaam waren. Interventies met de nadruk op extra oefenopgaven hadden een klein tot matig positief effect op leerlingen met geen tot weinig kennis. Er waren te weinig studies naar de effecten van dit type interventie op leerlingen met al enige bekwaamheid om hier conclusies uit te trekken.

Jacobse en Harskamp (2011) onderzochten het effect van instructie-interventies op rekenprestaties in het basisonderwijs. Alle studies in de meta-analyse hadden een controlegroep of een alternatieve interventie. De alternatieve interventie of de mate van instructie in de controlegroep wordt niet gespecificeerd in het artikel. De meta-analyse toonde geen significant verschil aan tussen directe en indirecte instructie. Zowel directe instructie, waarbij er expliciet verteld wordt aan de leerlingen wat de opdracht en de uitvoering is, als indirecte instructie, waarbij de docent een begeleidende rol aanneemt, hadden een positief effect op de rekenprestaties van de leerlingen. Uitsplitsing naar verschillende onderdelen van het rekencurriculum leverde geen significante verschillen.

In de meta-analyse van Şad, Kış en Demir (2016) is het effect van verschillende moderne instructievormen op rekenprestaties in het basis- en voortgezet onderwijs onderzocht. Zij maakten een onderscheid tussen moderne instructievormen, waarbij de leerling centraal staat en het samen werken en samen leren gestimuleerd wordt, en traditionele instructievormen. De instructievormen die Şad en collega's in hun meta-analyse onderzochten, waren constructivistisch leren, samenwerkend leren, probleemgestuurd leren en de meervoudige intelligentie theorie. Wanneer instructie gebaseerd is op de meervoudige intelligentie theorie wordt er eerst onderzocht hoe de leerling leert. Daarna wordt het instructieproces gevormd zodat het aansluit bij de individuele leerstijl van de leerling. Şad en collega's hebben verschillende onderzoeken met een voor- en nameting gebruikt in hun meta-analyse. In sommige onderzoeken was er ook een controlegroep. Voor alle vier de instructievormen afzonderlijk werd een matig tot groot positief effect op de rekenpresentaties gevonden bij de nameting, maar er was geen significant verschil tussen de verschillende instructievormen. Bovendien golden deze effecten zowel in het basisonderwijs als in het voortgezet onderwijs. Het aantal leerlingen dat deelnam aan het onderzoek had wel invloed op het effect van de instructievorm op de rekenprestaties: hoe minder leerlingen, hoe groter het effect.

Savelsbergh en collega's (2016) onderscheidden vijf verschillende innovatieve instructievormen: onderzoeksgericht, contextgericht, computergericht, samenwerkend leren en buitenschoolse activiteiten. Zij onderzochten de invloed van deze vijf instructievormen op leerhouding en leerprestaties in *science* en rekenen bij leerlingen van basisscholen en in het voortgezet onderwijs. Elk onderzoek had een voor- en nameting en een controlegroep. Er was geen verschil tussen het effect van de verschillende instructievormen op de leerprestaties in *science* en rekenen. Alle vijf de instructievormen hadden een positief effect op de leerprestaties. Het effect van de instructievormen was kleiner voor oudere leerlingen.

Samenvattend blijkt uit deze vier meta-analyses dat een interventie gericht op instructievorm een positief effect heeft op de rekenprestaties, ongeacht het type instructie dat centraal stond. Jacobse en Harskamp (2011) concluderen dat een effectieve instructie gebaseerd is op een systematisch leermodel waarbij de leertaken opeenvolgend zijn en er sprake is van visualisatie van een stappenplan om de rekenopgaven op te lossen. Dit houdt in dat de leerlingen stappen die zij zetten om de rekenopgaven op te lossen opschrijven of deze voor zich zien. Savelsbergh en collega's (2016) benadrukken dat leerkrachten gestimuleerd moeten worden te investeren in het onderhouden van hun instructieaanpak en aandacht moeten geven aan leerdoelen en de aansluiting van de instructie bij de leerlingen, ongeacht het type instructie.

Tabel 3.1: Meta-analyses met betrekking tot instructievorm

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Burns et al. (2010) instructievorm	17 studies (17 in PO)	vloeiendheid in basisbewerkingen	interventie met expliciete instructie: ▪ leerlingen met weinig of geen kennis: ES = 3.10 ▪ leerlingen met al enige bekwaamheid: ES = 1.12 interventie bestaande uit oefenopgaven: ▪ leerlingen met weinig of geen kennis: ES = 1.06	-
Jacobse & Harskamp (2011) instructie-interventies	40 studies (40 in PO)	rekenprestaties	directe instructie: ES = 0.58 indirecte instructie: ES = 0.61	-
Şad, Kış & Demir (2016) moderne instructievormen	47 studies (13 in PO)	rekenprestaties	samenwerkend leren: ES = 0.72 meervoudige intelligentie theorie: ES = 0.98 probleemgestuurd leren: ES = 0.94 constructivistisch leren: ES = 1.29	groepsgrootte: ES = -0.04
Savelsbergh et al. (2016) innovatieve instructievormen	40 studies (?[7] in PO)	algemene leerprestaties	ES = 0.78	leerjaar: ES = -0.10

Ontdekkend leren

Tegenover de traditionele directe instructie staat het ontdekkend leren, waarbij de leerlingen minder geholpen worden bij het leerproces en in plaats daarvan worden gestimuleerd om zelf te ontdekken hoe de oplossing van een probleem gevonden kan worden (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011). Het ontdekkend leren gaat uit van de nieuwsgierigheid van de leerling.

Alfieri en collega's (2011) hebben onderzocht of ontdekkend leren invloed heeft op de leerprestaties. Zij hebben hiervoor twee meta-analyses uitgevoerd (zie Tabel 3.2). In de eerste meta-analyse onderzochten zij het verschil tussen expliciete instructie en ontdekkend leren zonder enige vorm van ondersteuning op basis van studies waarin deze rechtstreeks met elkaar vergeleken werden. Het bleek

[7] Wanneer niet uit de publicatie te halen was hoeveel van de geincludeerde studies in (onder andere) het PO waren uitgevoerd, vermelden we dit door het vraagteken.

dat expliciete instructie tot betere prestaties op rekenen leidt dan ontdekkend leren zonder ondersteuning, zowel bij kinderen (jonger dan 12 jaar), adolescenten (tussen de 12 en 18), als bij volwassen (ouder dan 18). In de tweede meta-analyse gebruikten Alfieri en collega's (2011) studies waarbij ontdekkend leren mét ondersteuning rechtstreeks werd vergeleken met een andere vorm van instructie. Ontdekkend leren met ondersteuning leidde tot betere leerprestaties op het gebied van rekenen dan de andere onderzochte instructievormen, en dit effect was groter voor volwassenen. De andere onderzochte instructievormen werden echter niet uitgesplitst. Omdat de onderzoeksgroepen voornamelijk uit volwassenen bestonden, geven Alfieri en collega's (2011) aan dat dit het effect kan hebben vertekend.

Tabel 3.2: Meta-analyses met betrekking tot ontdekkend leren.

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Alfieri et al. (2011) ontdekkend leren zonder ondersteuning versus expliciete instructie	108 studies (? in PO)	algemene leerprestaties	ES = -0.38	vakgebied: ▪ rekenen/cijfers: ES = -0.16 ▪ *science*: ES = -0.39 ▪ probleemoplossen: ES = -0.48 ▪ verbale/sociale vaardigheden: ES = -0.95 leeftijd: ▪ adolescenten: ES = -0.53 ▪ volwassenen: ES = -0.26
Alfieri et al. (2011) ontdekkend leren met ondersteuning versus andere instructievormen	56 studies (? in PO)	algemene leerprestaties	ES = 0.30	vakgebied: ▪ rekenen/cijfers: ES = 0.29 ▪ computervaardigheden: ES = 0.64 ▪ *science*: ES = 0.11 ▪ motorische vaardigheden: ES = 1.05 ▪ verbale/sociale vaardigheden: ES = 0.58

Problem posing

Het vormen van problemen en vraagstukken is vaak de taak van de docent of de auteurs van methoden. De pedagoog Freire (1974) opperde dat leerlingen zelf opgaven en vraagstukken moeten vormen of hervormen. Hij noemde dit *problem posing*. Het concept van *problem posing* is terug te vinden in de grondslag van verschillende modellen binnen het (reken)onderwijs (Silver, 1994), denk hierbij aan realistisch rekenen en het Montessori-onderwijs.

Rosli, Capraro en Capraro (2014) onderzochten het effect van *problem posing* op het rekenleerproces bij leerlingen uit het basis- en voortgezet onderwijs en studenten van de lerarenopleiding (zie Tabel 3.3). De studies in de meta-analyse vergeleken een interventie in het klaslokaal met een

controlegroep of een voor- met een nameting. Het verschil tussen deze studie-designs is niet geanalyseerd. Het rekenleerproces werd beoordeeld op vier leeruitkomsten: rekenkennis, rekenvaardigheden, rekenbekwaamheid en attitudes en *beliefs* ten opzichte van rekenen. Uit de meta-analyse bleek dat wanneer leerlingen en studenten *problem posing* gebruikten, er een matig tot groot positief effect was op alle vier de leeruitkomsten.

Tabel 3.3: Meta-analyses met betrekking tot *problem posing.*

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Rosli et al. (2014) *problem posing*	14 studies (? in PO)	rekenprestaties	rekenkennis: ES = 1.31 rekenvaardigheden: ES = 0.83 rekenbekwaamheid: ES= 0.77 attitude-*beliefs*: ES = 0.76	-

Coöperatief leren

Coöperatief leren staat in het klaslokaal beter bekend als samenwerkend leren. Het is een populaire term voor verschillende werkvormen waarin het samenwerken van de leerlingen centraal staat. Er is een onderscheid te maken tussen twee categorieën binnen samenwerkend leren. De eerste categorie bevat de werkvormen waarin een groep (heterogene) leerlingen samenwerkt aan een opdracht waarbij er een gezamenlijk leerdoel of eindproduct is. In de tweede categorie neemt een leerling de taak van tutor op zich en bewaakt deze het leerproces van de andere leerlingen. Hierbij blijkt uit onderzoek dat niet alleen de leerling, maar ook de tutor-leerling vooruitgaat in rekenprestaties (Robinson, Schofield, & Steers-Wentzell, 2005).

Een onderzoek dat samenwerkend leren heeft gedefinieerd zoals in de eerste categorie is de meta-analyse van Johnson, Johnson en Stanne (2000). Zij vergeleken de leerprestaties bij samenwerkend leren met competitief en individueel leren. Zij maakten onderscheid tussen verschillende methoden van samenwerkend leren, die met elkaar gemeen hadden dat de leerlingen op een positieve manier afhankelijk van elkaar waren. Bij competitief leren was er sprake van een negatieve afhankelijkheid: de leerlingen werkten alleen, maar de uitkomsten of beloning waren afhankelijk van de andere leerlingen. De leerlingen die individueel leerden, werkten onafhankelijk van elkaar en werden ook onafhankelijk van elkaar beoordeeld. De leerprestaties zijn op verschillende onderwijsgebieden, waaronder rekenen, gemeten, maar worden niet uitgesplitst naar onderwijsgebied. Het gebruik van samenwerkend leren zorgde voor significant betere leerprestaties dan het gebruik van competitief of individueel leren (zie Tabel 3.4). Voornamelijk de werkvorm 'Samen Leren' had een groot positief effect op de leerprestaties. Bij deze werkvorm gaan de leerlingen in tweetallen aan de slag en neemt één leerling de rol van tutor op zich. De docent richt zich op het trainen van sociale vaardigheden bij de samenwerking.

Het onderzoek naar samenwerkend leren van Rohrbeck, Ginsburg-Block, Fantuzzo en Miller (2003) definieerde samenwerkend leren zoals in de tweede categorie: zij spreken van *peer-assisted learning*. De studies in deze meta-analyse waren allemaal gericht op interventies met nameting op de basisschool waarbij gebruik werd gemaakt van *peer tutoring*. De tutor kon zowel een oudere leerling zijn of een leerling uit dezelfde klas. Het gebruik van *peer tutoring* resulteerde in een klein positief effect op de algehele leerprestaties (zie Tabel 3.4). Ook binnen de studies die alleen gericht waren op rekenprestaties werd er een klein positief effect gevonden. Daarnaast waren de leerprestaties beter wanneer de tutor en de 'leerling' hetzelfde geslacht hadden en als de leerlingen een hoge mate van

autonomie voelden. Autonomie werd hierbij gemeten op vijf componenten, namelijk: (1) doel opstellen, (2) evaluatie, (3) monitoren, (4) beloning selecteren, (5) beloning registreren. Er was sprake van een hoge autonomie als de leerling zelf invloed had op meer dan drie componenten. Als er sprake was van individuele evaluatie van de leerling waren de leerprestaties ook hoger dan wanneer er sprake was van groepsevaluatie.

Met een meta-analyse onderzochten Capar en Tarim (2015) het effect van verschillende coöperatieve werkvormen op rekenprestaties (zie Tabel 3.4). In deze meta-analyse zijn studies meegenomen waarbij coöperatieve werkvormen vergeleken werden met traditionele werkvormen. Het gebruik van coöperatieve werkvormen had een matig positief effect. De meest effectieve coöperatieve werkvormen waren 'Samen Leren', het gebruik van ongestructureerde samenwerking, en het gebruik van werkgroepjes van leerlingen van verschillende leerniveaus. Zowel op basisscholen als in het voortgezet en hoger onderwijs had het gebruik van coöperatieve werkvormen een positief effect op de rekenprestaties.

Shenderovich, Thurston en Miller (2015) vonden dat het leeftijdsverschil tussen de leerling en de tutor-leerling geen effect had op rekenprestaties (zie Tabel 3.4). Er waren in deze meta-analyse echter maar twee studies meegenomen gericht op rekenprestaties.

In de meta-analyse van Leung (2015) werd een matig positief effect van *peer tutoring* op de rekenprestaties gevonden (zie Tabel 3.4). Verder bleek dat het rekenonderdeel dat aangeleerd werd door de tutor invloed had op het effect: gericht op enkel berekeningen had *peer tutoring* een groot positief effect, gericht op berekeningen en conceptuele kennis een matig positief effect, en gericht op enkel conceptuele kennis had het geen effect. Met betrekking tot algehele leerprestaties vond Leung (2015) een groter effect bij gestructureerde *tutoring* dan bij ongestructureerde *tutoring*, maar beide soorten hebben een matig positief effect. Het effect van *peer tutoring* verschilde niet tussen verschillende schoolvakken.

Samenvattend blijkt uit de meta-analyses dat samenwerkend leren een matig tot groot positief effect heeft op rekenprestaties op basisscholen. Zowel als werkvorm waarin een groep (heterogene) leerlingen samenwerkt aan een opdracht waarbij er een gezamenlijk leerdoel of eindproduct is, als in de werkvorm waarbij een leerling de taak van tutor op zich neemt. In twee van de vijf meta-analyses was 'Samen Leren' een van de meest effectieve werkvormen, zowel in vergelijking met andere coöperatieve werkvormen als met andere werkvormen.

Tabel 3.4: Meta-analyses met betrekking tot coöperatief leren.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Johnson et al. (2000) samenwerkend leren versus competitief en individueel leren	158 studies (73 in PO)	algemene leerprestaties	-	*constructive/academic controversy*: ▪ vs. comp.: ES = 0.67 ▪ vs. ind.: ES = 0.91 *cooperative integrated reading and composition*: ▪ vs. comp.: ES = 0.18 ▪ vs. ind.: ES = 0.18 *group investigation*: ▪ vs. comp.: ES = 0.37 ▪ vs. ind.: ES = 0.62 *Jigsaw*: ▪ vs. comp.: ES = 0.29 ▪ vs. ind.: ES = 0.13 *learning together*: ▪ vs. comp.: ES = 0.85 ▪ vs. ind: ES = 1.04 *student teams achievement divisions*: ▪ vs. comp.: ES = 0.51 ▪ vs. ind.: ES = 0.29 *teams-games-tournaments*: ▪ vs. comp.: ES = 0.48 ▪ vs. ind.: ES = 0.58 *team assisted individualization*: ▪ vs. comp.: ES = 0.25 ▪ vs. ind.: ES = 0.33
Rohrbeck et al. (2003) *peer-assisted learning*	81 studies (81 in PO)	algemene leerprestaties	ES = 0.33	geslacht tutor en leerling: ▪ gelijk: ES = 0.63 ▪ gemengd: ES = 0.30 autonomie leerling: ▪ laag: ES = 0.30 ▪ hoog: ES = 0.94 evaluatie: ▪ individueel: ES = 0.59 ▪ groep: ES = 0.27
	33 studies (33 in PO)	rekenprestaties	ES = 0.22	-

- vervolg Tabel 3.4: Meta-analyses met betrekking tot coöperatief leren. -

Capar & Tarim (2013) coöperatieve werkvormen	26 studies (10 effectgroottes voor PO)[8]	rekenprestaties	-	onderwijsniveau: ▪ basisonderwijs: ES = 0.50 werkvorm: ▪ Samen Leren: ES = 0.95 ▪ ongestructureerd: ES = 0.91 ▪ heterogene groepjes op leerniveau: ES = 0.71
Shenderovich et al. (2015) *peer tutoring*	15 studies (2 in PO)	rekenprestaties	ES = -0.02 (*ns*)	-
Leung (2015) *peer tutoring*	72 studies (46 in PO)	algemene leerprestaties	ES = 0.36	vorm: ▪ gestructureerde *peer tutoring*: ES = 0.53 ▪ ongestructureerde *peer tutoring*: ES = 0.33
	17 studies (? in PO)	rekenprestaties	ES = 0.34	*peer tutoring* gericht op: ▪ berekeningen: ES = 0.69 ▪ berekeningen en conceptuele kennis: ES = 0.43 ▪ conceptuele kennis: ES = 0.24 (*ns*)

Self explanation

Wanneer een leerling zelf een verklaring maakt voor een concept of voor nieuwe informatie en deze aan zichzelf uitlegt, is er sprake van *self explanation* (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). De leerling probeert hierbij nieuwe informatie te begrijpen aan de hand van voorkennis en maakt mogelijke gevolgtrekkingen. *Self explanation* kan aangemoedigd en begeleid worden door een leerkracht en kan dus ingezet worden als werkvorm en als instructie (Rittle-Johnson, Loehr, & Durkin, 2017).

Er zijn twee meta-analyses gevonden naar de effecten van *self explanation* op rekenprestaties (zie Tabel 3.5). In deze meta-analyses werden de rekenprestaties vergeleken tussen enerzijds leerlingen die wel werden aangespoord tot het geven van eigen verklaringen en anderzijds leerlingen die niet spontaan zelf verklaringen gaven (Durkin, 2011) of leerlingen die niet werden aangespoord om zelf verklaringen te geven (Rittle-Johnson et al., 2017). In beide meta-analyses zijn vooral studies met leerlingen in het voorgezet of hoger onderwijs opgenomen en is er niet apart gerapporteerd over basisschoolleerlingen.

In de meta-analyse van Durkin (2011) worden tegenstrijdige bevindingen gerapporteerd. In sommige geanalyseerde studies werd een positief effect gevonden, terwijl in andere studies geen effect of zelfs een negatief effect werd gevonden. Gemiddeld bleek spontane en gestuurde *self explanation* een klein tot matig positief effect te hebben op procedurele rekenkennis. Het aantal studies in deze meta-analyse is echter relatief laag, slechts enkele van de geanalyseerde studies zijn gericht op het

[8] Wanneer niet eenduidig vast te stellen was hoeveel studies een bepaalde selectie betrof (zoals in PO, of met uitkomstmaat rekenprestaties) vermelden we het aantal effectgrootten.

basisonderwijs en ook is er een publicatie-*bias* gevonden, waardoor de resultaten met voorzichtigheid geïnterpreteerd dienen te worden.

Rittle-Johnson en collega's (2017) onderzochten het effect van *self explanation* op drie vormen van kennis: conceptuele kennis, procedurele kennis en procedurele transfer (het aanpassen of integreren van procedurele kennis om nieuwe problemen op te lossen). Wanneer de kennis direct na de interventie werd getoetst, werd er voor alle drie de vormen van kennis een positief effect gevonden. Hierbij was het grootste effect op procedurele transfer en werd er een kleiner effect gevonden voor conceptuele kennis en procedurele kennis. Bij de studies waar het effect ook een langere tijd na de interventie werd gemeten, is er alleen een effect gevonden voor procedurele transfer. Vanwege het geringe aantal studies zijn de auteurs voorzichtig met het trekken van conclusies over het langdurige effect van *self explanation*. Verder was het effect van *self explanation* op conceptuele kennis sterker wanneer er binnen de interventie gebruikt werd gemaakt van *scaffolding* (training vooraf, of voorstructurering van de verklaringen van de leerlingen).

Samenvattend blijkt uit deze twee meta-analyses dat er een klein tot matig positief effect van spontane en aangespoorde *self explanation* op conceptuele rekenkennis, procedurele rekenkennis en procedurele transfer is bij leerlingen van de basisschool tot en met het hoger onderwijs. Bovendien lijkt het aanbieden van *self explanation* via *scaffolding* een positief effect te hebben, met name op de conceptuele kennis. Er is nog onvoldoende duidelijk in hoeverre deze conclusies geldig zijn voor het basisonderwijs, omdat maar weinig studies op de basisschool waren gedaan.

Tabel 3.5: Meta-analyses met betrekking tot self explanation.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Durkin (2011) *self explanation*	18 studies (? in PO)	procedurele rekenkennis	ES = 0.37	-
Rittle-Johnson et al. (2017) *self explanation*	9 studies (? in PO)	procedurele transfer	ES = 0.46	-
	16 studies (? in PO)	conceptuele kennis	ES = 0.28	-
	19 studies (? in PO)	procedurele kennis	ES = 0.33	-

3.1.3 Hulpmiddelen

Technologische hulpmiddelen

Technologische hulpmiddelen, zoals computers, mobiele telefoons en laptops, kunnen op verschillende manieren met verschillende programma's worden ingezet tijdens de rekenles. Redenen om technologie in te zetten in het onderwijs zijn bijvoorbeeld: meer leerlinggerichte instructie, meer projectonderwijs, meer betrokkenheid en enthousiasme bij de leerlingen en betere relaties creëren tussen leerkracht en leerling en thuis en school (Zheng, Warschauer, Lin, & Chang, 2016). Middels meta-analyses en reviews zijn de effecten van verschillende technologische hulpmiddelen onderzocht. Tabel 3.6 geeft de technologische hulpmiddelen die zijn onderzocht.

Tabel 3.6: Onderzochte technologische hulpmiddelen.

technologische hulpmiddel	uitleg	studies
Exploratory Computerized Environments (ECE)	technologische leeromgevingen waarin exploratief geleerd kan worden	Sokolowski, Li en Willson (2015) Li en Ma (2010)
supplemental Computer-Assisted Instruction (CAI)	CAI biedt extra instructie boven de reguliere instructie op specifieke leerlingniveaus	Cheung en Slavin (2013) Li en Ma (2010) Slavin en Lake (2008)
Computer-Managed Learning (CML)	programma's waarmee de rekenniveaus en voortgang van leerlingen in kaart worden gebracht en aan de hand waarvan het rekenmateriaal wordt aangepast.	Cheung en Slavin (2013)
comprehensive models	programma's waarbij instructie zowel via de computer als via leerkracht worden aangeboden	Cheung en Slavin (2013)
Intelligent Tutoring Systems (ITS)	computer-ondersteunende leeromgevingen waarin rekenen kan worden geoefend met oefeningen die aangepast zijn aan de domeinkennis, leervaardigheden en –strategieën, emoties en motivatie van de betreffende leerling	Steenbergen-Hu en Cooper (2013)
communicatiemedia	e-mail en internet	Li en Ma (2010)
tools	Microsoft Office, zoals Word, Excel en Powerpoint	Li en Ma (2010)
overige apparaten	laptops	Zheng, Warschauer, Lin en Chang (2016)
	mobiele apparaten zoals tablets, mobiele telefoons en *smartphones* en PDA (*personal digital assistent*)	Tingir, Cavlazoglu, Caliskan, Koklu en Intepe-Tingir (2017)
videogames	educatieve rekenspellen	Young en collega's (2012)

Li en Ma (2010) onderscheidden verschillende vormen van computertechnologie (CT), zoals tutorials (waarbij informatie, demonstratie en oefening wordt geboden), *computer-assisted instruction* (CAI), communicatiemedia, exploratieve leeromgevingen en andere digitale tools om leren leuker en efficiënter te maken. Het algemene effect van CT op rekenprestaties van leerlingen uit het basisonderwijs en de eerste jaren van het voortgezet onderwijs is klein tot matig positief. Het effect bleek groter voor leerlingen uit het basisonderwijs dan voor leerlingen uit het voortgezet onderwijs. Het effect was echter kleiner wanneer de CT langer werd ingezet. Wanneer CT één termijn werd ingezet was het effect het grootst. CT die ingezet was met een constructivistische benadering, waarbij de instructie meer op de leerling was gericht en strategieën benadrukt, had een groter effect dan CT die ingezet was met een traditionele benadering, waarbij instructie meer leerkracht gestuurd was en gericht aan de hele klas. Er bestond geen verschil in het effect van de verschillende vormen van CT. In studies over CT die voor 1999 zijn gepubliceerd werden hogere effectgrootten gerapporteerd dan studies die na 1999 zijn gepubliceerd.

Educatieve technologische applicaties zijn programma's die lesmateriaal leveren of de leerprocessen ondersteunen. Deze applicaties hadden een klein, maar positief effect op rekenprestaties van leerlingen uit het basisonderwijs en voortgezet onderwijs volgens de resultaten van de meta-analyse van Cheung en Slavin (2013) naar experimentele studies waarbij technologie langdurig (12 weken of meer) werd ingezet. Er bleek daarbij geen verschil te bestaan in de effecten van educatieve technologische applicaties tussen studies bij leerlingen uit het basisonderwijs of het voortgezet

onderwijs. Technologische programma's die meer dan 30 minuten per week werden ingezet als ondersteuning van de rekenlessen hadden een klein effect op de rekenprestaties. Programma's die minder dan 30 minuten per week werden aangeboden hadden geen effect op rekenprestaties.

Cheung en Slavin (2013) onderzochten ook het effect van drie soorten technologische applicaties. Aanvullende *computer-assisted instruction* (CAI)*,* waarbij extra instructie op specifieke leerlingniveaus wordt aangeboden bovenop de reguliere instructie op specifieke leerlingniveaus, had daarbij het grootste effect, maar had desondanks een klein effect op rekenprestaties. CAI biedt de mogelijkheid om de sterktes en zwaktes van leerlingen in kaart te brengen en aan de hand daarvan oefeningen te kiezen (Slavin & Lake, 2008). Ook in de meta-analyse van Slavin en Lake (2008) bij leerlingen uit het basisonderwijs had CAI een klein, positief effect. *Computer-managed learning* (CML), programma's waarmee de rekenniveaus en voortgang van leerlingen in kaart worden gebracht en aan de hand waarvan het rekenmateriaal wordt aangepast, hadden een klein positief effect op rekenprestaties. *Comprehensive models*, programma's waarbij instructie zowel via de computer als via de leerkracht in de klas worden aangeboden, hadden ook een klein effect op rekenprestaties. Verder bleek dat de effectgrootten van technologische applicaties kleiner waren wanneer de steekproeven uit de onderzoeken groot waren en enkel gerandomiseerde onderzoeksdesigns meegenomen werden in de meta-analyse.

Specifieke digitale leeromgevingen

Intelligent Tutoring Systems (ITS) zijn digitale leeromgevingen waarin rekenen kan worden geoefend, aangepast aan de domeinkennis, leervaardigheden en –strategieën, emoties en motivatie van de betreffende leerling (Steenbergen-Hu & Cooper, 2013). Het is anders dan CAI, zoals onderzocht door Cheung en Slavin (2013), omdat er bij ITS ook binnen de leertaak aandacht is voor het oplossen van het probleem via feedback, hints en foutmeldingen (Steenbergen-Hu & Cooper, 2013). Wanneer de effecten van alle studies meegenomen werden, bleek er geen effect te zijn van ITS op rekenprestaties. Er werd daarbij geen onderscheid gemaakt tussen ITS-leeromgevingen die ter aanvulling van reguliere rekenlessen werden ingezet of ITS-leeromgevingen die in plaats van de reguliere lessen werden ingezet. De effecten verschilden ook niet tussen rekenwiskundedomeinen (basiswiskunde en algebra) en tussen prestatieniveaus van de leerlingen (algemene en lage presteerders). Net als in het onderzoek van Cheung en Slavin (2013) bleek er geen verschil te zijn in effectgrootte tussen schoolniveaus (basisonderwijs en voortgezet onderwijs). Wanneer ITS een korte termijn werd ingezet, had het wel een effect op rekenprestaties. Dit komt overeen met de bevinding van Li en Ma (2010) dat technologische applicaties meer effect hebben als die kort wordt ingezet.

Exploratory Computerized Environment (ECE) is een leeromgeving op computers of Ipads die leerlingen gebruiken om patronen te herkennen en problemen op te lossen middels exploratief leren (Sokolowski et al., 2015). ECE had een matig tot groot positief effect op rekenprestaties bij leerlingen uit het basisonderwijs en de eerste klassen van het voortgezet onderwijs (Sokolowski et al., 2015). ECE had een klein tot matig positief effect op rekenprestaties wanneer het exploreren ondersteunde, door leerlingen te laten experimenteren met modellen en herkennen van onderliggende patronen. ECE had ook een klein tot matig positief effect op rekenprestaties wanneer het probleemoplossen ondersteunde, door een oplossing te zoeken voor bijvoorbeeld contextopgaven. ECE had een matig tot groot positief effect op rekenprestaties binnen het domein geometrie en een matig tot groot positief effect op het domein algebra. Wanneer de inzet van ECE werd ondersteund door de leerkracht, had het een hoger effect dan wanneer ECE ondersteund werd met behulp van computers. De auteurs concludeerden hieruit dat enkel het gebruik van computers als methode niet voldoende was om het rekenkundig begrip van leerlingen te verbeteren. Beide ondersteuningsvormen, leerkrachtondersteuning en computerondersteuning, hadden afzonderlijk een matig tot groot positief effect op rekenprestaties. ECE had het grootste effect (matig tot groot) wanneer het twee tot vijf weken aan de leerlingen werd aangeboden. Er werd in het onderzoek echter geen onderzoek gedaan

naar de verschillen in het gebruik van ECE. Er is daardoor bijvoorbeeld niet duidelijk gemaakt of ECE werd ingezet in aanvulling op of in plaats van de rekenles.

Het gebruik van mobiele apparaten in de rekenles

Technologische applicaties kunnen op verschillende apparaten worden aangeboden, zoals computers, laptops, tablets en mobiele telefoons. Het gebruik van laptops tijdens rekenlessen had een klein positief effect op rekenprestaties bij leerlingen van de basisschool en de eerste klassen van het voortgezet onderwijs (Zheng et al., 2016). Het effect van laptopgebruik kan variëren door de frequentie. In klassen waar laptops goed toegankelijk zijn en vaak werden gebruikt, had frequentie van laptopgebruik een positief effect op rekenscores. Wanneer laptops frequent werden gebruikt in klassen waar het niet gewoon is om een laptop te gebruiken, bleek dat een negatief effect te hebben op rekenscores (zie bijvoorbeeld Bebell & Kay, 2010).

In het rekenonderwijs kunnen naast computers en laptops ook andere mobiele apparaten worden ingezet, zoals tablets en mobiele telefoons. Het effect van het gebruik van tablets en mobiele telefoons in de rekenles bleek een klein effect te hebben op rekenprestaties van leerlingen van de basisschool en de eerste klassen van het voortgezet onderwijs (Tingir et al., 2017). Het effect van mobiele apparaten op rekenen was kleiner dan op *science* en lezen. Wanneer de prestaties op rekenen, lezen en *science* bij elkaar werden genomen, bleken er geen verschillen te zijn naar het type mobiele apparaat. Verder zijn er geen verschillen gevonden in het effect in het basisonderwijs en voortgezet onderwijs.

Videogames

Naast dat veel leerlingen in een informele setting, zoals thuis, videogames spelen in hun vrije tijd, worden videogames binnen het rekenonderwijs steeds vaker ingezet. De literatuurzoektocht heeft geen meta-analyses met effectgrootten opgeleverd over het effect van videogames op rekenprestaties, maar er is wel een review gevonden van Young en collega's (2012). Zij beschreven dat eerder gevonden positieve effecten op rekenprestaties van videogames die in de klas werden aangeboden in later onderzoek niet kon worden gerepliceerd (Young et al., 2012). Tot dusver bestaat er dus geen consensus over het effect van videogames op rekenprestaties

Conclusies technologische hulpmiddelen

Wanneer technologische hulpmiddelen worden ingezet, blijkt dat een positief effect te hebben op rekenprestaties bij leerlingen uit het basisonderwijs. Er bestaat echter nog geen consensus over de grootte van het effect en welk type technologie de meeste invloed heeft op rekenprestaties. Li en Ma (2010) en Cheung en Slavin (2013) vonden geen verschillen in de effecten van verschillende computerapplicaties op rekenprestaties. De resultaten van de meta-analyses lijken te impliceren dat het effect van technologie op rekenprestaties het grootst is wanneer technologie kort wordt ingezet (een aantal weken). Hier bestaat echter ook nog geen consensus over.

Tabel 3.7: Meta-analyses met betrekking tot technologische hulpmiddelen.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Cheung en Slavin (2013) educatieve technologische applicaties	74 studies (45 in PO)	rekenprestaties	ES = 0.16	applicatietype: ▪ CAI: ES = 0.19 ▪ CML: ES = 0.09 ▪ *comprehensive*: ES = 0.06 duur: ▪ <30 minuten per week: ES = 0.06 (*ns*) ▪ 30-75 minuten per week: ES = 0.20 ▪ >75 minuten per week: ES = 0.14
Li en Ma (2010) computer technologie	85 effectgrootten (48 in PO)	rekenprestaties	ES = 0.28[1]	onderwijsniveau: ▪ basisonderwijs: ES = 0.78 ▪ voortgezet onderwijs: ES = 0.61 duur: ▪ 1 semester: ES = 0.88 ▪ 1 semester tot 1 jaar: ES = 0.57 ▪ langer dan 1 jaar: ES = 0.55 instructiemethode: ▪ traditioneel: ES = 0.62 ▪ constructivistisch: ES = 1.29 type: ▪ *tutorial*: ES = 0.68 ▪ communicatiemedia: ES = 0.39 ▪ exploratieve leeromgeving: ES = 1.32 ▪ *tools*: ES = 0.54 onderzoeksdesign: ▪ experimenteel: ES = 0.42 ▪ quasi-experimenteel: ES = 0.80 publicatie: ▪ voor 1999: ES = 0.99 ▪ na 1999: ES = 0.42
Slavin en Lake (2008) *Computer Assisted Instruction* (CAI)	87 studies (87 in PO)	rekenprestaties	ES = 0.19	-

- vervolg Tabel 3.7: Meta-analyses met betrekking tot technologische hulpmiddelen. -

Sokolowski et al. (2015) *Exploratory Computerized Environments* (ECE)	24 studies (18 in PO)	rekenprestaties	ES = 0.60	groep: ▪ groep 3-5: ES = 0.61 ▪ groep 6-7: ES = 0.41 ▪ groep 8 en eerste klassen voortgezet onderwijs: ES = 0.65 ondersteuning: ▪ exploratie: ES = 0.62 ▪ probleem oplossen: ES = 0.54 domein: ▪ geometrie: ES = 0.67 ▪ algebra: ES = 0.61 ondersteuningsvorm: ▪ begeleid door de leerkracht: ES = 0.75 ▪ computerondersteuning: ES = 0.61. duur: ▪ kort: ES = 0.47 ▪ gemiddeld (2-5 weken): ES = 0.63 ▪ lang: ES = 0.62
Steenbergen-Hu en Cooper (2013) *Intelligent Tutoring Systems* (ITS)	26 studies (3 in PO)	rekenprestaties	ES = 0.01	domein: ▪ basiswiskunde: ES = 0.11 ▪ algebra: ES = -0.03 duur: ▪ minder dan 1 semester: ES = 0.52 ▪ 1 semester: ES = 0.06 ▪ 1 schooljaar of langer: ES = -0.08
Tingir et al. (2017) mobiele apparaten	14 studies (8 in PO)	prestaties op rekenen, lezen en *science*	ES = 0.48	type apparaat: ▪ *tablet*: ES = 0.34 ▪ PDA: ES = 0.61 ▪ *smartphone*: ES = 0.45 ▪ mobiel apparaat: ES = 0.54 schoolvak: ▪ rekenen: ES = 0.16 ▪ *science*: ES = 0.53 ▪ lezen: ES = 0.67 schoolniveau: ▪ *elementary school*: ES = 0.55 ▪ *middle school*: ES = 0.43 ▪ *high school*: ES = 0.28 onderzoeksdesign: ▪ experimenteel: ES = 0.63 ▪ quasi-experimenteel: ES = 0.42

- vervolg Tabel 3.7: Meta-analyses met betrekking tot technologische hulpmiddelen. -

Zheng et al. (2016) *one-to-one laptop environments*	7 studies (? in PO)	rekenprestaties	ES = 0.16	-

[1]Effectgrootte bepaald aan de hand van *WLS Regression Model.* Grootte gebaseerd op Rosenthal en Rosnow (1984): klein (<.30 SD), gemiddeld (.30 - .50) en groot (>.50).

Niet-technologische hulpmiddelen

Naast computerondersteuning en het gebruik van games, zijn er ook allerlei niet-technologische hulpmiddelen die leerkrachten kunnen gebruiken in hun rekeninstructie. Zoals hieronder verder beschreven zal worden, zijn er diverse studies gevonden waarin onderzoek is gedaan naar de effecten van specifieke hulpmiddelen. Daarnaast is er ook een meta-analyse gevonden waarin een combinatie van hulpmiddelen onderzocht is (Kablan, Topan, & Erkan, 2013). Kablan en collega's hebben zich in hun meta-analyse gericht op Turkse studies naar de effecten van hulpmiddelen op algemene schoolprestaties van leerlingen in het basisonderwijs tot en met het hoger onderwijs. Tot de onderzochte hulpmiddelen behoorden computerpresentaties, driedimensionale materialen, cartoons, *concept maps* of een combinatie hiervan. Zoals weergegeven in Tabel 3.8 bleken de hulpmiddelen een groot effect te hebben op schoolprestaties. Er werden geen verschillen in effectiviteit gevonden tussen de verschillende schoolniveaus, vakken of hulpmiddelen.

Tabel 3.8: Meta-analyses met betrekking tot niet-technologische hulpmiddelen (algemeen).

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Kablan et al. (2013) combinatie van niet-technologi-sche hulpmiddelen	57 studies (35 in PO)	algemene schoolprestaties	ES = 1.27	-

Het lijkt er dus op dat diverse hulpmiddelen van toegevoegde waarde zouden kunnen zijn in het verhogen van rekenprestaties. Hieronder zullen de resultaten van reviews en meta-analyses naar de effecten van driedimensionale materialen, geïllustreerde boeken en *concept maps* beschreven worden.

Driedimensionale materialen

Een veelgebruikt hulpmiddel in rekeninstructie is de toepassing van driedimensionale materialen (Carbonneau, Marley, & Selig, 2013). Dit hulpmiddel biedt leerlingen de mogelijkheid tot fysieke interactie met objecten om bepaalde informatie beter te begrijpen (Carbonneau & Marley, 2013). Hierbij kan bijvoorbeeld gedacht worden aan het spelen met blokken of nep-briefgeld om meer inzicht in getallen te krijgen.

Er is één meta-analyse gevonden waarin de effecten van het gebruik van driedimensionale materialen ten opzichte van het gebruik van abstracte rekenkundige symbolen op rekenprestaties onderzocht is (Carbonneau et al., 2013), gericht op leerlingen in het basisonderwijs, voortgezet onderwijs en hoger onderwijs (zie Tabel 3.9). Er werd een klein tot matig algeheel effect gevonden van het gebruik van driedimensionale materialen op rekenprestaties. Veel instructiebegeleiding bij het gebruik van de materialen was effectiever dan weinig instructiebegeleiding. Ook bleek dat er verschillen in effectiviteit waren tussen specifieke rekenonderwerpen. Het effect voor breuken was groter dan voor bewerkingen en algebra. Verder maakte ook de ontwikkelingsfase van de leerling uit. De effectiviteit was voor leerlingen van 7 tot 11 jaar oud groter dan voor leerlingen van 12 jaar en ouder. Tenslotte

bleek dat een instructieduur van minder dan 14 dagen of een instructieduur van 15 tot 45 dagen effectiever was dan een instructieduur van meer dan 46 dagen.

Daarnaast zijn verschillende uitkomstvariabelen binnen rekenen onderscheiden: basisvaardigheden, probleemoplossing, transfer en *justification* (waar leerlingen gevraagd werden hun oplossingsmethode uit te leggen). Voor basisvaardigheden werd er een matig tot groot algeheel effect gevonden van het gebruik van driedimensionale materialen, waarbij het uitmaakte door wie het gebruik van de materialen geïmplementeerd werd: als dat de onderzoeker was, was het effect kleiner dan wanneer dat de leerkracht was. Verder bleek dat voor basisvaardigheden veel instructiebegeleiding effectiever was dan weinig instructiebegeleiding, er grotere effecten waren voor breuken en algebra dan voor bewerkingen, perceptueel rijke materialen (materialen die representatief zijn voor een echt object, zoals nep-briefgeld) minder effectief waren dan perceptueel arme materialen (materialen die geen ander object moeten voorstellen, zoals blokken), vooral leerlingen van 7 tot 11 jaar baat hadden van de materialen en dat een instructieduur van maximaal 14 dagen effectiever was dan een instructieduur van 15 tot 45 dagen. Bij probleemoplossingsvaardigheden had gebruik van driedimensionale materialen op de leerlingen een klein tot matig positief effect. Uit de moderatoranalyses kwamen patronen naar voren die vergelijkbaar waren met die van basisvaardigheden. Op transfer had het gebruik van driedimensionale materialen een verwaarloosbaar tot klein positief effect. In tegenstelling tot bij de andere uitkomstvariabelen, was voor transfer weinig instructiebegeleiding effectiever dan veel instructiebegeleiding. Daarnaast bleek dat perceptueel rijke materialen effectiever waren dan perceptueel arme materialen. Op *justification* tenslotte had het gebruik van driedimensionale materialen een klein tot matig positief effect. Er zijn voor deze uitkomstvariabele geen moderatoranalyses uitgevoerd.

Op basis van de studie van Carbonneau en collega's kan geconcludeerd worden dat het gebruik van driedimensionale materialen een redelijk effectief hulpmiddel kan zijn om de rekenprestaties van basisschoolleerlingen te verhogen. Wel lijkt het erop dat de effecten op transfer redelijk beperkt zijn. Verder is uit deze meta-analyse gebleken dat het nuttig is om onderscheid te maken tussen de verschillende uitkomstvariabelen, mede doordat er uit de moderatoranalyses voor deze variabelen verschillende patronen naar voren kwamen.

Tabel 3.9: Meta-analyses met betrekking tot driedimensionale niet-technische hulpmiddelen.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Carbonneau et al. (2013) gebruik van driedimensio-nale materialen	55 studies (43 in PO)	algemene rekenprestaties	ES = 0.37	instructiebegeleiding: ▪ veel: ES = 0.46 ▪ weinig: ES = 0.29 onderwerp: ▪ decimale stelsel: ES = 0.58 ▪ bewerkingen: ES = 0.27 ▪ geometrie: ES = 0.37 ▪ breuken: ES = 0.69 ▪ algebra: ES = 0.21 ontwikkelingsfase: ▪ 3-6 jaar: ES = 0.33 ▪ 7-11 jaar: ES = 0.45 ▪ 12 jaar en ouder: ES = 0.16 instructietijd: ▪ ≤14 dagen: ES = 0.34 ▪ 15-45 dagen: ES = 0.45 ▪ ≥46 dagen: ES = 0.14 (*ns*)
	53 studies (? in PO)	basisvaardig-heden	ES = 0.59	implementator: ▪ onderzoeker: ES = 0.13 ▪ leerkracht: ES = 0.82 instructiebegeleiding: ▪ veel: ES = 0.90 ▪ weinig: ES = 0.19 onderwerp: ▪ place value: ES = 0.70 ▪ bewerkingen: ES = 0.39 ▪ geometrie: ES = 0.57 ▪ breuken: ES = 0.93 ▪ algebra: ES = 0.84 perceptuele rijkheid: ▪ rijk: ES = 0.28 ▪ arm: ES = 0.77 ontwikkelingsfase: ▪ 3-6 jaar: ES = -0.09 (*ns*) ▪ 7-11 jaar: ES = 0.81 ▪ 12 jaar en ouder: ES = 0.31 instructietijd: ▪ ≤14 dagen: ES = 0.59 ▪ 15-45 dagen: ES = 0.35 ▪ ≥46 dagen: ES = 0.49

- vervolg Tabel 3.9: Meta-analyses met betrekking tot driedimensionale niet-technische hulpmiddelen.-

	9 studies (? in PO)	probleem-oplossen	ES = 0.46	implementator: ▪ onderzoeker: ES = -0.39 (*ns*) ▪ leerkracht: ES = 0.82 instructiebegeleiding: ▪ veel: ES = 1.06 ▪ weinig: ES = 0.04 (*ns*) onderwerp: ▪ *place value*: ES = 0.48 (*ns*) ▪ bewerkingen: ES = 0.02 (*ns*) ▪ geometrie: ES = 0.72 ▪ breuken: ES = 2.50 perceptuele rijkheid: ▪ rijk: ES = -0.27 (*ns*) ▪ arm: ES = 0.80 instructietijd: ▪ ≤14 dagen: ES = 0.86 ▪ 15-45 dagen: ES = -0.62 ▪ ≥46 dagen: ES = 0.25 (*ns*)
	13 studies (? in PO)	transfer	ES = 0.13	instructiebegeleiding: ▪ veel: ES = 0.00 (*ns*) ▪ weinig: ES = 0.27 perceptuele rijkheid: ▪ rijk: ES = 0.48 ▪ arm: ES = -0.02 (*ns*)
	2 studies (? in PO)	*justification*	ES = 0.38	-

Geïllustreerde boeken

Naast het gebruik van driedimensionale materialen, zouden ook geïllustreerde boeken gebruikt kunnen worden om rekenkundige concepten visueel te ondersteunen (Flevares & Schiff, 2014). Afbeeldingen kunnen bijvoorbeeld helpen om te leren tellen, om geometrische vormen te leren of om sommen te oefenen. Zowel boeken die een combinatie van tekst en afbeeldingen bevatten als boeken die alleen afbeeldingen bevatten zouden bruikbaar kunnen zijn (Flevares & Schiff, 2014).

Er zijn geen meta-analyses gevonden waarin de effecten van het gebruik van geïllustreerde boeken op rekenprestaties kwantitatief getoetst zijn. Wel is er een narratieve review gevonden waarin dit onderwerp centraal staat (Flevares & Schiff, 2014). De studies die beschreven zijn in deze review zijn gericht op kinderen in de eerste jaren van het basisonderwijs. Flevares en Schiff beschrijven dat het gebruik van geïllustreerde boeken zou kunnen helpen bij het stimuleren van de interactie tussen leerlingen en hun ouders of leerkrachten, het creëren van visuele representaties van een concept, het leggen van relaties tussen concepten en eigen ervaringen en het verbeteren van probleemoplossingsvaardigheden, redenatievaardigheden en bewijsvaardigheden. Er zijn onvoldoende empirische studies gedaan om sterke conclusies te kunnen trekken over de daadwerkelijke effecten van het gebruik van geïllustreerde boeken op rekenprestaties.

Concept maps

Een ander hulpmiddel dat gebruikt kan worden in het onderwijsleerproces is het gebruik van *concept maps*. Een *concept map* kan worden omschreven als een diagram waarin verschillende concepten zijn

opgenomen waarbij door middel van links de onderlinge relaties tussen deze concepten kunnen worden aangegeven (Schroeder, Nesbit, Anguiano, & Adesope, 2017). *Concept maps* kunnen op twee manieren ingezet worden (Schroeder et al., 2017). Leerlingen kunnen bestaande *concept maps* bestuderen ofwel zelf een *concept map* construeren. Het gebruik van *concept maps* zou leren betekenisvoller kunnen maken en de cognitieve belasting voor leerlingen kunnen verminderen (Schroeder et al., 2017).

Er is geen meta-analyse gevonden naar de effecten van het gebruik van *concept maps* op rekenprestaties specifiek, maar wel een meta-analyse naar de effecten op algemene schoolprestaties (Schroeder et al., 2017). In deze meta-analyse is het bestuderen of construeren van *concept maps* vergeleken met andere leeractiviteiten zoals het voeren van discussies of het bestuderen van teksten. De geïncludeerde studies hebben betrekking op leerlingen vanaf groep 6 in het basisonderwijs tot en met leerlingen in het hoger onderwijs. Zoals weergegeven in Tabel 3.10 werd er een matig tot groot positief effect gevonden van het gebruik van *concept maps* op schoolprestaties. Er was geen significant verschil tussen de effecten op prestaties in STEM-onderwijs enerzijds en andere vakgebieden anderzijds. Hoe langer een leerling *concept maps* gebruikte als leerstrategie, hoe groter het effect op de prestaties was. Het bestuderen van bestaande *concept maps* had een klein tot matig positief effect op schoolprestaties, terwijl het zelf construeren van een concept map een matig tot groot positief effect had. Bij het bestuderen van *concept maps* bleek verder dat het effect voor leerlingen in het basisonderwijs en voortgezet onderwijs groter was dan voor leerlingen in het hoger onderwijs. Bij het construeren van *concept maps* was een gebruiksduur van één tot vier weken of vier weken of meer effectiever dan een gebruiksduur van één week of minder.

Samenvattend leek het gebruik van *concept maps* een goede leerstrategie te zijn om schoolprestaties te verhogen (al is onduidelijk wat precies de effecten op rekenprestaties waren). Het zelf construeren van *concept maps* leek effectiever te zijn dan het bestuderen van bestaande *concept maps*. Verder leken leerlingen er profijt van te hebben om langere tijd (meer dan één week) gebruik te maken van deze strategie.

Tabel 3.10: Meta-analyses met betrekking tot *concept maps*.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Schroeder et al. (2017) construeren of bestuderen van concept maps	142 effectgrootten (29 voor groep 6 tot en met de tweede klas)	algemene schoolprestaties	ES = 0.58	gebruiksduur: ▪ < 1 week: ES = 0.36 ▪ 1-4 weken: ES = 0.68 ▪ > 4 weken: ES = 0.72 ▪ onbekend: ES = 0.06 (*ns*)
Schroeder et al. (2017) bestuderen van concept maps	67 effectgrootten (? in PO)	algemene schoolprestaties	ES = 0.43	schoolniveau: ▪ groep 6-2e klas: ES = 0.82 ▪ 2e klas-6e klas: ES = 1.24 ▪ hoger onderwijs: ES = 0.32
Schroeder et al. (2017) construeren van concept maps	75 effectgrootten (? in PO)	algemene schoolprestaties	ES = 0.72	gebruiksduur: ▪ < 1 week: ES = 0.40 ▪ 1-4 weken: ES = 0.94 ▪ > 4 weken: ES = 0.72 ▪ onbekend: ES = 0.06 (*ns*)

3.1.4 Toetsing

Toetsing is een essentieel onderdeel van de reken-wiskundeles. Toetsresultaten kunnen op twee verschillende manieren gebruikt worden, namelijk summatief (evaluatie van het leerproces achteraf) en formatief (bijsturing van het leerproces). Summatieve toetsing vindt plaats aan het einde van het leerproces en heeft vaak selectie of niveaubepaling als doel op basis van de beoordeling van de leerprestaties. Voor de effecten van summatieve toetsing in de les op rekenprestaties zijn geen meta-analyses of reviews gevonden. Formatieve toetsing vindt plaats tijdens het leerproces en heeft als doel om het leerproces te ondersteunen en in de gewenste richting te sturen. In het algemeen is vastgesteld dat dit kan bijdragen aan verbeterde leerresultaten, vooral als leerlingen inhoudelijk feedback krijgen: taakgericht en met aanwijzingen wat ze moeten veranderen (Blok, Ledoux, & Roeleveld, 2015).

De invloed van formatieve toetsing op reken- en schoolprestaties is onderzocht in een meta-analyse van Kingston en Nash (2011) bij leerlingen uit de kleuterklas tot en met de laatste klas van de middelbare school (zie Tabel 3.11). Binnen onderzoek naar formatieve toetsing bestaat geen consensus over de definitie. Daarom bestaat veel variatie in de gehanteerde definities binnen de meta-analyse. Een overeenkomend kenmerk binnen de beschrijvingen van de formatieve toetsvormen is dat de informatie, die via verschillende vormen van toetsing werd verzameld, werd gebruikt om het leerproces te ondersteunen. Het effect werd onderzocht door groepen waarbij formatieve toetsing plaatsvond te vergelijken met groepen waarbij reguliere toetsingspraktijken of geen vorm van formatieve toetsing werd toegepast. Uit de analyse bleek het effect van formatieve toetsing op schoolprestaties in het algemeen, en voor rekenen specifiek, klein (Kingston & Nash, 2011). De vorm van formatieve toetsing maakte uit voor schoolprestaties in het algemeen. Het inzetten van toetssystemen om op specifieke momenten in het curriculum het leerproces van leerlingen in kaart te brengen had geen effect op schoolprestaties. Andere vormen van formatieve toetsing waarbij verzamelde data werd ingezet om de instructie aan te passen, waarbij reflectie plaatsvond of waarbij gesprekken werden gevoerd over de toets, hadden geen effect op schoolprestaties. Het effect op schoolprestaties van formatieve toetsing was klein wanneer gedetailleerde feedback werd gegeven over toetsen om leerlingen te begeleiden in het leerproces. Het inzetten van online of computeriseerde toetsingssystemen die verslagen boden van de leerlinguitkomsten had een klein tot matig effect op schoolprestaties. Wanneer docenten via professionele ontwikkeling tijd besteedden aan het uitvoeren van aspecten van formatieve toetsing was het effect op schoolprestaties gemiddeld matig. De specifieke effecten van de verschillende vormen van formatieve toetsing op rekenprestaties zijn in de meta-analyse niet gerapporteerd.

De invloed van het gebruik van digitale leerlingvolgsystemen door leerkrachten, waarmee leerkrachten op basis van toetsen feedback krijgen over de resultaten van het door hen aangeboden onderwijs, op de prestaties van leerlingen is onderzocht in een meta-analyse van Faber en Visscher (2014). Uit de meta-analyse bleek dat er gemiddeld over de vakgebieden een klein effect was, en alleen op het vakgebied rekenen bleek het effect significant (zie Tabel 3.11). Het effect was sterker bij studies met een relatief kleine steekproef en intensieve interventie dan in studies met grotere steekproef en minder intensieve interventie. Verder werden de grootste effecten op leerresultaten gevonden wanneer de digitale leerlingvolgsystemen frequent feedback gaven (ten minste maandelijks), gecombineerd werden met een interventie waarin advies over de instructie werd gegeven, en wanneer leerkrachten de resultaten niet vergelijken met normen of standaarden.

In plaats van de focus op de feedback die leerlingvolgsystemen geven aan leerkrachten, richtte de meta-analyse van Van der Kleij, Feskens, en Eggen (2015) zich op het effect van feedback-per-opgave aan leerlingen binnen digitale leeromgevingen (zie Tabel 3.11). De uitkomstmaten waren prestaties bij sociale vakken, rekenen-wiskunde, *science*, biologie, geografie en talen. De studies waren uitgevoerd bij leerlingen uit basisonderwijs tot en met volwassenenonderwijs. *Computer-based assessments*

(CBA's) bieden de mogelijkheid om op verschillende manieren feedback te geven. Enkel aangeven of het antwoord goed of fout is, had een klein effect op de verschillende uitkomstmaten. Wanneer daarnaast ook het correcte antwoord werd gegeven, had dat een klein tot matig effect op de uitkomsten. Feedback met uitgebreide uitleg had een matig effect. Deze uitgebreide uitleg kan gezien worden als een vorm van instructie, zoals hints, extra informatie of lesmateriaal, een uitgewerkt voorbeeld en een uitleg van het correcte antwoord (Shute, 2008). Hoewel deze resultaten laten zien dat binnen digitale leeromgevingen meer uitgebreide feedback een positieve invloed had op de resultaten dan enkel feedback van de correctheid van het antwoord (met of zonder het noemen van het correcte antwoord), is het belangrijk om op te merken dat er maar weinig onderzoeken zijn meegenomen naar het effect op rekenprestaties in het basisonderwijs.

Naast de verschillende feedbackvormen zijn ook de effecten van meerdere moderatoren onderzocht (niet specifiek bij de onderzoeken uit het basisonderwijs). Het geven van feedback had een groot positief effect op rekenwiskunde-uitkomsten. Het direct aanbieden van de feedback had daarnaast een groter effect dan uitgestelde feedback. Het effect van feedback bleek groter te zijn wanneer het werd onderzocht op hogere orde leeruitkomsten, zoals het toepassen van kennis in een nieuwe situatie dan op lagere orde leeruitkomsten, zoals reproduceren, herkennen en begrijpen van concepten zonder deze toe te passen. Tevens bleek dat uitgebreide feedback specifiek een groter effect had op hogere orde leeruitkomsten dan op lagere orde leeruitkomsten. Het effect voor leerlingen van de universiteit was iets groter dan in het basis- en voortgezet onderwijs. Feedback per opgave aan de leerling had dus een positieve invloed op rekenprestaties. Met name feedback met uitgebreide uitleg (zoals hints, extra informatie of lesmateriaal, een uitgewerkt voorbeeld en een uitleg van het correcte antwoord) had een positief effect op schoolprestaties in het algemeen.

Samenvattend blijkt formatieve toetsing, waarbij op de één of andere manier toetsgegevens worden gebruikt om het leerproces te verbeteren, positief samen te hangen met leerprestaties in het algemeen, en rekenprestaties specifiek. Twee ontvangers van de feedback kunnen worden geïdentificeerd: de leerkracht en de leerling. Het gebruik van de feedback uit digitale toets- of leerlingvolgsystemen door leerkrachten hangt positief samen met de rekenprestaties van hun leerlingen. Dit is vooral zo wanneer leerkrachten zich hierin professionaliseren, ze frequent feedback krijgen en een koppeling met de benodigde instructie wordt gemaakt. Als leerlingen in digitale toets- of leeromgeving per gemaakte opgave feedback ontvangen heeft dit een positief effect op hun rekenprestaties. Dit is vooral effectief als het uitgebreide feedback met uitleg is, in plaats van alleen feedback of het antwoord goed of fout is en wat het goede antwoord had moeten zijn.

Tabel 3.11: Meta-analyses met betrekking tot toetsing.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Kingston & Nash (2011) formatieve toetsing	42 effectgrootten (18 in K-7)	algemene schoolprestaties	ES = 0.20	verschillende vormen van formatieve toetsing: ▪ toetssystemen geïntegreerd in het curriculum: ES = -0.05 (*ns*) ▪ overig (reflectie of gesprekken over de toets): ES = 0.10 (*ns*) ▪ gedetailleerde feedback: ES = 0.03 ▪ gecomputeriseerde toetssystemen: ES = 0.28 ▪ professionele ontwikkeling van docenten: ES = 0.30
	19 effectgrootten (? in PO)	rekenprestaties	ES = 0.17	-
Faber & Visscher (2014) gebruik digitale leerlingvolgsy-stemen door leerkrachten	15 studies (14 in PO)	prestaties op verschillende uitkomstmaten	ES = 0.07	omvang steekproef: ▪ klein: ES = 0.40 ▪ groot ES = 0.06 feedbackfrequentie: ▪ laag: ES = -0.01 (*ns*) ▪ gemiddeld: ES = 0.05 (*ns*) ▪ hoog: ES = 0.17 inhoud feedback: ▪ leerstofinhoudelijk: ES = 0.05 (*ns*) ▪ scores: ES = 0.07 (*ns*) ▪ instructieadvies: ES = 0.31 ▪ ontwikkeling: ES = 0.42 (*ns*) inhoud interventie: ▪ begeleiding bij uitvoering: ES = 0.03 (*ns*) ▪ technisch: ES = 0.04 (*ns*) ▪ vertaling naar instructie: ES = 0.14 vergelijking met standaarden/normen: ▪ ja: ES = 0.05 ▪ nee: ES = 0.16
	18 effectgrootten (? in PO)	rekenprestaties	ES = 0.11	-

- vervolg Tabel 3.11: Meta-analyses met betrekking tot toetsing. -

Van der Kleij et al. (2015) feedback aan leerlingen in digitale toetssystemen	70 effectgrootten (2 in PO)	prestaties op verschillende uitkomsten	kennisgeving resultaten: ES = 0.05 (*ns*)
			kennisgeving van de correcte respons: ES = 0.33 (*ns*)
			uitgebreide feedback: ES = 0.49
				moment van feedback
				▪ direct: ES = 0.46
				▪ uitgesteld: ES = 0.22 (*ns*)
				type vaardigheden:
				▪ lagere orde: ES = 0.31
				  ▪ ES = 0.37 bij uitgebreide feedback
				▪ hogere orde (bevat zowel lagere en hogere orde): ES = 0.59
				  ▪ ES = 0.67 bij uitgebreide feedback
				onderwijsniveau:
				▪ universiteit: ES = 0.44
				▪ basis- en voortgezet onderwijs: ES = 0.34

	9 effectgrootten (? in PO)	rekenprestaties	ES = 0.93	-

3.1.5 Differentiatie

Leerlingen kunnen op allerlei aspecten verschillen van elkaar. Ze kunnen bijvoorbeeld verschillen in hun cognitieve vaardigheden, maar ook in hun interesses en doelen. Differentiatie heeft betrekking op hoe leerkrachten omgaan met deze verschillen en kan worden omschreven als een combinatie van kennis, houding en gedrag (Deunk, Annemieke, Roel, & Bosker, 2015). Leerkrachten moeten inzicht hebben in de verschillen tussen leerlingen en vervolgens bepalen wat en hoe diverse leerlingen moeten leren. Differentiatie kan op verschillende wijzen worden vormgegeven. Op basis van de zoekcriteria zijn alleen meta-analyses en reviews gevonden waarin de effecten van het werken in groepen op rekenprestaties onderzocht zijn.

Werken in groepen

Leerkrachten kunnen er tijdens een les voor kiezen om leerlingen in groepen te laten werken. Wanneer leerlingen worden ingedeeld op hun niveau, kunnen leerkrachten kiezen voor groepen met kinderen met een vergelijkbaar niveau (homogene groepen) of voor groepen met kinderen van verschillende niveaus (heterogene groepen; Deunk et al., 2015). Naar de effecten van het werken met groepen zijn een meta-analyse en een systematische review gevonden.

In de meta-analyse van Lou, Abrami en Spence (2000) zijn de effecten onderzocht van het vormen van groepen binnen klassen op de algemene schoolprestaties van leerlingen in het basisonderwijs tot en met het hoger onderwijs (zie Tabel 3.12). Het werken met groepen had een verwaarloosbaar tot klein positief effect ten opzichte van het werken met klassikale instructie. De effectgrootten waren echter groot onder de volgende implementatiecondities: er werd gebruik gemaakt van een lokaal ontwikkelde test in plaats van een gestandaardiseerde test, leerkrachten die werkten met de groepen kregen meer of andere training dan leerkrachten die werkten met klassikale instructie, het indelen van de groepen was gebaseerd op zowel niveau als andere aspecten zoals geslacht en groepscohesie en er werd gebruik gemaakt van coöperatief leren. Verder waren de effecten het grootst voor leerlingen met een hoog niveau en voor leerlingen in het basisonderwijs.

Deunk en collega's (2015) hebben een systematische review uitgevoerd naar de effecten van differentiatie, waarbij ook een beperkt aantal studies zijn beschreven die specifiek gericht zijn op rekenen in het basisonderwijs. Uit een interventiestudie bleek dat er geen significant verschil was op rekenprestaties tussen het werken met homogene of heterogene groepen binnen klassen. Uit studies

naar verschillen tussen klassen bleek dat het werken in homogene klassen in vergelijking met heterogene klassen een verwaarloosbaar tot klein negatief effect had op rekenprestaties. Het werken met een computerprogramma waarin kinderen oefenen met rekenen en waarbij leerkrachten op basis van de resultaten advies krijgen over differentiatie had een klein tot matig positief effect. Deunk en collega's beschrijven tenslotte dat het werken met groepen vooral effectief zou kunnen zijn wanneer het onderdeel is van een breder programma, waarin bijvoorbeeld ook aandacht is voor aangepaste instructies en voortgangscontroles.

Uit de resultaten van bovenstaande studies blijkt dat het werken met groepen effectief zou kunnen zijn, mits het goed geïmplementeerd wordt. Vooral wanneer het werken met groepen onderdeel is van een breder programma waarin ook veel aandacht is voor aangepaste instructie en waarin leerkrachten training krijgen, zou het tot positieve resultaten kunnen leiden. Uit deze studies blijken geen verschillen tussen het werken met homogene groepen en heterogene groepen binnen klassen.

Tabel 3.12: Meta-analyses met betrekking tot werken in groepen.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Lou et al. (2000) werken met groepen binnen een klas	51 studies (? in PO)	algemene schoolprestaties	ES = 0.16	type toets: ▪ niet gestandaardiseerd: ES = 0.36 ▪ gestandaardiseerd: ES = 0.07 (*ns*) leerkrachttraining: ▪ gelijkwaardig: ES = 0.14 ▪ meer of anders: ES = 0.42 groepsindeling: ▪ gebaseerd op niveau en andere criteria: ES = 0.39 ▪ gebaseerd op alleen niveau: ES = 0.14 type instructie: ▪ coöperatief leren: ES = 0.28 ▪ anders: ES = 0.15 niveau van de leerling: ▪ hoog: ES = 0.28 ▪ gemiddeld: ES = 0.19 ▪ laag: ES = 0.37

3.1.6 Huiswerk

Huiswerk wordt doorgaans gedefinieerd als taken opgegeven door leerkrachten aan leerlingen om te maken buiten schooltijd. Uit de synthese van Amerikaanse studies door Cooper, Robinson, en Patall (2006) bleek dat bij leerlingen in het basisonderwijs de tijd die gespendeerd werd aan huiswerk niet gerelateerd is aan de schoolprestaties op verschillende vakken. Fan, Xu, Cai, He, en Fan (2017) deden een meta-analyse naar het effect van huiswerk op de reken- en *science*prestaties van leerlingen op het basis- en voortgezet onderwijs (zie Tabel 3.13). De relatie tussen huiswerk en de prestaties was positief, maar zwak. Verder bleek de relatie in de Amerikaanse regio het sterkst, daarna in Europa en het kleinste in Azië. Er bleek geen verschil te zijn tussen vakgebieden (rekenen of *science*). De wijze waarop huiswerk werd gemeten, heeft invloed op de sterkte van de relatie. Huiswerkvoltooiing, huiswerkcijfer en huiswerk inspanning hadden een sterkere samenhang met de prestaties dan huiswerkfrequentie en tijd besteed aan huiswerk. De relatie was minder sterk in de onderbouw van het voortgezet onderwijs, dan in het basisonderwijs en de bovenbouw van het voortgezet onderwijs.

Uit deze meta-analyses blijkt dat er geen tot een positieve maar zwakke relatie is tussen huiswerk en (reken)prestaties. De meeste studies waren gebaseerd op correlationele designs, wat een causale interpretatie moeilijk maakt.

Tabel 3.13: Meta-analyses met betrekking tot huiswerk.

Studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Fan et al. (2017) Huiswerk	28 studies (? in PO)	rekenen en science	ES = 0.45	regio: ▪ Verenigde Staten: ES = 0.58 ▪ Europa: ES = 0.24 ▪ Azië: ES = 0.16 schoolniveau: ▪ basisschool: ES = 0.77 ▪ onderbouw VO: ES = 0.30 ▪ bovenbouw VO: ES = 0.63 type huiswerk: ▪ huiswerkvoltooiing: ES = 1.46 ▪ cijfer huiswerk: ES = 1.22 ▪ huiswerk inspanning: ES = 0.65 ▪ huiswerkfrequentie: ES = 0.24 ▪ tijd besteed aan huiswerk: ES = 0.30

3.1.7 Vakintegratie

Naast het aanbieden van rekenen als individueel vak, kunnen de rekenlessen ook geïntegreerd worden in andere lessen, zoals lessen in *science*. Geïntegreerd onderwijs, waarbij rekenen en *science* worden geïntegreerd, had echter geen effect op rekenprestaties bij leerlingen van de kleuterklas tot en met de universiteit, bleek uit een meta-analyse van Hurley (2001) van onderzoeken uit 1935 tot 1997, waarbij voor deze review ook apart de meest recente effectgrootten (1980-1997) zijn nagegaan (zie Tabel 3.14).

Uit de onderzoeken uit 1935 tot 1997 bleek dat vakintegratie kan worden opgedeeld in vijf niveaus die verschillend samenhangen met rekenprestaties. Wanneer *science* namelijk voorafgaand aan rekenen werd aangeboden of andersom (sequentieel niveau), is er wel een groot effect op rekenprestaties. Andere manieren van vakintegratie, zoals het integreren op basis van overeenkomstige onderwerpen (parallel niveau), het deels tegelijk aanbieden van de lesstof (partieel niveau), de twee vakken tegelijk aanbieden terwijl één vak als hoofdvak wordt benadrukt (benadrukt niveau) en complete vakintegratie op basis van gelijkheid (totaal niveau), bleken geen effect te hebben op rekenprestaties. Uit de onderzoeken naar de meest recente effectgrootten (1980-1997) bleek dat er geen effect was van vakintegratie op rekenprestaties. Er is bij deze analyse echter geen opdeling gemaakt in de vijf niveaus.

Tabel 3.14: Meta-analyses met betrekking tot vakintegratie.

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Hurley (2001) vakintegratie	10 studies (1 in PO)[1]	ES = 0.07 (*ns*)	rekenprestaties	studie in PO: ES = -0.13 (*ns*)
	29 studies (? in PO)[2]	ES = 0.27 (*ns*)	rekenprestaties	niveau: ▪ sequentieel niveau: ES = 0.85 ▪ parallel niveau: ES = -0.11 (*ns*) ▪ partieel niveau: ES = 0.13 (*ns*) ▪ benadrukt niveau: ES = 0.17 (*ns*) ▪ totaal niveau: ES = 0.20 (*ns*)

[1] Studies uit 1980-1997
[2] Studies uit 1935-1997

3.1.8 Rekenmethode

De gebruikte lesmethode is vaak onderdeel van debat over het rekenonderwijs (Slavin & Lake, 2008). Naar de effecten van lesmethoden op de rekenprestaties van basisschoolleerlingen is één review gevonden. Slavin en Lake (2008) concludeerden op basis van de mediaan van de gevonden effectgrootten dat rekenmethoden een verwaarloosbaar tot klein effect hebben op rekenprestaties. Dergelijke kleine positieve effecten werden voor diverse soorten rekenmethoden gevonden. Over het algemeen komt uit deze review dus weinig bewijs naar voren voor de stelling dat verschillende rekenmethoden verschillende effecten hebben op rekenprestaties. Uit een vergelijking met andere studies in de review bleek dat de bijbehorende instructiebegeleiding een belangrijkere factor is.

Uit deze review blijkt dat er tussen rekenmethoden geen grote verschillen bestaan in rekenprestaties. Het lijkt erop dat het belangrijker is hoe de leerkracht met een rekenmethode omgaat.

3.1.9 Samenvatting resultaten

Figuur 3.2 is een schematische weergave van de gevonden effecten van de kenmerken de les.

Figuur 3.2: Resultaten factoren op het niveau van de les

	categorie factor(en)		factor(en)
lesinhoud (wat?)	*opportunity to learn* (1)	leerstofaanbod (1)	▪ **aansluiting inhoud curriculum met inhoud toetsen**
			▪ *leerstofinhoud*
		tijd (0)	▪ *instructietijd* ▪ *oefentijd* ▪ *'time-on-task'*

lesvorm (hoe?)	instructie- en werkvormen (15)		▪ **interventie met bepaalde instructie-/werkvorm:**
			  ▪ **expliciete instructie, directe instructie**
			  ▪ **meer begeleid(end)e instructie**
			  ▪ **moderne instructievormen (leerling centraal)**
			  ▪ **innovatieve instructievormen**
			  ▪ **extra oefenopgaven**
			  ▪ **leerlingen zelf vraagstukken laten (her)vormen**
			  ▪ **'Cover-Copy-Compare' methode**
			  ▪ **samenwerkend leren (groepswerk, 'peer-tutoring')**
			  ▪ **aanmoedigen tot 'self explanation'**
			  ▪ **stimuleren zelfregulerend leren**
			▪ **expliciete instructie vs. ontdekkend leren**

			▪ directe instructie vs. begeleide/constructivistische instructie ▪ verschillende rekendidactieken (in NL onderzoek)
			▪ *probleemgestuurd leren* ▪ *voorbeeldgestuurd leren*
	hulpmiddelen (12)	technologisch (8)	▪ **computertechnologie, educatieve applicaties** ▪ **exploratieve digitale leeromgevingen** ▪ **gebruik van mobiele apparaten in de klas**
			▪ computerondersteunde leeromgevingen ▪ video-games
		niet-technologisch (4)	▪ **gebruik drie-dimensionale modellen** ▪ **gebruik concept maps**
inhoud- en vorm-overstijgend	toetsing (3)	summatief (0)	▪ *eindtoetsen*
		formatief (3)	▪ **verschillende vormen van formatieve toetsing** ▪ **gebruik (digitale) leerlingvolgsystemen door leerkracht** ▪ **feedback-per-opgave aan leerling**
			▪ *monitoren eigen prestaties door leerlingen*
	overige factoren (6)		▪ **differentiatie: werken in groepen** ▪ **huiswerk**
			▪ vakintegratie rekenen en 'science' ▪ rekenmethode

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties.
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties.
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

3.2 Factoren op het niveau van de leerkracht

De kenmerken van de leerkracht die als startpunt zijn gehanteerd staan in Figuur 3.3. In de volgende paragrafen bespreken we de gevonden meta-analyses en reviews.

Figuur 3.3. Startpunt factoren op het niveau van de leerkracht.

niveau	categorie factor(en)	factor(en)
leerkracht	kennis en vaardigheden	vakkennis, pedagogische vakkennis, klassenmanagement, helderheid uitleg
	professionele ontwikkeling	professionalisering, professionele leeromgevingen
	overtuigingen	prestatieverwachtingen, visie op (reken)didactiek, professionele motivatie, rekenangst

3.2.1 Kennis en vaardigheden

(Pedagogische) vakinhoudelijke kennis

Om de centrale rol van vakinhoudelijke kennis binnen lesgeven en de relatie daarvan met pedagogiek te benadrukken introduceerde Shulman (1986) het concept pedagogische vakinhoudelijke kennis (*pedagogical content knowledge*, PCK). Hij maakte hierbij onderscheid tussen twee categorieën: (1) kennis van leerstrategieën en representatie van de inhoud van het vak en (2) kennis van de (mis)concepties van de leerlingen. Ondanks het feit dat er in de literatuur verschillende definities worden gebruikt voor pedagogische vakinhoudelijke kennis, komen deze categorieën in vrijwel alle definities terug (Depaepe, Verschaffel, & Kelchtermans, 2013). Andere overeenkomsten zijn dat de kennis van de leerkracht het uitgangspunt is, vakinhoudelijke kennis een vereiste is, er een koppeling gemaakt wordt tussen vakinhoudelijke kennis en pedagogisch handelen en het specifiek is voor een vak. Er zijn ook onderzoekers die menen dat de categorieën van Shulman (1986) onvolledig zijn en dat er categorieën ontbreken. Zo spreken Ball, Thames en Phelps (2008) van een derde categorie die betrekking heeft op curriculumkennis.

Depaepe en collega's (2013) onderzochten hoe pedagogische vakinhoudelijke kennis onderzocht werd in empirisch onderzoek binnen het rekenonderwijs. In de literatuur is een groep studies naar de ontwikkeling van pedagogische vakinhoudelijke kennis en een groep studies naar de relatie tussen pedagogische vakinhoudelijke kennis en leeruitkomsten. Er was een positieve correlatie gevonden tussen de leeruitkomsten van de leerling en de pedagogische vakinhoudelijke kennis van de leerkracht. Daarnaast bleek dat deze correlatie significant hoger was dan de correlatie tussen leeruitkomsten en de vakinhoudelijke kennis van de leerkracht. Depaepe en collega's (2013) beschrijven onderzoeken die uitgevoerd zijn met leerkrachten in opleiding (LIO's) en leerkrachten in dienst. Ondanks dat specifieke verschillen in de review niet werden aangeduid, bleek de pedagogische vakinhoudelijke kennis van LIO's en leerkrachten niet compleet te zijn. Van sommige subdomeinen van rekenen hebben de leerkrachten (in opleiding) meer pedagogische vakinhoudelijke kennis dan van andere subdomeinen.

Leerkrachtvaardigheden

Walshaw en Anthony (2008; 2013) hebben onderzocht wat leerkrachten kunnen doen om een productieve leeromgeving te creëren. Aan de hand van studies die een verband leggen tussen het pedagogisch handelen van de leerkracht en rekenresultaten hebben zij de pedagogiek omschreven die een productieve leeromgeving heeft. Ten eerste was het belangrijk dat leerkrachten geven om de betrokkenheid die leerlingen hebben met rekenen. Ten tweede stelt de leerkracht duidelijke verwachtingen aan de leerling. Dit gaat niet alleen om duidelijke leerdoelen, maar de leerling heeft ook baat bij duidelijke richtlijnen voor zijn of haar gedrag in de klas, via de zogenoemde 'normen van deelname' die beschrijven wanneer en hoe een leerling als spreker of als luisteraar moet deelnemen

aan een klassikale discussie. Ten derde weet de leerkracht de grens van de rekenkennis van de leerlingen en zorgt hij of zij voor een omgeving waarin elke leerling een gelijke deelname heeft, ook de leerlingen met minder rekenkennis. Daarnaast moet er een respectvolle uitwisseling zijn van kennis en ideeën. De leerkracht moet er dus voor zorgen dat elke leerling zich veilig voelt om antwoord te geven en zich niet afgekeurd voelt als hij of zij een rekenconcept verkeerd uitlegt. De leerkrachten hebben ook een belangrijke rol in de communicatie-ontwikkeling van leerlingen (Walshaw & Anthony, 2013). Met klassikale discussies leren leerlingen rekenkundig redeneren. Verder noemen de auteurs de positieve resultaten uit onderzoeken naar werken in kleine groepjes, maar benadrukken dat effectief pedagogisch handelen van de leerkracht zowel klassikaal, als in groepjes, als op individueel niveau invloed heeft. De leerkracht moet balans brengen tussen het werken in de groep en zelfstandig denken.

Korpershoek, Harms, de Boer, van Kuijk en Doolaard (2016) onderzochten het effect van verschillende interventies met betrekking tot klassenmanagementstrategieën op leerprestaties binnen het basisonderwijs (zie Tabel 3.15). Voor het onderzoek hadden de auteurs klassenmanagementinterventies ingedeeld in vier categorieën: (1) gericht op leerkrachtvaardigheden, (2) gericht op de leerkracht-leerling relatie, (3) gericht op gedrag van de leerling en (4) gericht op de ontwikkeling van sociaal-emotionele vaardigheden van de leerling. Alle onderzochte interventies waren gericht op de klas als geheel, terwijl de onderzochte strategieën zowel konden worden toegepast op de klas als geheel, als op een individuele leerling. In alle studies was er sprake van een controlegroep. Uit het onderzoek bleek dat alle klassenmanagementinterventies een klein positief effect hadden op de leerprestaties (en ook op het gedrag, motivatie en de sociaal-emotionele vaardigheden van de leerling). Er was echter sprake van publicatie-*bias*: als er meer studies gevonden waren, was het effect kleiner geweest. Interventies gericht op de relatie tussen leerling en leerkracht hadden het grootste effect op algehele leerprestaties, maar dit was een klein effect. Interventies gericht op sociaal-emotionele vaardigheden van de leerlingen hebben over het algemeen de grootste invloed op de leerling, zowel op leerprestaties als op gedrag, motivatie, sociaal emotionele vaardigheden en op andere niet gespecificeerde factoren, maar dit was een klein effect. Leerlingkenmerken hadden geen invloed op de resultaten, waaruit blijkt dat alle leerlingen gebaat waren bij klassenmanagement interventies.

Samenvattend geven beide onderzoeken aan dat de relatie tussen leerling en leerkracht invloed heeft op de prestaties van de leerling. Er is geen eenduidig beeld voor het creëren van productieve leeromgevingen en in beide onderzoeken wordt benadrukt dat verschillende factoren meespelen in creëren van een omgeving waarin leerlingen optimaal leren en presteren.

Tabel 3.15: Meta-analyses met betrekking tot leerkrachtvaardigheden.

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Korpershoek et al. (2016) klassen-management interventies	54 interventie-studies (32 in PO)	algemene schoolprestaties	ES = 0.17	focus interventie: ▪ gedrag leerkracht: ES = 0.21 ▪ gedrag leerling: ES = 0.18 ▪ sociaal-emotionele ontwikkeling leerling: ES = 0.17 ▪ relatie leerkracht-leerling: ES = 0.24

3.2.2 Professionele ontwikkeling

Professionele ontwikkeling van vakinhoudelijke kennis

Blank en de las Alas (2009) onderzochten in hun meta-analyse het effect van professionele ontwikkelingsactiviteiten, gericht op vakinhoudelijke kennis, op leerprestaties binnen het basis- en voortgezet onderwijs (zie Tabel 3.16). Zij vonden bij de studies gericht op rekenen met een voor- en nameting een klein positief effect van het deelnemen van de leerkracht aan professionele ontwikkelingsactiviteiten op de rekenprestaties van de leerlingen, waarbij het effect groter was bij experimentele studies dan bij quasi-experimentele studies. De rekenstudies met alleen een nameting hadden een kleiner, maar positief effect. Als leerkrachten geen mentorprogramma volgen was het effect groter dan is zij wel een mentorprogramma volgen. Als leerkrachten stages liepen was het effect groter dan als zij dat niet deden.

Tabel 3.16: Meta-analyses met betrekking tot (pedagogische) vakinhoudelijke kennis.

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Blank & De las Alas (2009) professionele ontwikkelings-activiteiten	16 studies (10 in PO)	rekenen en *science*	-	onderzoeksdesign: ▪ voor- en nameting: ES = 0.21 ▪ alleen nameting: ES = 0.13 ▪ gerandomiseerd onderzoek met voor- en nameting en controlegroep: ES = 0.27 ▪ quasi-experimenteel onderzoek met voor- en nameting: ES = 0.17 mentorprogramma: ▪ ja: ES = -0.19 (*ns*) ▪ nee: ES = 0.16 stage: ▪ ja: ES = 0.21 (*ns*) ▪ nee: ES = 0.10

Professionele leergemeenschappen

Een professionele leergemeenschap is een samenwerkingsverband tussen vakmensen om informatie te delen en zo te groeien in hun eigen vakgebied. In termen van het onderwijs is er sprake van onder andere leerkrachten, directeuren en onderwijsadviseurs die kennis delen over hoe ze leerlingen het beste kunnen laten leren.

Mcclendon Patrick (2013) onderscheidde vijf componenten van een professionele leergemeenschap: (1) gedeelde en ondersteunende leiderschap, (2) gedeelde visie, (3) het delen van persoonlijke ervaringen, (4) collectief leren en (5) structurele ondersteunende condities, zoals waar, wanneer, wat en hoe. In hun meta-analyse naar de relatie tussen het implementeren van professionele leergemeenschappen en leerprestaties in rekenen en lezen binnen basisscholen en voortgezet onderwijs nemen zij alleen proefschriften mee waarin de componenten van professionele leergemeenschappen gemeten worden. In de meeste gevallen gebeurde dit met een vragenlijst voor de leerkracht. De overkoepelende meta-analyse over twee studies die een samengestelde score voor deze componenten gebruikten, toonde aan dat wanneer leerkrachtdeelname toenam, de rekenresultaten van de leerlingen verbeterde (zie Tabel 3.17). Dit effect was matig. Met de studies die afzonderlijke componenten beschreven werd voor elk afzonderlijk component een meta-analyses uitgevoerd. Wanneer de professionele leergemeenschappen gedeelde en ondersteunde leiderschap en gedeelde visie hadden, bleek er een matig tot groot positief effect te zijn op de rekenresultaten van de leerlingen in het basisonderwijs. Voor de componenten collectief leren en structurele

ondersteunende condities werd een klein negatief effect gevonden op de rekenresultaten in het basisonderwijs. Een samenvattende meta-analyse over zowel de studies die afzonderlijke componenten beschreven als de studies met een samengestelde score toonde echter geen significant effect van professionele leergemeenschappen op rekenresultaten.

Tabel 3.17: Meta-analyses met betrekking tot professionele leergemeenschappen

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Patrick et al. (2013) deelname leerkracht aan professionele leergemeenschap	13 studies (9 in PO)	rekenprestaties leerlingen basis- en voortgezet onderwijs	*composite* meta-analyse ES = 0.32 *summary* meta-analyse: ES = 0.06 (*ns*)	-
		rekenprestaties leerlingen binnen het basisonderwijs		component professionele leergemeenschap: ▪ gedeelde en ondersteunende leiderschap: ES = 0.23 ▪ gedeelde visie: ES = 0.58 ▪ collectief leren: ES = -0.02 ▪ structurele ondersteunende condities: ES = -0.19

3.2.3 Overtuigingen

Stereotyperingen

Omdat over het algemeen het stigma heerst dat jongens beter zijn in rekenen dan meisjes, hebben Flore en Wicherts (2015) onderzocht wat het effect van *stereotype threat* is op de rekenprestaties van meisjes in het basis- en voorgezet onderwijs (zie Tabel 3.18). Een *stereotype threat* is de negatieve invloed van stereotype gedachten. In hun meta-analyse zijn alleen studies met een controlegroep en een gemanipuleerde *stereotype threat* meegenomen. De gemanipuleerde conditie gaf bijvoorbeeld de toelichting op de test "meisjes presteren over het algemeen slechter op deze test". Uit het onderzoek bleek dat door een *stereotype threat* meisjes over het algemeen slechtere rekenprestaties hadden dan wanneer er geen *stereo threat* was. Dit was een klein effect. Flore en Wicherts (2015) geven echter aan dat er sprake kan zijn van publicatie-*bias*. Leeftijd van de leerlingen bleek geen invloed te hebben op de invloed van *stereotype threat.*

Tabel 3.18: Meta-analyses met betrekking tot stereotyperingen.

studie	aantal studies (aantal in PO)	uitkomstmaat	effectgrootte	uitgesplitst naar
Flore & Wicherts (2015) *stereotype threat*	26 studies (16 in PO)	rekenprestaties	ES = -0.22	-

3.2.4 Samenvatting resultaten

Figuur 3.4 is een schematische weergave van de gevonden effecten van de kenmerken de leerkracht.

Figuur 3.4. Resultaten factoren op het niveau van de leerkracht.

niveau	categorie factor(en)	factor(en)
leerkracht	kennis en vaardigheden (3)	▪ **vakinhoudelijke kennis** ▪ **pedagogische vakinhoudelijke kennis** ▪ **klassenmanagementvaardigheden** ▪ **pedagogisch handelen (creëren productieve leeromgeving)**
		▪ *helderheid uitleg*
	professionele ontwikkeling (2)	▪ **professionalisering (vakinhoudelijke kennis)** ▪ **professionele leergemeenschappen**
	overtuigingen (1)	▪ stereotypering (naar geslacht leerling)
		▪ *prestatieverwachtingen* ▪ *visie op (reken)didactiek* ▪ *professionele motivatie* ▪ *rekenangst*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

3.3 Factoren op het niveau van de klas

De kenmerken van de klas die als startpunt zijn gehanteerd staan in Figuur 3.5. Over de resultaten met betrekking tot kenmerken van de klas zijn geen relevante meta-analyses of reviews gevonden, zoals duidelijk wordt uit Figuur 3.6.

Figuur 3.5: Startpunt factoren op het niveau van de klas.

niveau	categorie factor(en)	factor(en)
klas	sfeer	werkklimaat, veiligheid, relaties, ondersteuning, regels, prestatiegerichtheid

Figuur 3.6: Resultaten factoren op het niveau van de klas.

niveau	categorie factor(en)	factor(en)
klas	sfeer (0)	▪ *werkklimaat* ▪ *veiligheid* ▪ *relaties* ▪ *ondersteuning* ▪ *regels* ▪ *prestatiegerichtheid*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

3.4 Factoren op het niveau van de school

De kenmerken van de school die als startpunt zijn gehanteerd staan in Figuur 3.7.

Figuur 3.7: Startpunt factoren op het niveau van de school.

niveau	categorie factor(en)	factor(en)
school	rekenbeleid	rekencoördinator, belang van rekenen binnen de school
	evaluatie van het (reken)leerproces	monitoren van prestaties, inzicht (reken)behoeftes leerlingen, opbrengstgericht werken
	klimaat	structuur

Blok en collega's (2015) verrichtten een *quick scan* van de onderzoeksliteratuur naar de effectiviteit van twee verschillende vormen van evaluatie van het onderwijs, de leerkracht of de school op basis van summatieve/gestandaardiseerde toetsgegevens. Eén vorm is *standards-based* schoolverbetering die scholen onderwerpt aan *accountability*-programma's waarin zij verplicht worden tot afname van toetsen en rapportage van de resultaten. Vaak zijn hier standaarden aan verbonden, die het mogelijk maken uitspraken te doen over goede en slechte scholen. De effectiviteit van deze vorm van evaluatie is omstreden en heeft bovendien een aantal negatieve bijeffecten zoals onzuivere toetspraktijken en verarming van het curriculum. Een tweede vorm van evaluatie is datagestuurde schoolverbetering, ook wel schoolprestatiefeedback genoemd, waarin vorderingsgegevens van leerlingen (en ook andere gegevens) benut kunnen worden voor onderwijsverbetering. Dit onderscheidt zich van *standards-*

based schoolverbetering doordat er in principe geen sprake is van sancties waardoor de kans op negatieve bijeffecten kleiner is. De werkzaamheid van datagestuurde schoolverbetering is echter nog niet overtuigend aangetoond (Blok et al., 2015). Vandaar dat in Figuur 3.8 de factoren *accountability*-programma's en schoolprestatiefeedback bij de factoren staan waarover geen meta-analyses of reviews zijn gevonden: Blok en collega's (2015) bepleiten vooral dat er onvoldoende overtuigend bewijs voor de werkzaamheid is. Over de resultaten met betrekking tot overige kenmerken van de school zijn geen relevante meta-analyses of reviews gevonden.

Figuur 3.8. Resultaten factoren op het niveau van de school.

niveau	categorie factor(en)	factor(en)
school	rekenbeleid (0)	▪ *rekencoördinator* ▪ *belang van rekenen binnen de school*
	evaluatie van het (reken)leerproces (1)	▪ *monitoren van prestaties* ▪ *'accountability' –programma's* ▪ *schoolprestatiefeedback* ▪ *inzicht (reken)behoeftes leerlingen* ▪ *opbrengstgericht werken*
	klimaat (0)	▪ *structuur*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

3.5 Factoren op het niveau van de leerling

De kenmerken van de leerling die als startpunt zijn gehanteerd staan in Figuur 3.9, die alle vallen onder leerlingovertuigingen. Overtuigingen van leerlingen vormen een breed begrip. Met betrekking tot rekenen kan het bijvoorbeeld gaan om zelfvertrouwen, gevoelens van controle en invloed en overtuigingen hoe je rekenen het beste zou kunnen leren (Muis, 2004). Behalve voor rekenangst, waar later op zal worden ingegaan, zijn er geen meta-analyses of reviews gevonden naar de effecten van specifieke overtuigingen. Wel zijn er overkoepelende reviews gevonden waarin de relatie tussen allerlei overtuigingen van leerlingen enerzijds en hun rekenprestaties anderzijds aan bod komt.

Figuur 3.9: Startpunt factoren op het niveau van de leerling.

niveau	categorie factor(en)	factor(en)
leerling	overtuigingen	motivatie, *self-efficacy*, attitude, rekenangst

3.5.1 *Beliefs* over rekenen

Uit de review van Muis (2004) blijkt dat leerlingen vaak (naïeve) epistemologische *beliefs* over rekenen hebben die hun rekenprestaties negatief beïnvloeden. Een voorbeeld is dat leerlingen die geloven dat rekenkennis bestaat uit losse stukjes informatie lager presteren dan leerlingen die geloven dat er samenhang bestaat in rekenkennis. Gedrag is een belangrijke mediator tussen overtuiging en leeruitkomsten. Wanneer een leerling er bijvoorbeeld van overtuigd is dat hij of zij snel een bepaalde taak kan oplossen, zonder daarbij rekening te houden met de complexiteit van de taak, zal de leerling niet snel geneigd zijn om er meer tijd dan gepland aan te besteden. Een tegenovergestelde overtuiging die leerlingen kunnen hebben, is dat een taak meer een proces is waarbij regelmatig de geplande tijd bijgesteld moet worden, afhankelijk van in hoeverre de leerling de stof begrijpt.

3.5.2 Attributie

Shores en Smith (2010) onderzochten de relatie tussen attributie en rekenprestaties. Attributie is de mate waarin het behaalde resultaat wordt toegeschreven aan externe en interne factoren. Uit de review blijkt dat leerlingen een positief of negatief resultaat voornamelijk toeschrijven aan hun individuele bekwaamheid en de ingestoken moeite. Kinderen zijn rond hun negende jaar bewust van het concept bekwaamheid. Leerlingen die geloven dat externe factoren meer invloed hebben op hun resultaten, zoals geluk of de moeilijkheidsgraad van de toets, halen gemiddeld lagere resultaten dan leerlingen die denken dat interne factoren een grotere rol spelen. Het is echter nog onduidelijk hoe attributie van intern naar extern veranderd kan worden. De auteurs concludeerden dat leerkrachten hierin een belangrijke rol kunnen spelen, met name voor rekenen. Door feedback te geven gericht op attributie tijdens klassikale lessen zouden leerkrachten de attributie van leerlingen kunnen beïnvloeden. Daarnaast geeft inzicht in attributie de leerkracht ook meer inzicht in het gedrag van de leerlingen.

3.5.3 Rekenangst

Rekenangst kan omschreven worden als een gevoel van spanning of vrees bij het maken van rekentaken (Ashcraft & Kirk, 2001). Rekenangst kan variëren van milde frustratie tot heftige emotionele reacties en kan zowel plaatsvinden op school als in alledaagse situaties (Ashcraft & Moore, 2009). Er zijn vier reviews gevonden waarin ingegaan wordt op de relatie tussen rekenangst en rekenprestaties.

Ashcraft en Moore (2009) beschrijven in hun review dat uit onderzoek is gebleken dat er een negatieve samenhang bestaat tussen rekenangst en rekenprestaties. Er was geen samenhang tussen rekenangst en intelligentie. Verder lijkt het erop dat kinderen in de eerste jaren van het basisonderwijs nog geen last van rekenangst hebben (of er in elk geval geen blijk van geven), maar dat rekenangst bij leerlingen zich meestal uit vanaf groep 6 of 7.

Ook Carey, Hill, Devine en Szücs (2016) beschrijven dat er een negatieve samenhang bestaat tussen rekenangst en rekenprestaties. Er was zowel bewijs gevonden dat rekenangst leidt tot slechte rekenprestaties als andersom. Het zou zo kunnen zijn dat voor sommige leerlingen slechte prestaties rekenangst kunnen uitlokken waarna er een vicieuze cirkel ontstaat waarbij rekenangst de prestaties nog verder verslechtert.

Chang en Beilock (2016) gaan in hun review eveneens in op de achterliggende oorzaken van de negatieve relatie tussen rekenangst en rekenprestaties. Zowel leerlingfactoren (zoals werkgeheugen, emotieverwerking en motivatie) als omgevingsfactoren (zoals attitudes van leerkrachten en ouders) lijken van invloed te zijn op deze relatie. Verschillende soorten interventies, zoals extra bijles, therapie en computerondersteuning, zouden rekenangst kunnen verminderen.

Tenslotte beschrijven Foley en collega's (2017) dat rekenangst een globaal fenomeen is. Het lijkt er op dat rekenangst zowel een oorzaak als gevolg kan zijn van slechte rekenprestaties. Factoren zoals de attitudes en ondersteuning van leerkrachten en ouders, sociale druk en stereotypering kunnen van invloed zijn op de ontwikkeling van rekenangst. Leerlingen met rekenangst zouden baat kunnen hebben bij psychologische ondersteuning, bijvoorbeeld voor emotieregulatie. Ook zouden interventies gericht op de interactie tussen ouders of leerkrachten met rekenangst en leerlingen positieve effecten kunnen hebben.

Samenvattend blijkt uit veel onderzoek dat er een negatieve relatie bestaat tussen rekenangst en rekenprestaties. De beschreven studies hadden zich vooral gericht op de mechanismen achter deze relatie. Het lijkt erop dat rekenangst kan leiden tot slechte rekenprestaties, maar dat slechte rekenprestaties ook kunnen leiden tot rekenangst. Er zijn diverse leerlingfactoren en

omgevingsfactoren die deze relatie beïnvloeden. Verschillende soorten interventies zouden kunnen helpen om rekenangst te verminderen.

3.5.4 Samenvatting van de resultaten

Figuur 3.10 is een schematische weergave van de gevonden effecten van de kenmerken van de leerling.

Figuur 3.10: Resultaten factoren op het niveau van de leerling

niveau	categorie factor(en)	factor(en)
leerling	overtuigingen (6)	▪ rekenangst ▪ externe (vs interne) attributie resultaten ▪ naïeve 'beliefs' over rekenen
		▪ *motivatie* ▪ *self-efficacy* ▪ *attitude*

Vetgedrukte factoren in het groen hadden een positieve samenhang met rekenprestaties
Onderstreepte factoren in het rood hadden een negatieve samenhang met rekenprestaties
Factoren in het grijs hadden nauwelijks tot geen samenhang met rekenprestaties.
Over de *cursief gedrukte factoren* is geen meta-analyse of review gevonden met de huidige zoekmethode.
Tussen haakjes staan de aantallen meta-analyses of reviews per categorie factor(en).

4 Onderzoeken met Nederlandse basisschoolleerlingen

Er is een brede verscheidenheid aan onderzoeken met Nederlandse basisschoolleerlingen gevonden. Deze hebben we gecategoriseerd in onderzoek naar specifieke reken(verbeter)programma's (2 studies), digitale rekenspellen en oefenplatforms (9 studies), trainingen die aansturen op het noteren van uitwerkingsstappen (2 studies), professionaliseringsactiviteiten gericht op differentiatie (1 studie), formatief toetsen (8 studies) of instructievaardigheden door teamleren (1 studie), schoolvisie (1 studie) en leerlingovertuigingen (3 studies).

4.1 Reken(verbeter)programma's

4.1.1 Onderwijsproject *Kwaliteitsversterking rekenen en wiskunde*

Houtveen, Van de Grift, en Creemers (2004) onderzochten het effect van het onderwijsproject "Kwaliteitsversterking rekenen en wiskunde". Dit experimentele programma had als doel de prestaties bij rekenen in groep 3 te verbeteren door in te zetten op adaptieve instructie. Het programma en de implementatie speelde op verschillende niveaus en bevatte verschillende factoren van het onderwijsleerproces. De school(directeur) en de leerkrachten participeerden in een intensief professionaliseringstraject, met het gehele team. Ook werd een coördinator aangesteld om de implementatie van het programma te ondersteunen. In de les richtte het programma zich op adaptieve instructie. Dit bestond uit elementen van (1) formatieve assessment (volgen en toetsen) om de leerbehoeften van kinderen in kaart te brengen, (2) specifieke instructiekenmerken (kwaliteit van de instructie, het optimaliseren van instructietijd en het creëren van succeskansen om zelfvertrouwen van de leerlingen te verhogen) en (3) elementen om actief leren te ondersteunen (zelfregulerend leren en het creëren van een "exploratieve leeromgeving", waarin leerlingen met moeilijkheden verlengde instructie kregen, terwijl de anderen verantwoordelijk waren voor hun eigen leerproces).

De implementatie op 14 scholen (237 leerlingen) verliep volgens een geplande veranderingsstrategie. Pas zodra het programma voldoende geïmplementeerd was op de deelnemende scholen, na 2-3 jaar, werd het evaluatieonderzoek uitgevoerd. Middels een voor- en nameting werd de ontwikkeling van de rekenscores (LVS-toetsen) gevolgd, en vergeleken met een random geselecteerde controlegroep (15 scholen, 311 leerlingen). Onderzoek naar de implementatie van het programma liet zien dat op de tien programmakenmerken de experimentele groep gemiddeld hoger scoorde dan de controlegroep, en dat de lespraktijk dus in de beoogde richting verschilde.

Op de nameting waren de leerlingprestaties van de experimentele groep hoger dan van de controlegroep (ES = 0.52, gecorrigeerd voor de voormeting, leeftijd, geslacht, intelligentie, SES en etniciteit). Ook waren er minder leerlingen met rekenproblemen in de experimentele groep dan in de controlegroep. Van de variantie in rekenprestaties lag 15 procent op het niveau van de school en 85 procent bij de individuele leerlingen. Ongeveer een kwart van de tussenscholenvariantie werd verklaard door de conditie (wel of niet het programma). Vijf van de tien onderzochte lesfactoren hingen positief samen met rekenprestaties. Dit waren twee elementen van formatieve assessment: het identificeren van de leerproblemen van de leerlingen en het implementeren van leerplannen voor risicoleerlingen. Verder was het één instructiekenmerk, namelijk de mate van "uitgebreide directe instructie": directe instructie voor gestructureerde taken en cognitieve strategie-instructie (expert-*scaffolding*) voor minder gestructureerde taken. Ten slotte hingen beide elementen van actief leren positief samen met rekenprestaties: het ondersteunen van het zelfvertrouwen van leerlingen door succeservaringen en het creëren van de eerder genoemde exploratieve leeromgeving.

Samenvattend, een experimenteel programma gericht op adaptieve instructie (bestaand uit elementen van formatieve assessment, specifieke instructiekenmerken en actief leren), en geïmplementeerd via een geplande veranderingsstrategie met een ruime implementatietijd (2-3 jaar) en intensief en

duurzaam ondersteund vanuit de directie en externe experts, had een matig positief effect op de rekenprestaties van leerlingen uit groep 3. Factoren in de les die positief samenhingen met de rekenprestaties kwamen uit alle drie de domeinen van adaptieve instructie. Echter, de helft van de factoren (waaronder ook het optimaliseren van instructietijd, terwijl daar wel een groot verschil tussen experimentele en controlegroep zat) hing niet met de rekenprestaties samen. De factor met de grootste samenhang met rekenprestaties was het diagnosticeren van de leerproblemen, maar dat was enkel op basis van natuurlijke variatie tussen de scholen en niet op basis van het wel of niet volgen van het experimentele programma.

4.1.2 Methodiek *Zo leer je kinderen rekenen*

Bosman (2011) onderzocht de effectiviteit van de op directe instructie gebaseerde methodiek *Zo leer je kinderen rekenen* op de basisvaardigheden (automatiseren optellen, aftrekken, vermenigvuldigen en delen). De kenmerken van de methodiek zijn (1) een systematische opbouw waarbij pas verder wordt gegaan als de voorwaardelijke vaardigheden zijn beheerst, (2) aanbieden van één strategie, (3) herhaling bij het oefenen, (4) groepsgewijze directe instructie waarbij de leerkracht als model fungeert door voor te doen en leerlingen dat nadoen, en (5) gebruik van de bal die naar leerlingen wordt toegegooid wanneer zij de opdracht krijgen tijdens het klassikaal oefenen.

Een groep van 50 leerlingen uit groep 5-8 (van 4 scholen uit het regulier basisonderwijs en 1 school uit het speciaal basisonderwijs) deed mee. De leerlingen zijn door de leerkrachten geselecteerd, en de meesten hadden enige of substantiële achterstand in hun basisvaardigheden. De leerlingen volgden 30 lessen van ongeveer 75 minuten volgens deze methodiek. Het is niet duidelijk of deze lessen bovenop of in plaats van de reguliere rekenlessen plaatsvonden. Er was ook geen controlegroep. Alle leerlingen gingen vooruit in hun basisvaardigheden van de voormeting naar de nameting. De gemiddelde leerwinst was 1.5 leerjaar, maar de variatie in vooruitgang was groot. Daarbij maakte het niet uit of de leerling bij aanvang een achterstand had of niet.

Samenvattend heeft het zes weken dagelijks toepassen van de methodiek *Zo leer je kinderen rekenen,* gekenmerkt door systematische opbouw, herhaling bij het oefenen, aanbieden van één strategie en groepsgewijze directe instructie, een positief effect op de basisvaardigheden (automatiseren optellen, aftrekken, vermenigvuldigen en delen) van kinderen uit de bovenbouw van het regulier en speciaal basisonderwijs die vrijwel allemaal een achterstand hadden in de basisvaardigheden. Daarbij moet opgemerkt worden dat de methodologie van de studie (geen controlegroep, onduidelijk of er meer lessen dan in het reguliere rekenonderwijs gegeven werden) het moeilijk maken te concluderen hoe effectief het programma nu echt is.

4.1.3 Samenvatting reken(verbeter)programma's

Deze twee studies laten zien dat de inzet van doelgerichte reken(verbeter)programma's effectief kan zijn, als de implementatie heel sterk wordt ondersteund (*Kwaliteitsversterking rekenen en wiskunde*) en/of het programma intensief wordt ingezet (*Zo leer je kinderen rekenen*). Omdat de programma's sterk verschilden qua doelstellingen en uitgangspunten (zoals actief leren in *Kwaliteitsversterking rekenen en wiskunde* versus groepsgewijze directe instructie in *Zo leer je kinderen rekenen*), en ook in methodologische aspecten van de onderzoeken, is het niet eenvoudig om algemene werkzame elementen te identificeren.

4.2 Rekenspellen en oefenplatforms op de computer

4.2.1 *Hit the target*

Kolovou (2011) onderzocht voor haar proefschrift onder andere de effecten van het spelen van een computerspel op de prestaties van leerlingen op niet-routinematige, puzzelachtige rekenopgaven. Bij dit soort opgaven is niet de uiteindelijke berekening het lastigste, maar het uitzoeken van de relaties

tussen getallen. Het spel dat gebruikt werd in het onderzoek, genaamd *Hit the Target,* is een interactief pijl-en-boogspel op de computer. In het spel kunnen leerlingen spelenderwijs leren dat variabelen met elkaar kunnen samenhangen: wanneer één variabele verandert, veranderen daarmee ook de andere variabelen. In het geval van dit spel gaat het om het aantal geschoten pijlen (zowel pijlen die het doel raken, als die het doel missen), de spelregels (hoeveel punten er worden toegekend of afgetrokken voor het raken of missen van het doel) en het totale aantal scorepunten. Een voorbeeldopgave is bijvoorbeeld dat leerlingen weten hoeveel punten er met iedere pijl te winnen of te verliezen zijn en wat de eindscore is, waarbij opgelost moet worden hoeveel pijlen het doel geraakt en gemist hebben. In het spel kunnen het aantal pijlen en de spelregels zowel automatisch als door de leerling worden aangepast. Leerlingen kunnen de consequenties van hun acties zien via de resultaten op het scorebord.

Om de effecten van het spel te analyseren waren verschillende studies opgezet. Eén van deze studies betrof een experiment waarin in totaal 236 leerlingen uit groep 8 van 10 verschillende scholen hadden deelgenomen (Kolovou, Van den Heuvel-Panhuizen, & Köller, 2013). De onderzoekers hadden eerst gezocht naar vijf scholen die wilden deelnemen aan het experiment, waarna vijf vergelijkbare scholen benaderd zijn voor de controlegroep. De experimentele groep volgde de interventie gedurende drie weken. Aan het begin van de eerste week werd een klassikale instructie door de onderzoeker gegeven, waarin uitgelegd werd hoe het spel werkt. In de tweede en derde week werd klassikaal ingegaan op de oplossingen van de problemen waaraan de leerlingen hadden gewerkt. Het spel zelf werd echter thuis (vrijwillig) gespeeld. Leerlingen konden in totaal aan 8 problemen (met subvragen) werken waarbij de antwoorden op een antwoordvel moesten worden ingevuld. Het gebruik van het spel (het aantal speelsessies en het aantal gebruikte pijlen) werden automatisch geregistreerd. De controlegroep volgde de interventie niet. De voor- en nameting bestond uit opgaven waarbij getallen met elkaar samenhingen.

Uit de resultaten bleek dat de experimentele groep significant hoger scoorde op de nameting dan de controlegroep (ES = 0.31), waarbij gecontroleerd werd voor de scores op de voormeting, algemene rekenvaardigheid en geslacht. Verder bleek dat 78% van de leerlingen in de experimentele groep het spel minstens één keer hadden gespeeld. Daarom is een onderscheid gemaakt binnen de experimentele groep tussen leerlingen die wel en niet ingelogd hebben in het spel. Leerlingen die wel hadden ingelogd, presteerden beter op de nameting dan leerlingen in de controlegroep (ES = 0.35). Er was geen significant verschil tussen leerlingen in de experimentele groep die niet hadden ingelogd en leerlingen in de controlegroep. Tenslotte bleek het niet uit te maken hoe intensief leerlingen met het spel oefenden.

Een andere studie (Van den Heuvel-Panhuizen, Kolovou, & Robitzsch, 2013) had een vergelijkbare onderzoeksopzet als het hierboven beschreven onderzoek, alleen participeerden hierin leerlingen uit groepen 6, 7 en 8 en volgden al deze leerlingen de interventie. De oorspronkelijke steekproef bestond uit 318 leerlingen van vijf verschillende scholen, maar omdat leerlingen die niet minstens één keer hadden ingelogd in het spel niet werden meegenomen in de analyses, zijn de resultaten uiteindelijk gebaseerd op 253 leerlingen. Uit een vergelijking tussen de scores op de voor- en nameting bleek dat leerlingen uit alle groepen significant vooruit waren gegaan, waarbij de vooruitgang het grootst was voor leerlingen in groep 8 (groep 6: ES = 0.24, groep 7: ES = 0.24, groep 8: ES = 0.37). Uit analyses naar het strategiegebruik van leerlingen bleek dat het leerlingen die extreme waarden gebruikten (de waarde 0 toekennen aan één van de onbekende variabelen) en leerlingen die *trial-and-error* gebruikten de meeste vaardigheidsgroei hadden (respectievelijk ES = 0.52 en ES = 0.29). Voor de groep als geheel had de mate van gebruik van het spel (een combinatie van oefenintensiteit en strategiegebruik) geen significante effecten op de vaardigheidsgroei.

Uit deze twee studies kan geconcludeerd worden dat het werken met het spel *Hit the Target* leerlingen uit groepen 6, 7 en 8 zou kunnen helpen bij het leren oplossen van niet-routinematige,

puzzelachtige rekenopgaven. Hoewel leerlingen profijt hadden van minstens één keer inloggen, leek de verdere oefenintensiteit minder van belang. Tenslotte bleek de effectiviteit van verschillende soorten strategieën te verschillen.

4.2.2 Mini-games

Binnen het promotieonderzoek van Bakker (2014) zijn verschillende studies opgezet naar de effecten van het spelen van mini-games op de multiplicatieve vaardigheden van leerlingen. Mini-games hebben een korte tijdsduur en zijn snel te leren. In de studies van Bakker zijn mini-*games* op Rekenweb als basis gebruikt, die vervolgens gedeeltelijk zijn aangepast (onder andere aanpassingen in de moeilijkheidsgraad en het toevoegen van een scoringsmechanisme). Leerlingen konden in de verschillende mini-games zowel nieuwe kennis opdoen via exploratie als bestaande rekenfeiten en rekenstrategieën automatiseren. Er zijn in de studies steeds verschillende condities met elkaar vergeleken: het spelen van de mini-games op school tijdens de les, het thuis spelen van de mini-games in combinatie met een nabespreking op school, het thuis spelen van de mini-games zonder nabespreking op school en ter controle het op school spelen van mini-games over andere rekenonderwerpen. Alle participerende leerkrachten werd gevraagd om de totale lestijd per rekenonderwerp hetzelfde te houden als voorafgaand aan het onderzoek. Binnen elke conditie waren er vier speelperioden van ieder tien weken (waarvan twee perioden in groep 4 en twee perioden in groep 5). Voor de voor- en nameting maakten de leerlingen online toetsen met multiplicatieve opgaven, bestaande uit een combinatie van kale sommen, contextopgaven en inzichtopgaven.

De eerste relevante studie binnen het kader van de huidige review betrof een studie naar de effecten van het spelen van de mini-games op de vaardigheid in het oplossen van multiplicatieve opgaven en het inzicht in multiplicatieve getalrelaties van leerlingen in groep 4 (Bakker, Van den Heuvel-Panhuizen, Van Borkulo, & Robitzsch, 2013). Uiteindelijk hebben 46 scholen (die random waren toegewezen aan één van de beschreven condities) met in totaal 1005 leerlingen de interventie voor meer dan de helft toegepast. Alleen deze scholen zijn meegenomen in de resultaten. Wanneer de resultaten van alle experimentele condities samen werden onderzocht, werd er geen effect gevonden op de multiplicatieve vaardigheden van de leerlingen. Daarnaast zijn alle experimentele condities ook apart onderzocht. Het thuis spelen van de mini-games in combinatie met een nabespreking op school bleek in vergelijking met de controlegroep een marginaal significant effect te hebben.

Ook de effecten op langere termijn zijn onderzocht (Bakker, Van den Heuvel-Panhuizen, & Robitzsch, 2015). Alleen scholen die tot het einde bij het project betrokken bleven en minstens de helft van de games behandelden, werden in deze uiteindelijke analyses meegenomen. Dit waren 35 scholen met in totaal 719 leerlingen die werden gevolgd vanaf het einde van groep 3 tot en met het einde van groep 5. Het thuis spelen van de mini-games in combinatie met het nabespreken op school bleek het meest effectief. Er werden positieve effecten gevonden op de vaardigheden van leerlingen in het oplossen van multiplicatieve opgaven en hun inzicht in multiplicatieve getalrelaties (significante effectgrootten uitgedrukt in ES lagen tussen 0.22 en 0.29). Alleen leerlingen in groep 4 hadden baat bij het uitsluitend spelen van mini-games op school en dan alleen voor hun inzicht in multiplicatieve getalrelaties (ES = 0.35). Het thuis spelen van mini-games zonder nabespreking op school had geen effect. De mate van gebruik van de mini-games had soms effect, maar deze resultaten waren niet eenduidig.

Samenvattend lijkt het thuis spelen van mini-games in combinatie met een klassikale nabespreking op school een positief effect te kunnen hebben op de multiplicatieve vaardigheden van leerlingen in groepen 4 en 5. Er zijn geen eenduidige effecten wat betreft de mate van gebruik van de mini-games gevonden.

4.2.3 *Snappet*

Snappet is een adaptief onderwijsplatform dat wordt aangeboden via tablets. Leerlingen kunnen via dit platform oefenopgaven maken die vergelijkbaar zijn met opgaven in een traditioneel werkboek. Kenmerkend voor Snappet is dat leerlingen directe feedback krijgen of het gegeven antwoord goed of fout is, dat de complexiteit van de opgaven automatisch kan worden aangepast aan het niveau van de leerling, dat leerlingen door middel van sterren inzicht krijgen in de voortgang op hun leerdoelen en dat ook leerkrachten direct inzicht hebben in de voortgang van zowel individuele leerlingen als de gehele klas (Faber & Visscher, 2016; Molenaar, Van Campen, & Van Gorp, 2015). Naar de effecten van het gebruik van Snappet op rekenprestaties zijn twee onderzoeksrapporten gevonden.

Molenaar en collega's (2015) hebben drie deelonderzoeken uitgevoerd, waarbij binnen twee onderzoeken de relatie tussen het gebruik van Snappet en rekenprestaties gelegd wordt. Het eerste deelonderzoek was quasi-experimenteel van opzet waarbij scholen zelf hebben gekozen om Snappet toe te passen in de les. Er participeerden 350 leerlingen uit groepen 4 en 6 van een onbekend aantal scholen (187 leerlingen in de experimentele groep en 163 leerlingen in de controlegroep). De experimentele groep werkte een jaar lang met Snappet terwijl de controlegroep dit niet deed en alleen de reguliere rekeninstructie toepaste. Aan de hand van Cito-toetsen zijn de rekenprestaties van leerlingen vastgesteld. Bij leerlingen in groep 4 bleek er tussen de experimentele groep en de controlegroep geen verschil te zijn in de groei van de rekenprestaties. Er bleek tussen de condities ook geen verschil te zijn in de rekenvaardigheidsgroei van leerlingen van verschillende niveaus. Voor leerlingen in groep 6 werd er wel een verschil in rekenvaardigheidsgroei gevonden, waarbij leerlingen die oefenden met Snappet sterker groeiden dan leerlingen die niet met Snappet oefenden. Verder bleek dat leerlingen in groep 6 van verschillende niveaus in de experimentele groep even sterk groeiden, terwijl leerlingen van verschillende niveaus in de controlegroep niet even sterk groeiden. Bij deze leerlingen was er sprake van convergerende prestaties waarbij de verschillen in niveaus dus steeds kleiner werden.

Het andere relevante deelonderzoek van Molenaar en collega's was gericht op de relatie tussen hoe Snappet gebruikt wordt en rekenprestaties. Gebruik werd hierbij gedefinieerd als het aantal klassikale en adaptieve opdrachten dat werd gemaakt. In dit onderzoek participeerden in totaal 636 leerlingen uit groepen 4, 5, 6 en 7 van een onbekend aantal scholen, waaronder ook de leerlingen uit het eerste deelonderzoek. Uit de resultaten van groep 4 bleek dat zowel het maken van meer klassikale opdrachten als meer adaptieve opdrachten leidde tot een grotere groei in rekenprestaties. Voor groep 5 en voor groep 6 bleek er geen relatie te bestaan tussen het aantal gemaakte opdrachten en groei in rekenprestaties. Tenslotte bleek voor groep 7 dat zowel het maken van meer klassikale opdrachten als meer adaptieve opdrachten van positieve invloed was op de groei in rekenprestaties. De verschillen werden dus alleen voor groep 4 en 7 gevonden, waarbij het verder niet uitmaakte of kinderen meer oefenden met klassikale opdrachten of met adaptieve opdrachten.

Faber en Visscher (2016) richtten zich in hun studie op leerlingen uit groep 5. Ze onderzochten 79 scholen met in totaal 1808 leerlingen. Scholen werden random toegewezen aan de experimentele groep (40 scholen) of de controlegroep (39 scholen). De experimentele groep gebruikte een half jaar lang Snappet, terwijl de controlegroep Snappet in deze periode niet mocht gebruiken. Om leerkrachten te kunnen ondersteunen in het gebruik van Snappet konden zij een basistraining en vervolgtraining volgen en hadden zij de beschikking over een coach. Om de rekenprestaties van de leerlingen aan het begin en aan het eind van het onderzoek vast te stellen zijn de scores op de Cito LVS-toetsen rekenen-wiskunde opgevraagd. Verder zijn in de experimentele groep bij een aantal leerkrachten lesobservaties uitgevoerd om te analyseren in hoeverre zij de voortgangsinformatie uit Snappet verder gebruikten voor differentiatie in de les.

Uit de studie van Faber en Visscher bleek dat het gebruik van Snappet over een periode van een half jaar positieve effecten had op de rekenprestaties van leerlingen uit groep 5. Alle leerlingen profiteerden van Snappet, maar de effecten waren het sterkst voor de 20 procent hoogst scorende leerlingen. Verder bleek dat wanneer leerkrachten de voortgangsinformatie uit Snappet veel gebruikten om te differentiëren in de les, leerlingen hoger scoorden dan wanneer leerkrachten Snappet minder gebruikten voor differentiatie. Tenslotte bleek ook oefenintensiteit nog van belang. Leerlingen die meer oefenopgaven in Snappet maakten scoorden hoger op rekenen dan leerlingen die minder oefenopgaven maakten.

Op basis van de twee bovenstaande onderzoeksrapporten kunnen de volgende conclusies getrokken worden. In beide onderzoeken werd Snappet voor langere tijd (een half jaar tot een jaar) gebruikt. Bij leerlingen uit groep 4 leek Snappet geen effect te hebben op rekenprestaties. Bij leerlingen in groepen 5 en 6 had Snappet wel positieve effecten op rekenprestaties. Uit onderzoek naar de oefenintensiteit van Snappet bleek dat leerlingen in groepen 4 en 7 profiteerden van het maken van meer oefenopgaven, waarbij het niet uitmaakte of dat klassikale of adaptieve opdrachten waren. Voor leerlingen in groep 5 werden wat betreft oefenintensiteit tegenstrijdige resultaten gevonden.

4.2.4 Rekentuin

Jansen, Hofman, Savi, Visser en Van der Maas (2016) hebben onderzoek gedaan naar de effecten van succesfrequenties in het spel *Rekentuin*. Rekentuin is een adaptief, online spel waarin de focus ligt op het oefenen van bewerkingen. De oefenopgaven in het spel kunnen automatisch worden aangepast aan het niveau van de leerling die het speelt. Een belangrijk aspect hierbij is de succesfrequentie. De succesfrequentie is het percentage goede antwoorden dat de leerling gemiddeld geeft. De moeilijkheid van de opgaven die de leerling krijgt kan aan dit percentage worden aangepast. De succesfrequentie zou van invloed kunnen zijn op de rekenmotivatie en uiteindelijke rekenprestaties van leerlingen. Er is in deze studie voor Rekentuin onder andere onderzocht of het wel of niet zelf kunnen kiezen van een succesfrequentie van invloed is op rekenprestaties en of een hogere succesfrequentie leidt tot hogere rekenprestaties.

In het onderzoek participeerden in totaal 192 leerlingen uit groepen 5, 6, 7 en 8 afkomstig van één basisschool. Door middel van loting werden de leerlingen toegewezen aan één van de volgende vier condities: zelf mogen kiezen van een succesfrequentie, hoge succesfrequentie (ongeveer 90 procent goed), middelmatige succesfrequentie (ongeveer 75% goed) of lage succesfrequentie (ongeveer 60 procent goed). De Tempo Toets Automatiseren werd als voor- en nameting afgenomen om de rekenvaardigheid van de leerlingen vast te stellen. De totale *time on task* lag voor de verschillende condities gemiddeld tussen 1,4 en 1,9 uur.

Uit de resultaten van de studie bleek dat de rekenprestaties van alle leerlingen significant verbeterden. Er was echter geen controlegroep die niet met Rekentuin heeft geoefend waarmee deze resultaten vergeleken zouden kunnen worden. Verder werden er geen verschillen tussen de condities gevonden. Dit betekent dat het voor de prestaties niet uitmaakte of leerlingen zelf een succesfrequentie mochten kiezen of dat ze werkten met een vaste succesfrequentie en dat het ook niet uitmaakte hoe hoog de succesfrequentie was.

4.2.5 *Tablet game* optellen en aftrekken tot 20

Van der Ven, Segers, Takashima en Verhoeven (2017) onderzochten de effecten van een *tablet game* waarin leerlingen uit groep 3 spelenderwijs kunnen oefenen met optellen en aftrekken tot 20. In het spel doet de leerling een race tegen een virtuele opponent waarbij de snelheid van deze opponent gelijk wordt gesteld aan voorgaande prestaties van de leerling. Andere kenmerken van het spel zijn dat er directe feedback is, dat er mogelijkheden zijn om een som opnieuw op te lossen indien het

gegeven antwoord fout is, dat er beloningen worden gegeven en dat het visueel aantrekkelijk is. Het spel draait om het oefenen van zowel accuratesse als snelheid.

Om te testen of het spelen van het spel van positieve invloed was op rekenprestaties hebben 103 leerlingen uit groep 3 deelgenomen aan een experiment. De vier verschillende klassen waar deze leerlingen in zaten werden random toegewezen aan de experimentele groep of de controlegroep. Over een periode van vijf weken oefende de experimentele groep vier keer per week met het spel met een tijdsduur van 15 minuten per keer. De controlegroep speelde niet met het spel en volgde alleen de reguliere rekeninstructie. Voorafgaand aan het experiment, direct na het experiment en 13 weken na afloop van het experiment maakten beide groepen een rekentoets. Deze rekentoets bestond uit vier delen van elk één minuut waarin leerlingen zoveel mogelijk bewerkingen goed moesten beantwoorden. Het ging hierbij om getallen lager dan 20 in de volgende categorieën: optellen en aftrekken in Arabische getallen en optellen en aftrekken van niet-symbolische getallen (in dit geval stippen).

Uit de resultaten van de accuratessescores bleek dat veel leerlingen 100 procent van de door hen beantwoorde opgaven correct hadden. Vanwege dit plafondeffect is accuratesse niet verder onderzocht en is alleen gekeken naar de efficiëntiescores (het aantal correct beantwoorde opgaven per minuut). Voor de efficiëntie in de bewerkingen met Arabische getallen werd op de directe nameting geen significant verschil gevonden tussen de experimentele groep en de controlegroep. Dit gold ook voor het optellen van niet-symbolische getallen, maar er was wel een significant verschil in het aftrekken van niet-symbolische getallen. Direct na de interventie waren leerlingen die het spel hadden gespeeld hier efficiënter in dan kinderen die het spel niet hadden gespeeld. Op de test 13 weken na afloop van het experiment werden voor beide groepen en voor alle bewerkingen geen significante verschillen meer gevonden. De gevonden effecten van de interventie verdwenen dus wanneer het spel niet meer werd gespeeld.

4.2.6 Samenvatting rekenspellen en oefenplatforms op de computer

De verschillende onderzoeken laten zien dat zowel speciaal ontworpen, specifieke, computerspellen als algemenere oefenplatforms een positief effect hebben op de rekenprestaties. Hierbij moet wel aangetekend worden dat dit vrijwel altijd een toevoeging was aan het reguliere curriculum, dus het is niet mogelijk om vast te stellen of het beter werkt dan een andere toevoeging

4.3 Trainingen die aansturen op het noteren van uitwerkingsstappen

Roorda, Kleefsman, en Witterholt (2014) onderzochten het effect van een programma waarin leerlingen zes weken, tweemaal per week, getraind werden in het gestructureerd, stapsgewijs uitwerken van samengestelde bewerkingsopdrachten (opgaven die met minstens twee bewerkingen moeten worden opgelost). In totaal name 59 leerlingen uit groep 7 deel, van twee scholen. Binnen elke school volgde één klas het experimentele programma en in de andere klas maakten de leerlingen dezelfde opgaven maar zonder de structureringsaanpak; zij dienden als controlegroep. Het uitvoeren van het experimentele programma had geen eenduidige invloed op de resultaten: op één school verbeterden leerlingen in beide condities hun prestaties en uitwerkingsstappen tussen voor- en nameting (waarbij de experimentele groep qua uitwerkingsstappen wat meer vooruitging dan de controlegroep), op de andere school was in beide condities juist geen verbetering te zien. Verder leken vooral de rekenzwakkere leerlingen meer vooruitgang te boeken (enkel onderzocht voor experimentele groep) qua prestaties en uitwerkingsstappen. Samenvattend lijkt vooral op een school waar men niet gewend is leerlingen te stimuleren hun rekenstappen te noteren, winst geboekt te kunnen worden door te oefenen met samengestelde bewerkingsopdrachten waarbij de leerkracht op een bepaalde manier aandacht geeft aan het noteren van stappen.

Fagginger Auer (2016) onderzocht het effect van een training gericht op het stimuleren van het opschrijven van berekeningsstappen bij het oplossen van opgaven meercijferig delen. In totaal volgden 147 leerlingen uit groep 8 met een beneden gemiddeld rekenniveau een training van drie sessies tussen de voor- en nameting. De helft van de leerlingen volgde een experimentele training waarin ze hun berekeningen zodanig moesten opschrijven dat andere leerlingen konden zien hoe ze het hadden uitgerekend, en in het geval de eerste (zelfstandige) poging de opgave op te lossen niet tot het goede antwoord had geleid ze systematische feedback kregen om de rekenstappen gestructureerd te noteren. Leerlingen in de controletrainingsconditie oefenden dezelfde opgaven, maar zonder instructie over het opschrijven van hun berekeningen en zonder systematische feedback over de gestructureerde notatie van de rekenstappen. De groepen leerlingen gingen in dezelfde mate vooruit in prestaties en aantal geschreven rekenstappen. De extra, individuele, oefening in het oplossen van dit soort opgaven in beide condities lijkt dus een positief effect te hebben op zowel de prestaties, als het noteren van rekenstappen bij leerlingen met een benedengemiddeld rekenniveau. Het expliciet stimuleren de rekenstappen op te schrijven en de systematische feedback op het gestructureerd noteren van de rekenstappen uit de experimentele training hadden daarbij geen aanvullend effect.

Samenvattend geven deze twee studies waarin het effect van het stimuleren van het noteren van rekenstappen is onderzocht aanwijzingen dat vooral op een school waar men niet gewend is leerlingen te stimuleren hun rekenstappen te noteren winst geboekt kan worden. Er werd echter geen overtuigend bewijs gevonden dat het expliciet stimuleren de rekenstappen op te schrijven of systematische feedback op het gestructureerd noteren van de rekenstappen te geven, meer effect had dan enkel het oefenen met de opgaven.

4.4 Professionaliseringsactiviteiten

Een aantal recente studies onderzochten de effecten van professionaliseringstrajecten voor leerkrachten(teams) en/of scholen op de leerresultaten van de leerlingen.

4.4.1 Professionalisering gericht op differentiatie

Van de Weijer-Bergsma en collega's (2016) richtten zich in hun studie op de effecten van een nascholingstraject voor differentiatie. Het traject, genaamd GROW, was gericht op het leren werken met een differentiatiecyclus waarin achtereenvolgens aandacht was voor het vaststellen van de onderwijsbehoeften van leerlingen, het stellen van doelen, het aanbieden van gedifferentieerde instructie en verwerking en tenslotte de evaluatie van het proces en de voortgang. Binnen deze cyclus werd onder andere gewerkt met homogene subgroepen van leerlingen op basis van niveau, waarbij flexibiliteit in de indelingen essentieel was. Er werd onderscheid gemaakt tussen een intensieve subgroep (voor zwakkere rekenaars), een basisgroep (met gemiddelde rekenaars) en een gevorderde subgroep (voor sterkere rekenaars). Het nascholingstraject was voor zowel schoolleiders als leerkrachten bestemd en bestaat uit verschillende soorten bijeenkomsten zoals teambijeenkomsten, coachbijeenkomsten en beleidsbesprekingen. Met de effectstudie is onder andere onderzocht of GROW leidt tot betere differentiatievaardigheden van leerkrachten en tot sterkere rekenprestaties van leerlingen.

Om dit te testen is een longitudinaal onderzoek opgezet waarbij in totaal 31 scholen via loting zijn toegewezen aan drie verschillende cohorten. Het onderzoek, dat een looptijd had van drie jaar, had betrekking op groep 1 tot en met groep 8. Het eerste cohort volgde het nascholingstraject in het eerste jaar. Het tweede cohort diende als controlegroep in het eerste jaar en volgde het nascholingstraject in het tweede jaar. Het derde cohort diende alle jaren als controlegroep. Er zijn verschillende soorten meetinstrumenten gebruikt. De eerste resultaten van de studie, die hieronder beschreven zullen worden, zijn gebaseerd op video-opnames van leerkrachten en Cito-toetsuitslagen

van leerlingen. Om de rekenprestaties van de leerlingen te controleren voor intelligentie en werkgeheugen hebben zij hier nog aanvullende taken voor uitgevoerd.

Met betrekking tot de differentiatievaardigheden van leerkrachten blijkt uit de eerste resultaten van de studie dat leerkrachten in de verschillende cohorten aan het eind van het tweede jaar niet van elkaar verschilden wanneer zij instructie gaven aan de intensieve subgroep, maar wel van elkaar verschilden wanneer zij instructie gaven aan de gevorderde subgroep. Voor deze groep scoorden leerkrachten uit het eerste cohort hoger op differentiatievaardigheden, zoals het bieden van gerichte aandacht en extra uitdaging, dan leerkrachten uit het tweede en derde cohort. Deze effecten waren dus niet zichtbaar in het jaar dat het nascholingstraject plaatsvond, maar in het jaar daarna en alleen voor de groep sterkere rekenaars. De effecten voor de basisgroep werden in deze eerste resultaten nog niet beschreven.

Met betrekking tot de rekenprestaties van leerlingen in groep 3 tot en met groep 8, hadden leerlingen in het eerste cohort aan het einde van het eerste jaar meer vooruitgang dan leerlingen in het tweede en derde cohort. Deze resultaten werden gevonden voor leerlingen van alle niveaus, al waren de effecten het sterkst voor zwakkere rekenaars. Aan het einde van het tweede jaar werden tussen alle cohorten echter geen significante verschillen gevonden. Kortom, er zijn aanwijzingen dat homogene subgroepen op basis van niveau positieve effecten kunnen hebben als flexibiliteit in de indelingen mogelijk is, er klassikale instructie is en er goed wordt ingespeeld op de onderwijsbehoeften van leerlingen. Aangezien het nascholingstraject geen directe effecten op de differentiatievaardigheden van leerkrachten leek te hebben en de resultaten met betrekking tot de prestaties van de leerlingen ook niet een eenduidig beeld lieten zien, is er nog geen sterke ondersteuning voor deze voorlopige conclusies.

4.4.2 Professionalisering gericht op formatief toetsen

Veldhuis en van den Heuvel-Panhuizen (2015) onderzochten in hoeverre het ondersteunen van leerkrachten in het gebruik van formatieve toetstechnieken bij rekenen-wiskunde een positief effect kan hebben op de leerresultaten van leerlingen. Aan het onderzoeksexperiment in groep 5 met voormeting (LVS medio groep 5) en nameting (LVS einde groep 5) deden 30 leerkrachten en hun 616 leerlingen mee. De leerkrachten werden random toegewezen aan één van vier condities: drie experimentele condities die van elkaar verschilden in het aantal professionaliseringsworkshops dat de leerkrachten volgden (één, twee of drie sessies van één uur) en een controleconditie zonder workshops. In de workshops stond het reken-wiskundeprogramma van de komende weken centraal, in het bijzonder de belangrijkste struikelblokken, om zo in samenspraak met de leerkrachten te bepalen welke informatie over de leerlingen nodig was om hun leerproces richting de leerdoelen te bevorderen. Ten slotte kregen de leerkrachten toelichting over een aantal voorbeelden van op deze onderwerpen gerichte formatieve toetstechnieken. Relevant van deze toetstechnieken was dat ze snel in de klas toepasbaar waren om meteen aangrijpingspunten voor verdere instructie te geven, en de benadrukking dat leerkrachten de technieken konden aanpassen aan hun onderwijspraktijk. De resultaten op de nameting lieten zien dat de leerlingen van leerkrachten die drie workshops hadden gevolgd meer vooruit waren gegaan dan leerlingen van leerkrachten uit de andere drie condities (ES = 0.26). De professionaliseringsactiviteiten gericht op het gebruik van formatieve toetspraktijken in de klas hadden dus een positief effect op de leerlingprestaties, maar enkel bij het volgen van drie workshops. Het volgen van een minder intensief programma van één of twee sessies resulteerde niet in meer leerwinst dan de controlegroep. Omdat geen betrouwbare meting van de feitelijke uitvoering van formatieve toetstechnieken beschikbaar was, is het niet eenduidig vast te stellen in hoeverre het daadwerkelijk het gebruik van de toetstechnieken was die de leerwinst veroorzaakten.

4.4.3 Professionalisering gericht op opbrengstgericht werken

Opbrengstgericht werken wordt vaak gezien als een manier van werken waarmee de schoolprestaties van basisschoolleerlingen verhoogd zouden kunnen worden. Hoewel er meerdere definities bestaan, kan opbrengstgericht werken over het algemeen omschreven worden als een manier van werken waarin het analyseren van toetsgegevens, het stellen van doelen en het maken van groepsplannen centraal staan. Verschillende Nederlandse onderzoeksprojecten zijn opgezet om de effecten van opbrengstgericht werken te meten, die hieronder beschreven worden.

Opbrengstgericht leiderschap

Het project 'Opbrengstgericht leiderschap' was gericht op het bevorderen van opbrengstgericht werken bij basisscholen door schoolbesturen hiertoe aan te moedigen en hen hier zelf plannen voor op te laten stellen (Verbeek, Ledoux, & Glaudé, 2012). Bij het project, dat een looptijd had van 3,5 jaar, waren in totaal ruim 500 scholen betrokken. De schoolbesturen hadden de regie en kregen veel vrijheid in hoe zij hun plannen voor de toepassing van opbrengstgericht werken wilden vormgeven en uitvoeren. Vanuit het project kregen zij hierbij financiële ondersteuning (mits de school ook eigen middelen vrijmaakte voor het project) en advies van experts (bij het opstellen van het plan van aanpak en vervolgens eenmaal per jaar ter begeleiding). Evaluaties van de voortgang van de toepassing van opbrengstgericht werken werden zowel door scholen zelf uitgevoerd als door de onderzoekers. Cito-eindtoetsscores werden gebruikt om de taal- en rekenprestaties van de leerlingen vast te stellen.

Uit de resultaten bleek dat er gemiddeld geen toenames waren in de Cito-eindtoetsscores gedurende de loop van het project. Uit vragenlijstonderzoek naar schoolkenmerken bleek dat scholen die positief waren over in hoeverre opbrengstgericht werken bij hen werd toegepast hogere Cito-eindtoetsscores hadden. Ook tevredenheid bij de scholen over de rol van het bestuur in het project hing positief samen met de leerlingprestaties.

Streef-project

Het project *Streef* van de Rijksuniversiteit Groningen in samenwerking met twee onderwijsadvies-organisaties bestond uit meerdere deelonderzoeken gericht op interventies voor opbrengstgericht werken (Doolaard, 2013b). Het onderzoek was quasi-experimenteel van opzet. Scholen die mee wilden doen konden binnen het Streef-project zelf aangeven welk deelonderzoek hun interesse had. De meeste scholen volgden een deelproject met ofwel de bovenbouw ofwel de onderbouw, waarbij de andere bouw van de school fungeerde als controlegroep voor één van de andere deelprojecten. De data werden verzameld middels vragenlijsten voor leerkrachten en directies voor en na afloop van de projecten en via toetsen uit het Cito LVS. De deelprojecten die relevant zijn in het kader van het huidige onderzoek zullen hieronder worden beschreven.

Het eerste project, *Werken met referentieniveaus* was gericht op het in kaart brengen van de kennis van leerkrachten op het gebied van referentieniveaus en het vervolgens verhogen van deze kennis middels een interventie (Doolaard, 2013b). Aan dit deelproject namen 20 scholen deel. De interventie bestond uit twee regionale trainingsbijeenkomsten met meerdere scholen tegelijkertijd en twee bijeenkomsten op iedere school apart. Met betrekking tot rekenen leerden leerkrachten onder andere welke opgaven tot een bepaald referentieniveau behoorden en aan welke referentieniveaus hun eigen leerlingen voldeden. Uit de resultaten bleek dat het volgen van de interventie leidde tot een beter begrip van de referentieniveaus en meer realistische verwachtingen van de leerlingen. Er waren echter geen effecten zichtbaar op de schoolprestaties van leerlingen.

In een tweede deelproject werd een andere interventie onderzocht: *Gebruik maken van opbrengsten* (Doolaard, 2013a). Voor dit project hadden zich 40 scholen aangemeld. In deze interventie leerden leerkrachten Cito-toetsuitslagen te analyseren en deze te gebruiken om groepsplannen te maken en te

differentiëren. Ook deze interventie bestond uit een combinatie van een aantal regionale trainingsbijeenkomsten en voor iedere school aparte begeleidingssessies (met daarbij nog een aantal huiswerkopdrachten). In totaal stond er 40 uur voor de interventie. De resultaten lieten zien dat leerkrachten die de interventie hadden gevolgd uitgebreidere toetsanalyses maakten en daarbij een breder reflectiekader hanteerden (door bijvoorbeeld eerdere resultaten of landelijke scores te betrekken in de analyses) dan leerkrachten in de controlegroep. Ook maakten leerkrachten in de interventiegroep vaker groepsplannen, maar de uiteindelijke manier van differentiëren verschilde niet tussen de twee groepen. Wat betreft de leerlingprestaties voor taal en rekenen bleken voor zowel het jaar dat de interventie plaatsvond als voor het jaar daarna er geen significante verschillen te zijn tussen de interventiegroep en de controlegroep.

Focus-project

Het Focus-project van de Universiteit Twente was gericht op opbrengstgericht werken in het basisonderwijs op basis van prestatiefeedback. Prestatiefeedback wordt geleverd aan de school door het leerling monitor systeem. Schoolleiders, intern begeleiders en leerkrachten kunnen zo reflecteren op de prestaties van de klassen en school in het algemeen en deze vergelijken met het nationale gemiddelde. In de Focus training leren schoolleiders, intern begeleiders en leerkrachten de prestatiefeedback grondig en correct te analyseren, te interpreteren en toe te passen. Tussen 2010 en 2015 zijn er vier projecten geweest; drie projecten waarin teamtrainingen gegeven werden en één waarbij individuele trainingen aan leerkrachten gegeven werden. Visscher (2015) geeft een overzicht van alle vier de Focus projecten. Alleen bij Focus-project I werd er geen verschil gevonden in de rekenprestaties tussen de controlegroep en de experimentele groep. De overige drie Focus-projecten hadden wel een positief effect op de rekenprestaties van leerlingen.

De interventies in Focus II en III duurden twee jaar en richtten zich op klas en school. Aan de interventie namen schoolleiders, intern begeleiders en leerkrachten deel. Beide onderzoeken hadden een *multiple single subject design*. Dit betekent dat meerdere scholen individueel gevolgd werden en de resultaten van de school met voorgaande prestaties binnen dezelfde school vergeleken werden. De rekenresultaten van de leerlingen uit groep 3-8 werden gemeten met Cito-LVS toetsen. In totaal zijn rekenresultaten van 7500 leerlingen in Focus II en van 4750 leerlingen in Focus III gemeten. Bij beide projecten werd gemiddeld één maand leerwinst gevonden over een periode van twee jaar. Het effect van de interventie was het grootst op de rekenprestaties van leerlingen met een lage SES. Verder had de interventie het meeste effect op scholen met veel gewichtenleerlingen.

In Focus IV kregen de deelnemende leerkrachten, die allemaal les gaven aan groep 6, individuele trainingen. Deze studie duurde één jaar en leerkrachten werden random aan de experimentele of controlegroep toegewezen. In totaal zijn de rekenresultaten van 673 leerlingen gemeten. Er werd alleen een positief effect gevonden voor zwakker presterende leerlingen. Zij behaalden twee maanden leerwinst door de interventie. Er was geen effect voor de overige leerlingen.

Bovenstaande Focus-projecten hadden gemiddeld een positief effect, maar er waren ook scholen die deelnamen waarbij de rekenresultaten niet stegen. Keuning, Van Geel en Visscher (2017) deden onderzoek naar de oorzaak van deze verschillen binnen de Focus-projecten. Hiervoor hebben zij de invloed van school- en leerkrachtfactoren geanalyseerd en hiervan een vergelijking gemaakt tussen de tien scholen met de meeste groei in rekenresultaten en de tien scholen met de minste groei. Uit kwantitatieve en kwalitatieve data bleek dat de kwaliteit van de leerkracht, de houding van de leerkrachten tegenover de interventie en de datacultuur invloed hadden op het effect van de interventie. Op scholen waarbij de interventie een groot effect had op de rekenprestaties hadden de leerkrachten vooraf aan de interventie al een positievere houding tegen over de interventie dan op scholen waarbij er geen effect was. De leerkrachten waren leergierig en veranderde hun lespraktijk naar aanleiding van de feedback die gekregen was in de interventie. Daarnaast werd de kwaliteit van lesgeven als hoger beoordeeld tijdens observaties, verricht door de trainers vanuit het focusproject,

en was de (data)cultuur binnen de school al meer resultaatgericht. Tijdens de interventieperiode bleef bij de scholen met een groot effect de houding positief. Bij de scholen met een groot effect waarbij de houding voor de interventie negatief was, werd de houding tijdens de interventie steeds positiever. Bij de scholen met geen effect veranderde de houding niet of werd deze zelfs negatiever. Voor een succesvolle implementatie van opbrengstgericht werken, zijn dus leergierige leerkrachten nodig die open zijn voor veranderingen binnen hun lespraktijk en het nut in zien van zowel opbrengstgericht werken, als een training daarin. In zowel de scholen met als de scholen zonder effect werd de (data)cultuur steeds resultaatgerichter. Aan het einde van de interventie waren de scholen met effect resultaatgerichter dan de scholen zonder effect. Ook de kwaliteit van lesgeven werd tijdens de interventie beter op zowel scholen met als scholen zonder effect, maar aan het einde van de interventie was de geobserveerde kwaliteit hoger op scholen met effect.

4.4.4 Professionalisering gericht op instructievaardigheid

Logtenberg en de Lange (2014) onderzochten het effect op leeropbrengsten van het inzetten van het professionaliseringsinstrument *Lesson Study* bij rekenlessen op de basisschool. Het doel van *Lesson Study* is het verbeteren van de instructievaardigheid van de leerkracht aan de hand van herhalende cycli waarin teammomenten bij de voorbereiding en nabespreking afgewisseld worden door uitvoeringsmomenten met observatie. In de voorbereidingsfase komen leraren, een vakinhoudelijke expert en een vertegenwoordiger van de schoolleiding bijeen om een doel vast te stellen en grondig een reken-wiskundeconcept te bestuderen. Aan de hand van de opgedane kennis en de gedeelde kennis tussen de leraren en de vakinhoudelijk expert wordt de les ontworpen. Deze les wordt getest in de klas waarna het team aan de hand van observaties en vragen van de leerlingen de les evalueert en doorontwikkelt. Op deze manier doorloopt het team meerdere cycli. In onder andere Japan, de Verenigde Staten, Engeland en Scandinavië wordt *Lesson Study* al gebruikt en levert het positieve resultaten op.

Om te testen of het gebruik van *Lesson Study* ook effectief is in Nederland hebben Logtenberg en de Lange (2014) een pilotstudy opgezet in het (speciaal) basisonderwijs. Vijf basisscholen werden in de *Lesson Study* cyclus begeleid door vier procesbegeleiders. De deelnemende klassen, met in totaal 233 leerlingen, waren verspreid over de onder- en bovenbouw. Aan de hand van een vragenlijst en resultaten van Cito-leerlingvolgsysteemtoetsen is er een voor- en nameting gedaan. Het beschreven effect van *Lesson Study* op de deelnemende scholen is vergeleken met een normgroep, maar is niet statistisch getoetst.

De door de leerkracht ingevulde vragenlijsten wijzen op weinig tot geen verschil tussen voor- en nameting in de interactievaardigheden van de leerkracht, maar wel een kleine verandering in de instructievaardigheden. De leerkrachten gaven aan dat zij de methode meer los durfde te laten. Daarnaast hadden de leerkrachten het gevoel meer verdieping en inzicht te hebben in de lesstof en leerlijn door de voorbereidingsfase. De leerlingen die deelnamen aan de *Lesson Study* lessen hebben een gemiddelde vaardigheidsgroei van 7,3 procent. De leerwinst is hierbij 22 procent hoger dan de normgroep. Dit verschil is echter niet statistisch getoetst. Logtenberg en de Lange (2014) hebben daarnaast geen rekening gehouden met andere factoren die het resultaat mogelijk beïnvloed hebben en geven zelf ook aan dat het resultaat niet perse uitsluitend door het gebruik van *Lesson Study* komt, maar mogelijk de focus op rekenen binnen de school. Het gebruik van *Lesson Study* wordt door de deelnemende leerkrachten als positief ervaren, met name door de samenwerking. Door de vraag naar actieve bijdrage en samenwerking van de leerkrachten, experts en schoolleiding sluit *Lesson Study* ook aan bij het werken in een leergemeenschap.

Samenvattend zijn de resultaten van de pilot naar het inzetten van het professionaliseringsinstrument *Lesson Study* veelbelovend, maar laten de verzamelde gegevens en analyses geen harde uitspraken over de effectiviteit toe.

4.4.5 Samenvatting professionaliseringsactiviteiten

De studies naar professionaliseringsactiviteiten voor leerkrachten gericht op differentiatievaardigheden en instructievaardigheden laten geen overtuigend beeld van de effecten op rekenprestaties zien. Professionaliseringsactiviteiten op het gebied van formatief toetsen kunnen effectief zijn voor de rekenprestaties van leerlingen, waarbij de intensiviteit van de training en het soort begeleiding een rol spelen.

4.5 Schoolvisie: Daltononderwijs

Er is één studie gevonden waarin het effect van de visie van de school, en specifiek het Daltononderwijs, is onderzocht. Daltononderwijs behoort, net als bijvoorbeeld jenaplan- en montessorionderwijs, tot het traditioneel vernieuwingsonderwijs in Nederland. Daltononderwijs wordt gekenmerkt door vrijheid, zelfstandigheid en samenwerken (Van der Ploeg, 2010). Sins en Van der Zee (2015) onderzochten de verschillen tussen daltonscholen en traditionele scholen op allerlei niet-cognitieve en cognitieve aspecten, waaronder ook rekenprestaties. De data voor deze studie waren afkomstig van het COOL$^{5\text{-}18}$ project, waaraan in totaal 550 basisscholen deelnamen. Scholen die aangaven daltononderwijs te geven en ook voldeden aan de kwaliteitseisen van de Nederlandse Dalton Vereniging werden geselecteerd voor het huidige onderzoek. Deze scholen, 31 in totaal, zijn vergeleken met 26 traditionele scholen. Scholen werden als traditioneel aangemerkt indien zij aangaven geen vernieuwende schoolconcepten of elementen daaruit toe te passen en indien zij ook zelf aangaven traditioneel ingericht te zijn. De rekenprestaties van leerlingen zijn vastgesteld aan de hand van scores op toetsen uit het Cito-leerlingvolgsysteem. Daarnaast is informatie verzameld over de volgende achtergrondkenmerken van de leerlingen: geboortejaar, geslacht, intelligentie, afkomst ouders, opleidingsniveau ouders en mate van ouderbetrokkenheid (aangegeven door de leerkracht).

Uit de resultaten bleek ten eerste dat leerlingen op daltonscholen en traditionele scholen van elkaar verschilden op meerdere achtergrondkenmerken. Op de daltonscholen was er meer ouderbetrokkenheid, waren ouders vaker hoog opgeleid en vaker van autochtone afkomst en hadden de leerlingen een hogere intelligentie. Multilevelanalyses waarbij gecontroleerd werd voor deze achtergrondkenmerken zijn vervolgens voor groep 2, 5 en 8 apart uitgevoerd. Bij leerlingen in groep 2 bleken er tussen daltonscholen en traditionele scholen geen verschillen in rekenprestaties te zijn. Dit was ook het geval bij leerlingen in groep 5, al werd hier wel gevonden dat specifiek voor leerlingen met hoogopgeleide ouders de rekenprestaties op daltonscholen hoger waren dan op traditionele scholen. Bij leerlingen met laagopgeleide ouders werden geen verschillen gevonden. Tenslotte bleek schooltype niet van belang voor de rekenprestaties van leerlingen in groep 8. Kortom, op basis van deze studie lijkt het erop dat schooltype (daltononderwijs of traditioneel onderwijs) niet of nauwelijks samenhangt met de rekenprestaties van basisschoolleerlingen.

Samenvattend bleken er gecorrigeerd voor verschillen in achtergrondkenmerken van de leerlingpopulaties nauwelijks verschillen te zijn in de rekenprestaties van leerlingen op Daltonscholen en traditionele scholen.

4.6 Leerlingovertuigingen

4.6.1 Motivatie, zelfvertrouwen en welbevinden

Peetsma, Roeleveld en Stoel (2003) onderzochten onder andere de relatie tussen psychosociaal functioneren van basisschoolleerlingen en hun rekenprestaties. Psychosociaal functioneren bestond in dit onderzoek uit motivatie (of de leerling bijvoorbeeld inzet en volharding laat zien), zelfvertrouwen (in hoeverre de leerling snel angstig of van streek is) en welbevinden (of de leerling zich goed voelt op school en een goede relatie met de leerkracht heeft). Het onderzoek was gebaseerd op gegevens uit het PRIMA-cohortonderzoek. Er zijn 2693 leerlingen gedurende de basisschool gevolgd, waarbij de

metingen plaatsvonden in groepen 2, 4, 6 en 8. De rekenprestaties van de leerlingen zijn gemeten aan de hand van diverse rekentoetsen die binnen het PRIMA-onderzoek zijn afgenomen (welke deels voor PRIMA zijn geconstrueerd en deels afkomstig zijn uit het Cito-leerlingvolgsysteem). De psychosociale kenmerken van de leerlingen zijn vastgesteld op basis van oordelen van leerkrachten via vragenlijsten. Geslacht, opleiding van ouders en herkomst van de leerlingen zijn als achtergrondkenmerken meegenomen in de analyses.

Uit de resultaten bleek ten eerste dat de niet-beïnvloedbare achtergrondkenmerken van invloed waren op rekenprestaties. Er werden hogere rekenprestaties behaald door autochtone leerlingen, leerlingen met hoger opgeleide ouders en (behalve in groep 2) jongens. Met betrekking tot de psychosociale kenmerken bleken zelfvertrouwen en vooral motivatie positieve effecten te hebben op rekenprestaties. Zelfvertrouwen en motivatie waren ook van invloed op welbevinden, maar welbevinden zelf had geen invloed op rekenprestaties. Motivatie had bij leerlingen in groepen 6 en 8 sterkere effecten op rekenprestaties dan bij leerlingen in groepen 2 en 4.

Samenvattend hadden motivatie en zelfvertrouwen van leerlingen een positieve samenhang met rekenprestaties.

4.6.2 Rekenangst

Erturan en Jansen (2015) onderzochten door middel van twee studies de relatie tussen de emotionele beleving van leerlingen op het gebied van rekenen enerzijds en rekenprestaties anderzijds, en onderzochten daarbij ook mogelijke verschillen tussen jongens en meisjes. Aan de eerste studie namen 134 leerlingen van in totaal drie scholen deel. Het waren leerlingen uit groepen 5, 6 en 8 en leerlingen uit de eerste en tweede klas van het voortgezet onderwijs (VWO). In deze studie zijn drie variabelen onderzocht: rekenprestaties (Tempo Test Rekenen, die de basisbewerkingen meet), rekenangst en toetsangst (beide met vragenlijsten). Uit de resultaten bleek dat jongens en meisjes niet van elkaar verschilden in rekenprestaties en rekenangst, maar dat meisjes wel significant hoger scoorden op toetsangst. Verder bleek dat er voor jongens geen relatie was tussen rekenangst en rekenprestaties, maar dat dit voor meisjes wel het geval was. Bij meisjes ging een hogere rekenangst samen met lagere rekenprestaties, ook wanneer er gecontroleerd werd voor toetsangst.

In de tweede studie beschreven door Erturan en Jansen (2015) participeerden 208 leerlingen uit groepen 5 tot en met 8 van drie verschillende basisscholen. Ook in deze studie werden rekenprestaties en rekenangst gemeten, maar in tegenstelling tot in de eerste studie werd niet toetsangst maar door de leerlingen ervaren rekencompetentie gemeten. Er bleken tussen jongens en meisjes geen verschillen te zijn in rekenprestaties, rekenangst en ervaren rekencompetentie. Zowel voor jongens als meisjes hingen hogere rekenprestaties samen met hogere ervaren rekencompetentie. Wanneer er gecontroleerd werd voor ervaren rekencompetentie was er voor zowel jongens als meisjes geen relatie tussen rekenangst en rekenprestaties.

Samengevat bleek uit de twee studies dat jongens en meisjes niet van elkaar verschilden in rekenangst en ervaren rekencompetentie, maar dat meisjes wel hogere toetsangst hadden. In de eerste studie leek er voor meisjes een negatieve relatie te zijn tussen rekenangst en rekenprestaties, maar in de tweede studie, waarbij gecontroleerd werd voor ervaren rekencompetentie, bleek er geen relatie te zijn tussen rekenangst en rekenprestaties.

4.6.3 Samenvatting leerlingovertuigingen

Samenvattend lijken de leerlingovertuigingen zelfvertrouwen en motivatie samen te hangen met rekenprestaties. De negatieve samenhang tussen rekenangst en rekenprestaties loopt mogelijk via de ervaren rekencompetentie.

5 Resultaten nadere analyses recente peilingen

De meest grootschalige informatie die beschikbaar is over de 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; en de *Trends in International Mathematics and Science Studies* (TIMSS), die een internationale vergelijking mogelijk maken. Om de huidige review naar beïnvloedbare factoren in het onderwijsleerproces die samenhangen met rekenprestaties in perspectief te kunnen plaatsen wordt in het Hoofdstuk 5.1 beschreven hoe Nederlandse basisschoolleerlingen tot nu toe gepresteerd hebben bij deze peilingen van het rekenniveau. Voor de volledigheid worden ook niet-beïnvloedbare factoren die samenhingen met rekenprestaties in deze peilingen besproken.

Bij beide peilingen is niet alleen informatie verzameld over prestaties en niet-beïnvloedbare factoren, maar ook over factoren die wel beïnvloedbaar zijn. In de originele rapportages worden deze factoren echter niet in verband gebracht met de rekenprestaties. Omdat dit direct relevant is voor de huidige review zijn nadere analyses verricht op de data van PPON-2011 en TIMSS-2015, gericht op de relatie tussen het onderwijsaanbod (leerkrachtvragenlijsten) en de leerlingprestaties. Die worden in het tweede en derde deel van dit hoofdstuk beschreven.

5.1 Stand van zaken rekenniveau Nederlandse basisschoolleerlingen.

5.1.1 Periodieke Peiling van het Onderwijsniveau (PPON) in groep 8

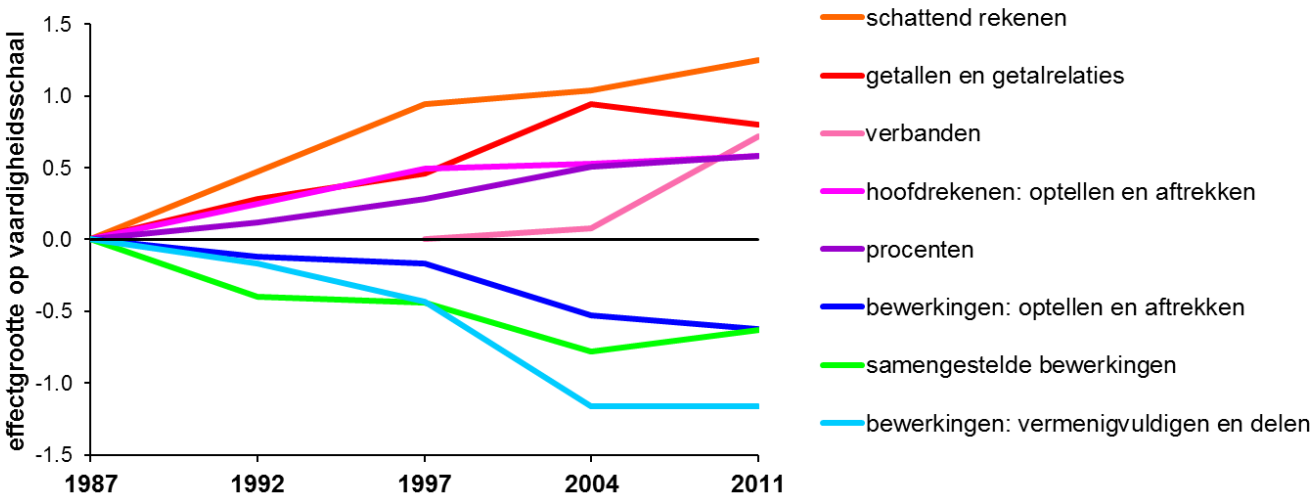
In de jaren 1987, 1992, 1997, 2004 en 2011 zijn door Cito peilingen van het reken-wiskundeniveau aan het einde van Nederlandse basisscholen verricht (Bokhove, Van der Schoot, & Eggen, 1996; Janssen, Van der Schoot, & Hemker, 2005; Janssen & Verhelst, 1999; Scheltens et al., 2013; Wijnstra, 1988). In 2011 namen er 2403 groep 8-leerlingen van 108 basisscholen deel aan de peiling. In dit peilingsonderzoek worden met opgaven over 22 onderwerpen de kerndoelen voor rekenen-wiskunde aan het einde van de basisschool gedekt.

Voor elk van de 22 onderwerpen is in kaart gebracht welke veranderingen in rekenprestaties zich hebben voorgedaan tussen 1987 en 2011 (Scheltens et al., 2013). Bij sommige onderwerpen zijn de veranderingen over deze periode gering, terwijl bij andere onderwerpen prestaties aanzienlijk zijn gedaald of gestegen (Figuur 5.1). De veranderingen tussen de twee meeste recente peilingen in 2004 en 2011 waren voor alle onderwerpen beperkt, behalve voor Verbanden.

- In het domein *Getallen en bewerkingen* is de vaardigheid van leerlingen wat betreft basisoperaties en hoofdrekenend vermenigvuldigen en delen weinig veranderd tussen 1987 en 2011. De vaardigheid in bewerkingen met grotere getallen en kommagetallen (zowel voor optellen, aftrekken, vermenigvuldigen, delen en samengestelde bewerkingen) kende sterke dalingen. Positieve ontwikkelingen zijn gevonden voor getallen en getalrelaties, hoofdrekenend optellen en aftrekken en schattend rekenen.
- In het domein *Verhoudingen, breuken en procenten* ontwikkelde de vaardigheid van leerlingen op het gebied van verhouding en breuken zich weinig over de hele periode van 1987 tot 2011. Bij procenten vond een continu positieve ontwikkeling plaats, die afvlakte tussen 2004 en 2011.
- In het domein *Meten en meetkunde* deden zich weinig veranderingen voor in de vaardigheid van leerlingen op vrijwel alle onderwerpen: lengte, oppervlakte, inhoud, gewicht, toepassingen, meetkunde, tijd en geld.

- In de het domein *Verbanden* nam de vaardigheid van leerlingen sterk toe tussen 2004 en 2011.

Figuur 5.1: Verandering van vaardigheidsniveaus voor onderwerpen met grootste veranderingen (effectgrootte minstens 0.5) tussen PPON-1987 en PPON-2011.



Naast het vergelijken van prestaties van leerlingen met eerdere prestaties, kunnen ze ook worden afgezet tegen de referentieniveaus (Expertgroep doorlopende leerlijnen, 2008). De referentieniveaus 1F en 1S aan het eind van de basisschool worden geïllustreerd aan de hand van opgaven uit PPON 2004 die door respectievelijk 75 en 50 procent van de groep-8-leerlingen beheerst werden (Scheltens et al., 2013). Gezien de beperkte veranderingen in prestaties tussen PPON-2004 en PPON-2011, is het percentage leerlingen dat de niveaus beheerst bij beide peilingen veelal vergelijkbaar en is er in de relatief korte tijd dat referentieniveaus gehanteerd worden dus geen sprake van een voor- of achteruitgang ten opzichte van deze gestelde niveaus.

Ten slotte is onderzocht in hoeverre het prestatieniveau van leerlingen samenhangt met diverse factoren: geslacht, leertijd, formatiegewicht en schoolstratum in kaart gebracht (Scheltens et al., 2013). Hierbij werd telkens gecorrigeerd voor de andere variabelen die in de analyse zijn meegenomen. Jongens presteerden op bijna alle onderwerpen hoger scoren dan meisjes (zeer kleine tot matige verschillen), uitgezonderd bewerkingen en meetkunde. Leerlingen met een langere leertijd, die dus ergens in hun schoolloopbaan vertraging hebben opgelopen, presteerden slechter dan leerlingen die niet vertraagd zijn (matig tot grote verschillen). Het formatiegewicht van leerlingen zegt iets over hun sociaaleconomische achtergrond. Hoe hoger het formatiegewicht (hoe lager de sociaaleconomische klasse), hoe lager de rekenprestaties. Ook zijn op basis van de formatiegewichten van leerlingen scholen ingedeeld in drie strata, die globaal de samenstelling van de schoolbevolking kenmerken. De leerlingpopulatie op stratum 1-scholen heeft gemiddeld een lager formatiegewicht dan op stratum 2 en 3-scholen. Als toevoeging aan het effect van formatiegewicht van de individuele leerlingen waren de gemiddelde prestaties lager op stratum-2-scholen (geen tot kleine verschillen) en stratum-3-scholen (kleine tot matige verschillen) dan op stratum-1-scholen.

5.1.2 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 2015 werd daarbij een vergelijking tussen 49 landen gemaakt en namen er wereldwijd 312000 leerlingen uit 'grade 4' (groep 6) deel, waarvan 4634 leerlingen van 128 Nederlandse basisscholen.

In 2015 scoorden Nederlandse leerlingen ver boven het algemene TIMSS-schaalgemiddelde voor rekenen. De score was wel significant lager dan die in 2011 en is de laagste die Nederland in de afgelopen twintig jaar heeft behaald (maar de verschillen zijn klein). In 16 landen scoorden leerlingen in 2015 gemiddeld significant beter dan Nederlandse leerlingen. Van de Noord-Europese landen hebben alleen leerlingen uit Duitsland en Zweden in 2015 significant slechter gepresteerd. 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.

In TIMSS worden drie inhoudelijke rekendomeinen getoetst: *Getallen*, *Geometrische vormen en meten* en *Gegevensweergave*. Binnen elk inhoudsdomein zijn er drie typen vragen (cognitieve dimensies): weet-, toepassings- en redeneervragen. Nederlandse leerlingen presteerden relatief het best bij de inhoudsdomeinen *Getallen* en *Gegevensweergave*, maar hun prestaties bij deze domeinen zijn in 2015 wel gedaald ten opzichte van 2011. Het domein *Geometrische vormen en meten* past volgens curriculumexperts en leerkrachten het minst bij het Nederlandse curriculum en Nederlandse leerlingen presteerden daar ook relatief het slechtst (maar ongeveer hetzelfde bij de afgelopen drie peilingen). Net als bijvoorbeeld leerlingen uit Scandinavische landen, Duitsland en Finland, presteerden Nederlandse leerlingen beter op redeneervragen dan op weet- en toepassingsvragen. Nederlandse leerlingen presteerden het minst goed in weetvragen en presteerden op deze vragen in 2015 minder goed dan in 2011.

Ook wordt in TIMSS in kaart gebracht hoeveel leerlingen een bepaalde rekenstandaard halen. Deze standaarden variëren van een basisniveau tot een geavanceerd niveau. In Nederland haalde een relatief groot gedeelte van de leerlingen het basisniveau (99 procent in 2015), maar een relatief kleine gedeelte het geavanceerde niveau (4 procent in 2015). In 1995 lag het percentage leerlingen dat het geavanceerde niveau haalde hoger (12 procent), maar dat halveerde in 2003 en bleef daarna laag.

Het rekenprestatieniveau van Nederlandse leerlingen hangt samen met diverse factoren. In TIMSS-2015 werden onder andere de effecten van geslacht, thuistaal en schoolkenmerken in kaart gebracht (Meelissen & Punter, 2016). Jongens presteerden significant beter bij rekenen dan meisjes, maar het verschil is klein en is in de afgelopen twintig jaar iets in grootte afgenomen. De jongens presteerden vooral beter bij het inhoudsdomein *Getallen*. De prestaties van leerlingen die thuis soms of nooit Nederlands spreken liggen lager dan die van leerlingen die thuis (bijna) altijd Nederlands spreken, in het bijzonder bij opgaven uit het inhoudsdomein *Gegevensweergave* en bij redeneeropgaven. Er was geen effect op rekenprestaties van hoeveel procent van de leerlingen op een school Nederlands als eerste taal spreken (meer of minder dan 90 procent), hoeveel procent van de leerlingen op een school uit een gezin met economische achterstand komen (meer of minder dan 10 procent) of de omgeving van een school (verstedelijkt of niet-verstedelijkt).

5.1.3 Conclusies peilingsonderzoeken

In het algemeen geven de grootschalige peilingsonderzoeken het beeld dat Nederlandse leerlingen in internationaal perspectief relatief hoog presteren, maar dat deze prestaties dalen. In TIMSS wordt dit algemener waargenomen en in PPON komt een gedifferentieerder beeld naar voren: op bepaalde onderdelen gaan de prestaties achteruit, maar op andere onderdelen juist vooruit (zie ook Van Weerden, Janssen, & Scheltens, 2016). Verder hangen de rekenprestaties samen met een aantal niet-beïnvloedbare factoren: geslacht van de leerling (in het voordeel van jongens over de hele linie in TIMSS en op de meeste onderdelen in PPON), thuistaal van de leerling in TIMSS en de sociaaleconomische status van de leerling en de school in PPON.

5.2 Nadere analyses PPON-2011

Cito heeft voor de huidige review nadere analyses uitgevoerd op een vragenlijst die bij PPON-2011 is voorgelegd aan de leerkrachten van wie leerlingen deelnamen aan de peiling. Het doel van deze

vragenlijst was om het rekenonderwijsaanbod in kaart te brengen. De antwoorden van leerkrachten op deze vragenlijst werden gerelateerd aan de rekenprestaties van leerlingen in verschillende rekendomeinen.

5.2.1 Methode

De leerkrachtvragenlijst bestond uit 50 vragen, waarvan sommige bestonden uit meerdere onderdelen. Deze vragen zijn teruggebracht tot 48 variabelen, die betrekking hadden op verschillende aspecten van het rekenonderwijs (zie Bijlage 2): algemene factoren (bijv. groepsgrootte); rekenspecifieke schoolfactoren (bijv. de gebruikte rekenmethode); rekenspecifieke leerkrachtfactoren (bijv. hoeveelheid klassikale instructie); onderwijstijd (bijv. hoeveelheid tijd voor rekenen per week); rekenspecifieke-domeinspecifieke factoren (bijv. in hoeveel strategieën leerlingen instructie krijgen). Deze variabelen werden in verkennende analyses door middel van multilevelmodellen gerelateerd aan de rekenprestaties van 2532 leerlingen van 108 scholen. Naast prestaties op rekenen in zijn totaliteit werden prestaties op verschillende deelaspecten onderzocht, waarvoor tien vaardigheidsschalen werden samengesteld op basis van 22 rekenonderwerpen (zie Tabel 5.1). Eerst werd voor elke leerkrachtvariabele los getoetst of deze significant samenhing met elk van de tien vaardigheidsschaalscores. De variabelen die samenhingen met rekenprestaties, werden samen in een multilevelmodel per schaal opgenomen. In deze modellen werd gecorrigeerd voor samenstelling van de leerlingpopulatie van scholen qua sociaal-economische status (SES). Voor elke schaal werd op de variabelen in het model *backwards selection* toegepast, wat inhoudt dat steeds de variabele met de grootste *p*-waarde uit het model wordt verwijderd, tot alleen significante variabelen in het model overblijven. Rekenmethode en sociaal-economische status van de school bleven ongeacht hun *p*-waarde in het model.

5.2.2 Resultaten

Tabel 5.1 beschrijft de tien schalen en geeft de proportie van de variantie van schaalscores op schoolniveau (de intraklassecorrelatie: ICC), en daarmee de hoeveelheid variantie die potentieel te verklaren is door leerkrachtvariabelen. Bij alle schalen behalve S05 (het domein Verbanden) is die hoeveelheid variantie dusdanig dat het design-effect groter is dan 2, wat betekent dat voor deze schalen een multilevelanalyse zinvol is. Ook wordt in Tabel 5.1 weergegeven hoeveel van de 48 leerkrachtvariabelen een significant verband hadden met de schaalscore. Bij de schaal *Basisoperaties* was een ander patroon te zien dan bij de andere schalen en was er als enige een significant overall verschil tussen rekenmethodes. Bij de schaal *Bewerkingen* waren er significante verschillen in scores afhankelijk van het aantal strategieën dat leerlingen gebruiken bij bewerkingen met vermenigvuldigen en delen (bij beide bewerkingen één, bij één bewerking meerdere of bij beide bewerkingen meerdere, zoals gerapporteerd door de leerkracht).

Op basis van deze correlaties werden multilevelmodellen met de relevante variabelen per schaal samengesteld. In de resterende kolommen van Tabel 5.1 is te zien welke variabelen in welk model werden opgenomen, welke daarvan kwamen te vervallen door *backwards selection*, en welke variabelen in het uiteindelijke model positief of negatief samenhingen met prestaties (bij rekenmethode en SES zijn er ook niet-significante coëfficiënten, omdat op deze variabelen geen *backwards selection* werd toegepast).

Te zien is dat schaalscores meestal niet verschilden naar rekenmethode (Pluspunt, Rekenrijk, Alles Telt, vergeleken met Wereld in Getallen). Alleen bij de schaal *Basisoperaties* werd met zowel Pluspunt, Rekenrijk als Alles Telt significant lager gepresteerd dan met Wereld in Getallen (maar dit geldt niet voor de overige methodes). Bij *Verhoudingen* en *Schattend rekenen* presteerden leerlingen met Pluspunt ook significant lager en bij *Hoofdrekenen* leerlingen met Rekenrijk, terwijl leerlingen met 'overige methoden' significant hoger presteerden bij *Meten en Meetkunde* en *Hoofdrekenen*. Bij alle

schalen behalve *Schattend rekenen* werd slechter gepresteerd door leerlingen op scholen met een leerlingpopulatie met een lagere SES.

Hoewel groepsgrootte met diverse schalen negatief correleerde, vervalt deze variabele uit de modellen voor alle schalen met *backwards selection*. De leeftijd van de leerkracht vervalt uit vier modellen, maar hangt bij vier andere schalen, *Totaalschaal*, *Getallen*, *Hoofdrekenen* en *Schattend rekenen,* positief samen met leerlingprestaties. De tevredenheid van leerkrachten met de steun die leerlingen thuis ontvangen vervalt uit vijf modellen en hangt bij drie andere schalen, *Totaalschaal*, *Verhoudingen* en *Schattend rekenen,* positief samen met leerlingprestaties. Hoe goed de methode bij de leerkracht past vervalt uit zes modellen en hangt bij drie schalen, *Totaalschaal*, *Verhoudingen* en *Meten en Meetkunde*, positief samen met leerlingprestaties. De frequentie waarmee fouten die veel gemaakt worden leerlingen klassikaal worden besproken vervalt uit vier modellen en hangt bij vier schalen, *Totaalschaal*, *Verhoudingen*, *Meten en Meetkunde* en *Getallen en Getalsrelaties,* negatief samen met leerlingprestaties. Het aantal strategieën dat leerlingen gebruiken bij bewerkingen met vermenigvuldigen en delen correleerde negatief met scores op de schaal *Bewerkingen*, maar vervalt daaruit. Het percentage leerlingen dat externe steun ontvangt zit alleen in het model voor de schaal *Basisoperaties* en hangt daar negatief samen met prestaties. De lestijd die wordt besteed aan individuele instructie correleerde negatief met scores op de schaal *Basisoperaties*, maar vervalt daaruit.

5.2.3 Conclusies nadere analyses PPON-2011

Slechts een beperkte hoeveelheid van de variantie in rekenprestaties bij PPON-2011 is op schoolniveau (tussen de 4 en 9 procent). Daarmee is er slechts een klein deel van rekenprestaties dat potentieel te verklaren is door leerkrachtvariabelen uit de aanbodvragenlijst van PPON-2011. De correlaties tussen leerkrachtenvariabelen en diverse schaalscores verschillen slechts in enkele gevallen significant van nul (bij 3 tot 5 van de 48 variabelen per schaal). Rekenmethode lijkt vooral samen te hangen met rekenprestaties bij Basisoperaties, en niet zozeer bij andere onderwerpen. Hierbij verschilde Wereld in Getallen positief van Pluspunt, Rekenrijk en Alles Telt. In het verslag van PPON-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. De sociaaleconomische status van de leerlingpopulatie van scholen had bij alle schalen (op één na) een significant effect en is dus een belangrijke variabele om voor te corrigeren.

Verschillende significante effecten kwamen naar voren uit de leerkrachtenvariabelen die op grond van hun correlaties met schaalscores samen werden opgenomen in multilevelmodellen. 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 gebruiken beter bij zich vinden 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 zaken daadwerkelijk een negatief effect op prestaties hebben, maar slechts dat ze ermee samengaan (zo is meer externe hulp voor leerlingen die slechter presteren niet meer dan logisch).

Al met al hangen slechts enkele van de tientallen variabelen uit de leerkrachtvragenlijst op een beperkt aantal gebieden 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.

Tabel 5.1: Vaardigheidsschalen PPON met samenhang met leerkrachtvariabelen

Schaal	ICC (proportie variantie op schoolniveau)	Aantal leerkracht-variabelen met significante correlatie met schaalscore	Leerkrachtvariabelen in multilevelmodel per schaal												
			Pluspunt (vs. WiG)	Rekenrijk (vs. WiG)	Alles Telt (vs. WiG)	Overige methoden (vs. WiG)	SES leerlingpopulatie	Groepsgrootte	Leeftijd leerkracht	Tevreden met steun thuis	Hoe goed past rekenmethode bij leerkracht	Klassikaal bespreken veelgemaakte fouten	Aantal strategieën bewerkingen x :	Percentage leerlingen met externe steun	Lestijd individuele instructie
Totaalschaal (excl. basis-operaties)	0,09	5	0	0	0	0	-	v	+	+	+	-	n	n	n
Getallen (domein)	0,09	5	0	0	0	0	-	v	+	v	v	v	n	n	n
Verhoudingen (domein)	0,08	5	-	0	0	0	-	v	v	+	+	-	n	n	n
Meten en Meetkunde (domein)	0,07	5	0	0	0	+	-	v	v	v	+	-	n	n	n
Verbanden (domein)	0,04	te weinig variantie op schoolniveau voor multilevelanalyse													
Getallen en getalsrelaties	0,05	3	0	0	0	0	-	v	v	v	v	-	n	n	n
Hoofdrekenen	0,06	4	0	-	0	+	-	v	+	v	v	v	n	n	n
Bewerkingen	0,09	4	0	0	0	0	-	v	v	v	v	v	v	n	n
Schattend rekenen	0,05	5	-	0	0	0	0	v	+	+	v	v	n	n	n
Basisoperaties	0,09	4	-	-	-	0	-	v	n	n	v	n	n	-	v

+ = significant positieve regressiecoëfficiënt
- = significant negatieve regressiecoëfficiënt
0 = regressiecoëfficiënt niet significant anders dan 0
v = leerkrachtvariabele verwijderd uit model in *backwards selection*
n = leerkrachtvariabele niet opgenomen in model

5.3 Nadere analyses TIMSS-2015

Onderzoekers van Universiteit Twente hebben in opdracht van het ministerie van OCW nadere analyses uitgevoerd op de data van TIMSS 2015 (Rebber et al., 2017). Hierbij werden onder andere de rekenprestaties van leerlingen gerelateerd aan twee zaken: de percepties van leerkrachten over hun didactische vaardigheden en het leerklimaat op hun school, en het uitgevoerde curriculum (percentage onderwezen onderwerpen).

5.3.1 Leerkrachtpercepties

Leerkrachten vulden voor TIMSS een leerkrachtvragenlijst in over verschillende procesfactoren: perceptie van werkdruk, zelfvertrouwen in didactische vaardigheden, waargenomen mate van toerusting per inhoudsdomein, frequentie van bijscholing, en perceptie van de prestatiegerichtheid van het schoolklimaat. Met behulp van een *partial credit model* werden contextvariabelen samengesteld op basis van antwoordpatronen op sets items die een overkoepelend construct representeren.

De analyses met deze procesfactoren werden gedaan met behulp van multilevelmodellen met drie niveaus (leerling, leerkracht en school). Naast de procesfactoren werd hierbij gekeken naar zogenaamde inputfactoren: leerling- en schoolkenmerken die nauwelijks of moeilijk manipuleerbaar zijn (sekse, taal en aantal boeken thuis (als indicator van SES) op leerlingniveau en percentage achterstandsleerlingen en urbanisatiegraad op schoolniveau). Met een combinatie van *forwards* en *backwards selection*, om te voorkomen dat ten onrechte variabelen werden uitgesloten vanwege een niet-significant effect, kwamen de onderzoekers tot een uiteindelijk model met uitsluitend significante effecten van factoren.

De potentieel te verklaren variantie op schoolniveau was in Nederland 9 procent en de variantie op klasniveau 3,5 procent. In de andere vier landen die in het nadere onderzoek werden meegenomen (Vlaanderen, Engeland, Finland in Zweden) was de gezamenlijke variantie op school- en klasniveau hoger (tot wel 32,8 procent). Verschillende inputfactoren op leerlingniveau hingen in de Nederlandse data samen met prestaties. Jongens presteerden beter dan meisjes, kinderen die thuis de toetstaal spraken presteerden beter dan kinderen die dat niet deden, en kinderen met meer boeken thuis presteerden beter dan kinderen met minder boeken. Bij Vlaanderen, Engeland en Zweden hing het percentage achterstandsleerlingen op school ook samen prestaties, maar bij Nederland niet.

Van alle procesfactoren bleek in Nederland alleen de leerkrachtperceptie van de prestatiegerichtheid van het schoolklimaat samen te hangen met rekenprestaties, in positieve zin. Dit zou kunnen betekenen dat leerlingen beter presteren bij hogere verwachtingen, maar ook dat beter presterende leerlingen zorgen voor hogere verwachtingen. Nederland was in TIMSS-2015 een land waar de prestatiegerichtheid relatief laag is, en bij de andere vier landen in het nadere onderzoek was het verband met prestaties niet significant. Andere procesfactoren zijn ook voor andere landen niet significant, behalve een negatief effect van bijscholing voor rekenen in Vlaanderen. De werkdruk ligt in Nederland hoger dan internationaal gemiddeld, maar in de analyses werd geen relatie gevonden tussen deze factor en prestaties.

5.3.2 Uitgevoerd curriculum

Aan de leerkrachten werd voor het in kaart brengen van het uitgevoerde curriculum ook gevraagd in hoeverre verschillende leerstofgebieden behandeld waren (vorig schooljaar, dit schooljaar en (nog) niet behandeld). Op basis hiervan werd het percentage behandelde leerstofgebieden (dit of vorig schooljaar) per inhoudsdomein bepaald. De effecten van het uitgevoerde curriculum werden net als bij leerkrachtpercepties in kaart gebracht met multilevelmodellen met inputfactoren.

Bij het inhoudsdomein *Getallen* was 8,2 procent van de variantie op schoolniveau en 3,5 procent op klasniveau. Bij *Geometrische vormen en meten* was dat respectievelijk 10 en 3 procent en bij *Gegevensweergave* 8,5 en 2,2 procent. Al zijn de verschillen dus beperkt, de grootste variantie op schoolniveau was bij het domein *Geometrische vormen en meten*. Dat is het domein dat het minst goed bij het Nederlandse beoogde en uitgevoerde curriculum past (Meelissen & Punter, 2016) en volgens de onderzoekers verschillen scholen daarom waarschijnlijk ook wat meer in de mate waarin ze hier aandacht aan besteden. De mate waarin de leerstof van elk inhoudsdomein is behandeld bleek niet samen te hangen met rekenprestaties in het algemeen of specifiek met de prestaties per domein. De inputvariabelen sekse, thuistaal en aantal boeken thuis hadden ook een effect op de prestaties per domein.

5.3.3 Conclusies nadere analyses TIMSS-2015

Ook in TIMSS-2015 is een relatief klein percentage van de variantie in rekenprestaties van leerlingen op schoolniveau (rond de 10 procent). Slechts één procesvariabele (waarover leerkrachten zelf in een vragenlijst rapporteerden) hangt samen met de rekenprestaties van hun leerlingen: de waargenomen prestatiegerichtheid van het schoolklimaat. De aansluiting van de inhoud van de toets bij het curriculum hing niet samen met de rekenprestaties.

5.4 Conclusies nadere analyses PPON-2015 en TIMSS-2011

Zowel in PPON-2011 (groep 8) als TIMSS-2015 (groep 6) ligt van de verschillen in rekenprestaties tussen leerlingen slechts een klein deel (4-10%) op het niveau van de school. Dit is dus ook de maximale hoeveelheid te verklaren variantie door (beïnvloedbare en niet-beïnvloedbare) kenmerken van het onderwijsleerproces. Rekenmethode (PPON) en behandelde leerstofgebieden (TIMSS) hingen niet of nauwelijks samen met de rekenprestaties. De factoren die wel positief met rekenprestaties samenhingen geven tezamen niet een heel coherent of overtuigend beeld. Al met al hebben de in deze peilingen gehanteerde leerkrachtvragenlijsten slechts beperkte informatie opgeleverd over de relatie tussen factoren van het onderwijsleerproces en rekenprestaties van leerlingen.

6 Resultaten veldraadpleging

Om de op basis van wetenschappelijke literatuur verkregen inzichten te toetsen aan de praktijk van het reken-wiskundeonderwijs is een veldraadpleging gehouden, in samenwerking met Stichting Leerplanontwikkeling (SLO). Aan de veldraadpleging namen rekenexperts deel met verschillende achtergronden zoals PABO-docenten, onderwijsadviseurs, rekencoördinatoren, leerkrachten uit groep zeven en acht, uitgevers, onderzoekers, het ministerie van OCW en de PO-Raad. Het voornaamste doel van de veldraadpleging was om inzicht te krijgen in wat volgens praktijkexperts de belangrijkste factoren zijn binnen het onderwijsleerproces die van invloed zijn op rekenprestaties, zodat we onze bevindingen konden valideren en aanvullen.

In totaal zijn 82 opmerkingen gemaakt. Deze zijn gecategoriseerd en geordend van meest genoemd naar minst genoemd. De resultaten van deze analyse wordt weergegeven in Tabel 6.1.

Tabel 6.1. Door de praktijkexperts als meest belangrijk genoemde beïnvloedbare factoren.

thema	%	toelichting
leerkracht	22,0	De kwaliteit van de leerkracht, de opleiding, kennis, attitude en overtuigingen, relaties tussen de leerkracht en de leerlingen en professionalisering.
toetsing	12,2	Toetsing met als doel zicht op voortgang en evaluatie. Formatief toetsen en feedback worden ook genoemd.
rekenbeleid	11,0	Rekengesprekken en rekencoördinator (met name ook de invulling van de functie en de achtergrond van de coördinator doen ertoe).
instructie	8,5	Instructievormen, met name ook diversiteit in leer- en werkvormen.
klimaat	8,5	Klimaat en structuur, veiligheid en sfeer.
inhoud	7,3	Opportunity to learn en leerstofaanbod.
rekentijd	6,1	De tijd die aan rekenen wordt besteed.
motivatie	4,9	Intrinsieke motivatie, zowel voor de leerling, als voor de leerkracht, ouders en omgeving.
leerling-overtuiging	3,7	Geloof in eigen kunnen.
ouders	3,7	Ouderbetrokkenheid en houding van ouders ten opzichte van school en rekenen.
niet-beïnvloedbare factoren	3,7	Geslacht en leeftijd van de leerkracht en leerling, groepsgrootte.
hulpmiddel	2,4	Niet gespecificeerd.
leren leren	2,4	Leerling als eigenaar van zijn eigen leerproces.
afstemmen	2,4	Adequate instructie geven, aanpassen aan de behoeften van de leerlingen.
differentiatie	1,2	Gedifferentieerd onderwijs, ook voor de plusklassen.

Volgens de praktijkexperts hebben dus de (kenmerken van de) leerkracht, zoals de kennis, houding, opleiding en professionalisering, een zeer belangrijke invloed op de rekenprestaties. Van de leerkrachtfactoren wordt de kennis van de leerkracht het meest frequent genoemd, zoals te zien is in Tabel 6.2. Onder kennis van de leerkracht wordt bijvoorbeeld de pedagogische en didactische (vak)kennis en kennis over de leerlijnen verstaan.

Tabel 6.2. Door praktijkexperts genoemde relevante leerkrachtfactoren.

categorie	%
algemeen	27,8%
kennis	27,8%
attitude en overtuiging	16,7%
opleiding	11,1%
professionalisering	11,1%
relatie	5,6%

Naast de rol van de leerkracht wordt ook de rol van toetsing, en dan met name de manier waarop toetsing wordt ingezet, zoals voor voortgang, evaluatie, feedback of het veranderen van het leerproces (formatief toetsen) veelvuldig genoemd. Als derde wordt rekenbeleid en de rekencoördinator genoemd. Kenmerken van de daadwerkelijke rekenles, zoals het type instructie, de inhoud van de les en de tijd die aan de les wordt besteed, staan niet in de top drie van factoren die er volgens de praktijkexperts het meest toe doen.

De praktijkexperts mochten ook aangeven welke factoren volgens hen nog ontbraken. Het overzicht van de opmerkingen die de experts maakten staat thematisch geordend in Tabel 6.3. Sommige factoren vallen wel onder de categorieën in het factorenschema (zoals klassikale instructie, geïndividualiseerde instructie/differentiatie, beleidsplan), andere factoren vallen buiten de huidige kaders van "beïnvloedbare factoren van het onderwijsleerproces" (zoals de rol van de ouders). De andere factoren hebben we waar mogelijk meegenomen in de aanbevelingen voor nader onderzoek binnen en buiten de peilingen (zoals methodegebruik).

Tabel 6.3. Volgens praktijkexperts nog ontbrekende factoren.

thema	voorbeelden
inhoudoverstijgend	▪ samenhang leerstofaanbod ▪ methode-analyse ▪ materiaalgebruik (wat naast of in plaats van methodes) ▪ rekenen leren, geen ideologische fratsen en franjes
instructie	▪ aandacht voor informeel handelen in de instructie, ▪ bevorderen rekentaal (bijvoorbeeld keer = vermenigvuldigen) ▪ klassikale instructievormen, dus juist niet individualiseren/differentiëren. ▪ zelfontdekkend leren ▪ verwerkingsvorm (digitaal) ▪ problem-solving/denkinstructie, omdat probleem-oplossend vermogen een 21e eeuwse vaardigheid is. ▪ leerkracht [moet] zorgen voor succeservaringen.

- vervolg Tabel 6.3. Volgens praktijkexperts nog ontbrekende factoren. -

tijd	▪ tijdverhouding groepsleren en individueel leren, omdat kinderen ook van elkaar leren en altijd alleen leren kan leiden tot verkeerde concepten. ▪ tijdstip van de les ▪ goede tijdsplanning vooral bij gecombineerde groepen ▪ planning vs. realisatie
hulpmiddelen/ leeromgeving	▪ rijke leeromgeving,is een voorwaarde om leerling eigen initiatief te laten tonen en betrokkenheid te creëren ▪ inrichting van de klas (bijvoorbeeld tijdelijke hulpmiddelen zichtbaar ook echt tijdelijk zichtbaar) ▪ hulpmiddelen. Denk ook aan bijvoorbeeld tafelkaarten, rekenmachine als hulpmiddelen voor zwakke rekenaars.
beleid	▪ duidelijk beleidsplan omdat weten waar je mee bezig bent, gezamenlijke koers en doorgaande lijn ▪ invloed keuzes van het besturen
leerproblemen/ zwakke rekenaars	▪ vroeg signaleren en interventie om opeenstapeling van leerproblemen te voorkomen ▪ is pre-teaching aanwezig voor zwakke leerlingen? ▪ handelingensplannen ▪ remedial Teaching ▪ kwaliteit van de behandeling
sterke rekenaars	▪ verrijkingsprogramma voor de betere rekenaars. ▪ inzet van een plusklas (in het algemeen, waar onder andere executieve vaardigheden ondersteund worden en leren leren etc).
professionalisering	▪ teamscholing doorgaande lijn, open discussies en toewijding. ▪ peer review (als specificatie binnen professionalisering in algemenere zin)
ervaring leerkracht	▪ effect onbevoegde leerkracht/stagiaires ▪ factor duobaan, omdat ik denk dat de continuïteit van het leerproces van belang is ▪ de aanstelling van de leerkracht (voltijd/deeltijd) ▪ onderwijservaring (aantal jaren)
welbevinden (toetsmoment)	▪ welbevinden van de leerling op het moment van toetsen: ruzie op het plein, thuissituatie. Meerdere toetsmomenten/ breder tijdsvak. Omdat het prestaties heel erg beïnvloedt. Welbevinden van de leerling zorgt voor betere prestaties
ouders	▪ omdat die veel invloed hebben op wat, hoe en waarom kinderen iets wel of niet leren.

Samenvattend noemden de praktijkexperts als belangrijkste factor kenmerken van de leerkracht, met name de kennis van de leerkracht, gevolgd door (formatieve) toetsing en rekenbeleid. Voor formatieve toetsing en (pedagogisch) vakinhoudelijke kennis sluiten de resultaten aan bij die uit de internationale en Nederlandse onderzoeken. Aangezien voor rekenbeleid van de school geen onderzoeksresultaten beschikbaar waren, verdient het aanbeveling deze blinde vlek in de komende peiling mee te nemen en hier nader gericht onderzoek naar te doen.

7 Conclusies en aanbevelingen

7.1 Samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en rekenprestaties.

De analyse van internationale onderzoeksliteratuur (33 meta-analyses en 19 reviews), onderzoek bij Nederlandse basisschoolleerlingen (26 (deel)studies) en nadere analyses van de gegevens van de meest recente peilingen TIMSS-2015 (groep 6) en PPON-2011 (groep 8) leverde een grote verzameling van beïnvloedbare factoren in het onderwijsleerproces die samenhangen met rekenprestaties. Tegelijkertijd bleek uit de nadere analyses van TIMSS-2015 en PPON-2011 dat slechts een beperkte hoeveelheid (4-10 procent) van de verschillen in rekenprestaties tussen Nederlandse leerlingen op het niveau van de school lag. Dat betekent dat in de huidige situatie hoogstens 10 procent van de verschillen in rekenprestaties verklaard kan worden door (beïnvloedbare en niet-beïnvloedbare) factoren uit het onderwijsleerproces.

7.1.1 Niveau van de les.

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 constructivistische instructie. Bovendien zijn er zeer 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 KNAW (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-technologische hulpmiddelen* zoals 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 Rekentuin. 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 den Heuvel-Panhuizen, in druk).

- 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 geen robuuste onderzoeksresultaten zijn met betrekking tot leerstofaanbod of rekenmethode: noch in de internationale literatuur, noch in de nadere analyses van PPON-2011 en TIMSS-2015. Hoewel het belang van deze factoren voor de hand ligt (zie ook Hiebert & Grouws, 2007; Van Zanten & van den Heuvel-Panhuizen, 2014) lijkt het moeilijk deze gericht te onderzoeken. Mogelijk komt dit doordat de begrippen heel 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.

7.1.2 Niveau van de leerkracht

Wat betreft de leerkracht hangen zowel meer algemene l*eerkrachtvaardigheden* (effectief pedagogisch handelen, klassenmanagementvaardigheden) als *rekenspecifieke kennis* (vakinhoudelijke kennis en pedagogisch vakinhoudelijke kennis) 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 verschilden 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, maar de resultaten zijn (nog) niet sterk overtuigend. Intensieve en duurzame ondersteuning lijken belangrijk. Over het effect van *leerkrachtovertuigingen* vonden we weinig onderzoeksliteratuur, behalve een negatieve invloed van het induceren van stereotype 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.

7.1.3 Niveau van 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 PPON gevonden. Omdat klassenkenmerken waarschijnlijk niet los gezien kunnen worden van de leerkracht denken wij niet dat het nodig is hier specifieke hypothesen voor te formuleren.

7.1.4 Niveau van 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 en schoolklimaat.

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. Een ander onderzoek met Nederlandse leerlingen liet zien dat *schoolvisie*, en dan specifiek Daltononderwijs versus traditionele scholen, geen relatie had met rekenprestaties. 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. Deze kenmerken van de school verdienen dus nader onderzoek.

7.1.5 Niveau van 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 samenhangt met rekenprestaties. Rekenangst lijkt zowel oorzaak als gevolg van 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.

7.2 Aanbevelingen voor toekomstig peilingsonderzoek

Voordat we ingaan op aanbevelingen over welke factoren van het onderwijsleerproces meegenomen zouden moeten worden in toekomstig peilingsonderzoek, is het raadzaam nogmaals stil te staan bij de bevinding dat in de meest recente peilingen slechts 4-10 procent van de variantie in rekenprestaties in Nederland op het niveau van de school lag. Dit is daarmee ook het maximum wat in het komende peilingsonderzoek mogelijk verklaard kan worden door beïnvloedbare factoren in het onderwijsleerproces. Als we daarbij ook nog in ogenschouw nemen dat het waarschijnlijk is dat een aantal niet-beïnvloedbare factoren, zoals de sociaaleconomische status van de leerlingpopulatie, ook een deel van de schoolvariantie zullen verklaren, geeft dat aan dat we bescheiden moeten zijn in de verwachtingen wat het 'meenemen van factoren in het onderwijsleerproces' zal opleveren.

Daarnaast is uit de nadere analyses van zowel PPON-2011 als TIMSS-2015 duidelijk geworden dat de daar gehanteerde vragenlijsten slechts beperkte resultaten gaven wat betreft de samenhang tussen factoren uit het onderwijsleerproces en de rekenprestaties van leerlingen. Vragenlijstonderzoek naar instructiepraktijken heeft een aantal beperkingen (Porter, 2002). Ten eerste zijn ze beperkt tot wat de onderzoekers a priori hebben besloten te bevragen. Ten tweede heeft zelfrapportage een aantal inherente beperkingen, zoals dat leerkrachten rapporteren wat zij denken dat gepast is, dat wat leerkrachten denken dat ze doen afwijkt van wat ze daadwerkelijk doen, of dat de gebruikte terminologie niet duidelijk is voor leerkrachten. Ten derde zijn ze beperkt in de mate van complexiteit van instructiepraktijken die gemeten kan worden. Leslogboeken of lesobservaties kunnen een meer valide alternatief zijn (Heidelberg, Van den Bergh, Kuhlemeier, & Hoeks, 1993). Deze alternatieven zijn uiteraard moeilijker praktisch te realiseren in nationale peilingsonderzoeken. Porter (2002) beargumenteert dat vragenlijsten naar de *inhoud* van instructie valide en daarmee de beste optie kunnen zijn, maar niet als het gaat om de *kwaliteit* van de instructie.

Naast deze beperkingen van (leerkracht)vragenlijsten in het algemeen is onze indruk verder dat de kwaliteit van de vraagstellingen in PPON-2011 voor verbetering vatbaar is. Sommige vragen zijn erg

breed gesteld en daarmee weinig onderscheidend, andere vragen hadden weinig variatie in antwoorden of leken sterk gevoelig voor sociaalwenselijke anwtoorden. Ten slotte zijn in PPON-2011 enkel gegevens verzameld bij leerkrachten, en niet op het niveau van de school of de leerling.

Wij raden daarom aan om (a) meer metingen te doen dichter bij het onderwijsleerproces, bijvoorbeeld middels lesobservaties en/of leslogboeken, met name voor de kenmerken van de les, (b) wanneer vragenlijsten worden gebruikt gerichtere, meer onderscheidende vragen te stellen en (c) gegevens bij alle actoren in het onderwijsleerproces te verzamelen, dat wil zeggen naast van leerkrachten ook van leerlingen en schooldirectie.

7.2.1 Aanbevelingen over te verzamelen gegevens van kenmerken van de les

Gegevens over de kenmerken van de les kunnen het best verzameld worden met lesobservaties en/of leslogboeken, met name als het vragen over het 'hoe' betreft (zoals mate van leerkrachtondersteuning). Vragen over het 'wat' kunnen ook met gerichte vragen in een vragenlijst voor de leerkracht beantwoord worden.

lesfactor	**hoe?**
rekenmethode	
▪ welke rekenmethode wordt gebruik in groep 8?	▪ vragenlijst
▪ in hoeverre wordt de rekenmethode gevolgd (*gecombineerd met rekenmethode-analyse als indicatie van leerstofaanbod*)	▪ observaties/logboek
▪ welk evt. aanvullend schriftelijk materiaal?	▪ vragenlijst
onderwijstijd, inclusief *time on task*	observaties/logboek
gebruik technologische hulpmiddelen: wat	
▪ educatieve applicaties zoals Snappet, Rekentuin en methode-software	▪ vragenlijst
▪ (exploratieve) digitale leeromgevingen	▪ vragenlijst
▪ gebruik van mobiele apparaten in de klas	▪ vragenlijst
gebruik technologische hulpmiddelen: hoe	
▪ hoe vaak	▪ observaties/logboek
▪ met welke functie(s)	▪ observaties/logboek
▪ met welke mate van leerkrachtondersteuning	▪ observaties/logboek
gebruik niet-technologische hulpmiddelen zoals driedimensionale materialen	observaties/logboek
in hoeverre wordt het directe instructiemodel gevolgd?	observaties/logboek
mate van ondersteuning bij meer zelfstandige leeractiviteiten als ontdekkend leren of computerprogramma's	observaties/logboek
gebruik van verschillende vormen van formatieve toetsing (feedback aan leerkrachten en/of leerlingen)	
▪ formatieve toetspraktijken in de dagelijkse klassenpraktijk (met voorbeelden)	▪ observaties/logboek
▪ het gebruik van gegevens van leerlingvolgsystemen en methode-toetsen bij het vormgeven van de les	▪ vragenlijst
differentiatie in niveaugroepen	
▪ homogeen/heterogeen	▪ vragenlijst
▪ flexibel of (relatief) vast	▪ vragenlijst
▪ op basis waarvan de indeling?	▪ vragenlijst
▪ hoe geeft de leerkracht de verschillende groepen instructie?	▪ observaties/logboek
hoeveelheid huiswerk per week (inschatting tijdsbesteding gemiddelde leerling)	▪ vragenlijst

7.2.2 Aanbevelingen over te verzamelen gegevens van kenmerken van de leerkracht

Gegevens over de kenmerken van de les kunnen het beste verzameld worden door gerichte vragen in een vragenlijst voor de leerkracht, liefst aangevuld met kennistoetsen die de leerkrachten zelf maken. Kanttekening is dat op deze manier enkel kenmerken van de leerkracht van groep 8 meegenomen worden, terwijl de leerlingen al 6-8 jaar rekenonderwijs hebben gekregen.

leerkrachtfactor	hoe?
vakinhoudelijke kennis rekenen-wiskunde (niveau groep 8, alle vier de domeinen)	rekentoets (vergelijkbaar met leerlingentoets)
pedagogisch vakinhoudelijke kennis rekenen-wiskunde: ▪ kennis van leerstrategieën ▪ kennis van misconcepties bij leerlingen ▪ kennis van leerlijnen rekenen-wiskunde	toets/gevalideerd meetinstrument
professionaliseringsactiviteiten op het gebied van rekenwiskundeonderwijs ▪ indien ja: welk thema/onderwerp?	vragenlijst
actieve deelname aan professionele leergemeenschappen / teamoverleg rondom rekenonderwijs ▪ indien ja: ervaren nut?	vragenlijst
prestatieverwachtingen leerlingen (fixed / growth)	vragenlijst/gevalideerd meetinstrument
in hoeverre past rekenmethode bij leerkracht?	vragenlijst

7.2.3 Aanbevelingen over te verzamelen gegevens van kenmerken van de school

Gegevens over de kenmerken van de school kunnen het beste verzameld worden door gerichte vragen in een vragenlijst aan de directeur, eventueel aangevuld met een analyse van de schoolgids.

schoolfactor	hoe?
rekenvisie en rekenbeleid, o.a. ▪ aanwezigheid rekencoördinator ▪ beschikbare tijd (formatie) taken rekencoördinator ▪ taken rekencoördinator ▪ opleiding tot rekencoördinator ▪ door de leerkrachten ervaren ondersteuning door rekencoördinator ▪ hoe wordt keuze voor rekenmethode gemaakt?	 ▪ vragenlijst ▪ vragenlijst ▪ vragenlijst ▪ vragenlijst ▪ vragenlijst (leerkracht) ▪ vragenlijst (evt. open vraag)
belang van rekenen, o.a. ▪ hoeveelheid buitencurriculaire activiteiten op het gebied van rekenen ▪ aandacht in de schoolgids? ▪ stimuleren van nascholing binnen het team op het gebied van rekenen	 ▪ vragenlijst ▪ analyse schoolgids ▪ vragenlijst
mate van opbrengstgericht werken bij het rekenen	vragenlijst (evt. vragen naar concrete voorbeelden)
prestatiegerichtheid van de school	vragenlijst

7.2.4 Aanbevelingen over te verzamelen gegevens van kenmerken van de leerling.

Gegevens over de kenmerken van de leerling kunnen het beste verzameld worden door gerichte vragen in een vragenlijst aan de leerling. Daarbij is het aan te bevelen op zoek te gaan naar gevalideerde meetinstrumenten, en/of de vragenlijst van TIMSS te gebruiken. Dat laatste heeft als voordeel dat de resultaten van TIMSS (groep 6) en Peil (groep 8) met elkaar te vergelijken zijn.

leerlingfactor	hoe?
beliefs over het vak rekenen	vragenlijst
motivatie voor rekenen	vragenlijst
mate van rekenangst	vragenlijst
attributie van rekenprestaties	vragenlijst
attitude ten opzichte van rekenen	vragenlijst
plezier in rekenen	vragenlijst
zelfvertrouwen in rekenen	vragenlijst
self efficacy met betrekking tot rekenen	vragenlijst

7.3 Beperkingen

Een eerste categorie beperkingen van deze review 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, dit niet betekent dat dit niet onderzocht is. Mogelijk bestaan er meta-analyses of reviews die wij gemist hebben, maar waarschijnlijk bestaan er sowieso individuele studies die waardevolle informatie geven over de bijdrage van deze factor. De breedte van het onderzoeksdomein 'factoren in het onderwijsleerproces' en de beschikbare tijd voor het review lieten het niet toe dit soort individuele studies op te sporen, maar 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.

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 correlationeel. Dat betekent dat wel kan worden vastgesteld of een bepaalde factor met rekenprestaties samenhangt, maar niet of de factor de uitkomsten beïnvloedt (een causaal verband). Als bijvoorbeeld de hoeveelheid tijd die wekelijks wordt besteed aan rekenlessen positief samenhangt met de rekenprestaties, kan het zijn dat door 'aan de knop rekenonderwijstijd te draaien' de prestaties omhoog gaan (een causaal verband), maar het kan ook zijn dat de scholen met een leerlingpopulatie met 'slimmere' leerlingen ook meer tijd aan rekenen besteden (bijvoorbeeld omdat minder tijd voor taal nodig is). In het laatste scenario lijkt het enkel zo dat meer tijd aan rekenonderwijs betere prestaties tot gevolg heeft, terwijl de samenhang enkel te verklaren is doordat onderwijstijd en leerlingkenmerken samenhangen. Een omgekeerde bevinding,

een negatieve samenhang tussen rekenonderwijstijd en rekenprestaties, kan ook op twee manieren verklaard worden: door het (niet zo aannemelijke) causale verband 'meer tijd resulteert in lagere opbrengsten', of door een onderliggende variabele zoals opnieuw aanleg van de leerlingen: met meer 'minder slimme' leerlingen is meer rekenonderwijstijd nodig, terwijl de resultaten nog steeds achterblijven bij die van scholen met meer 'slimmere' leerlingen. Men moet dus heel voorzichtig zijn met de conclusies en praktijkaanbevelingen die op basis van dit soort correlationele onderzoeken gedaan worden.

Ten slotte beperkt de focus op empirische onderzoeken de reikwijdte van het review. Zo blijft de literatuur over leertheorieën in het algemeen en rekenwiskundedidactiek in het bijzonder hoofdzakelijk buiten beschouwing.

7.4 Aanbevelingen voor overig nader onderzoek.

Zoals al eerder genoemd beperkt de correlationele aard van peilingsonderzoeken de interpretatie van verbanden tussen factoren in het onderwijsleerproces en rekenprestaties. 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. Om robuustere conclusies te trekken zijn meer gecontroleerde onderzoeken nodig. Experimenten of quasi-experimenten waarbij een bepaalde (les)factor systematisch gevarieerd wordt is methodologisch gezien een sterke manier, maar is in de onderwijspraktijk niet altijd praktisch haalbaar en 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 formatief toetsen en differentiëren.

Wij bevelen nader onderzoek op in ieder geval de volgende thema's aan:

- De rol 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.

Referenties

Alfieri, L., Brooks, P., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? A meta-analysis. *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*(2), 224–237.

Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational Assessment*, *27*(3), 197–205. https://doi.org/10.1177/0734282908330580

Bakker, M. (2014). *Using mini-games for learning multiplication and division: A longitudinal effect study*. Proefschrift Universiteit Utrecht.

Bakker, M., Van den Heuvel-Panhuizen, M., & Robitzsch, A. (2015). Effects of playing mathematics computer games on primary school students' multiplicative reasoning ability. *Contemporary Educational Psychology*, *40*, 55–71.

Bakker, M., Van den Heuvel-Panhuizen, M., Van Borkulo, S., & Robitzsch, A. (2013). Effecten van online mini-games op multiplicatieve vaardigheden van leerlingen in groep 4. *Pedagogische Studiën*, *90*(3), 21–36.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Bebell, D., & Kay, R. (2010). One to One Computing: A summary of the quantitative results from the Berkshire wireless learning initiative. *The Journal of Technology, Learning and Assessment*, *9*(2), 60. https://doi.org/10.1556/ComEc.10.2009.1.10

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. https://doi.org/10.1111/j.1741-5446.2006.00241.x

Blank, R. K., & de las Alas, N. (2009). *Effects of teacher professional development on gains in student achievement.* Washington, DC: Council of Chief State School Officers. Retrieved from www.ccsso.org

Blok, H., Ledoux, G., & Roeleveld, J. (2015). Opbrengstgericht werken in het primair onderwijs: Een effectieve weg naar onderwijsverbetering? *Pedagogische Studiën*, *92*(3), 167–178.

Bokhove, J., Van der Schoot, F., & Eggen, T. (1996). *Balans van het rekenonderwijs aan het einde van de basisschool 2. Uitkomsten van de tweede peiling rekenen/wiskunde einde basisonderwijs*. Arnhem.

Bosman, A. M. T. (2015). Zo leer je alle kinderen rekenen. *Orthopedagogiek: Onderzoek En Praktijk*, *54*(10), 413–424.

Burns, M. K., Codding, R. S., Boice, C. H., & Lukito, G. (2010). Meta-analysis of acquisition and fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. *School Psychology Review*, *39*(1), 69–83.

Capar, G., & Tarim, K. (2015). Efficacy of the cooperative learning method on mathematics achievement and attitude: A meta-analysis research. *Educational Sciences: Theory & Practice*, *15*(2), 553–559. https://doi.org/10.12738/estp.2015.2.2098

Carbonneau, K. J., & Marley, S. C. (2013). Activity-based learning strategies and academic achievement. In J. A. C. Hattie & E. M. Anderman (Eds.), *The international handbook of student*

achievement (pp. 282–284). New York: Routledge.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, *105*(2), 380–400. https://doi.org/10.1037/a0031084

Carey, E., Hill, F., Devine, A., & Szücs, D. (2016). The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance. *Frontiers in Psychology*, *6*(JAN), 1–6. https://doi.org/10.3389/fpsyg.2015.01987

Chang, H., & Beilock, S. L. (2016). The math anxiety-math performance link and its relation to individual and environmental factors: A review of current behavioral and psychophysiological research. *Current Opinion in Behavioral Sciences*, *10*, 33–38. https://doi.org/10.1016/j.cobeha.2016.04.011

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, *9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. https://doi.org/10.1016/0364-0213(89)90002-5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987-2003. *Review of Educational Research*, *76*(1), 1–62.

Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systmatic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, *34*, 12–25. https://doi.org/10.1016/j.tate.2013.03.001

Deunk, M., Annemieke, S. D., Roel, S.-J., & Bosker, J. (2015). *Differentiation within and across classrooms: A systematic review of studies into the cognitive effects of differentiation practices*. Groningen: GION. Retrieved from https://www.nro.nl/wp-content/uploads/2015/03/Roel-Bosker-Effectief-omgaan-met-verschillen-in-het-onderwijs-review.pdf

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.

Durkin, K. (2011). The self-explanation effect when learning mathematics: a meta-analysis. *Conference on Building and Educatioan Science: Investigating Mechanisms*, 1–5. Retrieved from http://eric.ed.gov/?id=ED518041

Erturan, S., & Jansen, B. (2015). An investigation of boys' and girls' emotional experience of math, their math performance, and the relation between these variables. *European Journal of Psychology of Education*, *30*, 421–435.

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.

Faber, J. M., & Visscher, A. J. (2014). *Digitale leerlingvolgsystemen: een review van de effecten op leerprestaties*. Kennisnet. Retrieved from https://onderwijsdatabank.s3.amazonaws.com/downloads/Digitale_leerlingvolgsystemen_Kennis

net_eindversie_20022014.pdf

Faber, J. M., & Visscher, A. J. (2016). *De effecten van Snappet: Effecten van een adaptief onderwijsplatform op leerresultaten en motivatie van leerlingen*. Enschede: Universiteit Twente.

Fagginger Auer, M. F. (2016). *Solving multiplication and division problems. Latent variable modeling of students' solution strategies and performance.* Proefschrift Universiteit Leiden.

Fan, H., Xu, J., Cai, Z., He, J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year meta-analysis, 1986-2015. *Educational Research Review*, *20*, 35–54. https://doi.org/10.1016/j.edurev.2016.11.003

Flevares, L. M., & Schiff, J. R. (2014). Learning mathematics in two dimensions: A review and look ahead at teaching and learning early childhood mathematics with children's literature. *Frontiers in Psychology*, *5*(May), 1–13. https://doi.org/10.3389/fpsyg.2014.00459

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, *53*(1), 25–44. https://doi.org/10.1016/j.jsp.2014.10.002

Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The math anxiety-performance link. *Current Directions in Psychological Science*, *26*(1), 52–58. https://doi.org/10.1177/0963721416672463

Freire, P. (1974). *Pedagogie van de onderdrukten*. Baarn: In den Toren.

Friso-Van Den Bos, I., Van Der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Heidelberg, J., Van den Bergh, H., Kuhlemeier, H., & Hoeks, J. (1993). Tijdsbesteding van docenten in het voortgezet onderwijs op verschillende manieren gemeten: een valiatiestudie. *Tijdschrift Voor Onderwijsresearch*, *18*, 306–3016.

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.

Houtveen, A. A. M., Van de Grift, W. J. C. M., & Creemers, B. P. M. (2004). Effective School Improvement in Mathematics. *School Effectiveness and School Improvement*, *3453*(3–4), 337–376. https://doi.org/10.1080/09243450512331383242

Hurley, M. M. (2001). Reviewing integrated science and mathematics: The search for evidence and definitions from new perspectives. *School Science and Mathematics*, *101*(May), 259–268. https://doi.org/10.1111/j.1949-8594.2001.tb18028.x

Inspectie voor 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

Jacobse, A. E., & Harskamp, E. (2011). *A Meta-Analysis of the Effects of Instructional Interventions on Students' Mathematics Achievement*. Groningen: GION. Retrieved from https://pdfs.semanticscholar.org/5ed9/f9f0518c51cff0d883d6491fd026d2d6ddfd.pdf

Jansen, B. R. J., Hofman, A. D., Savi, A., Visser, I., & Van der Maas, H. L. J. (2016). Self-adapting the success rate when practicing math. *Learning and Individual Differences*, *51*, 1–10. https://doi.org/10.1016/j.lindif.2016.08.027

Janssen, J., Van der Schoot, F., & Hemker, B. (2005). *Balans van het reken-wiskundeonderwijs aan*

het einde van de basisschool 4. Arnhem.

Janssen, J., & Verhelst, N. (1999). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 3*. Arnhem.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. University of Minnesota, Minneapolis: Cooperative Leaming Center. Retrieved from http://www.tablelearning.com/uploads/File/EXHIBIT-B.pdf

Kablan, Z., Topan, B., & Erkan, B. (2013). The effectiveness level of material use in classroom instruction: A meta-analysis study. *Kuram ve Uygulamada Egitim Bilimleri*, *13*(3), 1638–1644. https://doi.org/10.12738/estp.2013.3.1692

Keuning, T., Van Geel, M., & Visscher, A. (2017). Why a data-based decision-making intervention works in some schools and not in others. *Learning Disabilities Research & Practice*, *32*(1), 32–45. https://doi.org/10.1111/ldrp.12124

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*(4), 28–37. https://doi.org/10.1111/j.1745-3992.2011.00220.x

KNAW. (2009). *Rekenonderwijs op de basisschool: Analyse en sleutels tot verbetering*. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen.

Kolovou, A., Van den Heuvel-Panhuizen, M., & Köller, O. (2013). An intervention including an online game to improve grade 6 students' performance in early algebra. *Journal for Research in Mathematics Education*, *44*(3), 510–549. https://doi.org/10.5951/jresematheduc.44.3.0510

Kolovou, A., Ven den Heuvel Panhuizen, M., & Bakker, A. (2011). Mathematical problem solving in primary school. *Mediterranean Journal for Research in Mathematics Education*, *8*(2), 29–66.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 1–38. https://doi.org/10.3102/0034654315626799

Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, *107*(2), 558–579. https://doi.org/10.1037/a0037698

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, *22*, 215–243. https://doi.org/10.1007/s10648-010-9125-8

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Logtenberg, H., & de Lange, S. (2014). Lesson Study: methodiek voor teamleren bij instructie. *Reken-Wiskundeonderwijs: Onderzoek, Ontwikkeling, Praktijk*, *33*, 67–77.

Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *The Journal of Educational Research*, *94*(2), 101–112. https://doi.org/10.1080/00220670009598748

Marzano, R. J. (2014). *Wat werkt op school*. Rotterdam: Bazalt Educatieve Uitgaven.

Mcclendon Patrick, S. (2013). *A meta-analysis of dissertation research on the relationship between professional learning community implementation and student achievement*. dissertation University of Alabama.

Meelissen, M., & Punter, A. (2016). *Twintig jaar TIMSS. Ontwikkelingen in leerlingprestaties in de*

exacte vakken. Enschede: Universiteit Twente.

Molenaar, I., Van Campen, C., & Van Gorp, K. (2015). *Onderzoek naar Snappet; gebruik en effectiviteit*. Nijmegen: Radboud Universiteit.

Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research*, *74*(3), 317–377.

Peetsma, T., Roeleveld, J., & Stoel, R. (2003). Stabiliteit en verandering in de samenhang tussen psychosociaal functioneren en schoolprestaties gedurende het basisonderwijs. *Pedagogische Studiën*, *80*, 4–23.

Porter, A. C. (2002). Measuring the Content of Instruction: Uses in Research and Practice. *Educational Researcher*, *31*(7), 3–14. https://doi.org/10.3102/0013189X031007003

Rebber, A., Van den Berg, S., & Meelissen, M. (2017). *Secundaire analyses op de data van TIMSS-2015: een nadere analyse van leerkrachtpercepties en de referentieniveaus voor rekenen en natuuronderwijs*. Enschede.

Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles. *ZDM Mathematics Education*, *49*(4), 599–611. https://doi.org/10.1007/s11858-017-0834-z

Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review*, *17*(4), 327–362. https://doi.org/10.1007/s10648-005-8137-2

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psyhology*, *95*(2), 240–257. https://doi.org/10.1037/0022-0663.95.2.240

Roorda, G., Kleefsman, T., & Witterholt, M. (2014). Invloed structureringsactiviteiten bij uitvoeren van samengestelde bewerkingsopdrachten in groep 7. *Reken-Wiskundeonderwijs: Onderzoek, Ontwikkeling, Praktijk*, *33*, 55–61.

Rosli, R., Capraro, M. M., & Capraro, R. M. (2014). The effects of problem posing on student mathematical learning: A meta-analysis. *International Education Studies*, *7*(13), 227–241. https://doi.org/10.5539/ies.v7n13p227

Şad, S. N., Kış, A., & Demir, M. (2016). A meta-analysis of the effect of contemporary learning approaches on students' mathematics achievement. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *32*(1), 209–227. https://doi.org/10.16986/HUJE.2016017222

Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, *19*, 158–172. https://doi.org/10.1016/j.edurev.2016.07.003

Scheerens, J. (2016). Opportunity to learn, instructional alignment and test preparation: a research review. In *Closing the gaps? Differential accountability and effectiveness as a road to school improvement.* (pp. 1–166). Oslo: University of Oslo.

Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). *Review and meta-analysis of school and teaching effectiviness.* Enschede: Universiteit Twente.

Scheltens, F., Hemker, B., & Vermeulen, J. (2013). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 5*. Arnhem: Cito.

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2017). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, 1–25.

https://doi.org/10.1007/s10648-017-9403-9

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does Working Memory Training Transfer? A Meta-Analysis Including Training Conditions as Moderators. *Educational Psychologist*, *50*(2), 138–166. https://doi.org/10.1080/00461520.2015.1036274

Shenderovich, Y., Thurston, A., & Miller, S. (2015). Cross-age tutoring in kindergarten and elementary school settings: A systematic review and meta-analysis. *International Journal of Educational Research*, *76*, 190–210. https://doi.org/10.1016/j.ijer.2015.03.007

Shores, M. L., & Smith, T. (2010). Attribution in mathematics: A review of literature. *School Science and Mathematics*, *110*(1), 24–30. https://doi.org/10.1111/j.1949-8594.2009.00004.x

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Silver, E. (1994). On mathematical problem posing. *For the Learning of Mathematics*, *14*(1), 19–28.

Sins, P. H. M., & Van der Zee, S. (2015). De toegevoegde waarde van traditioneel vernieuwingsonderwijs: een studie naar de verschillen in cognitieve en niet-cognitieve opbrengsten tussen daltonscholen en traditionele scholen voor primair onderwijs. *Pedagogische Studiën*, *92*, 254–273.

Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best evidence synthesis. *Review of Educational Research*, *78*(3), 427–455. https://doi.org/10.1017/CBO9781107415324.004

Sokolowski, A., Li, Y., & Willson, V. (2015). The effects of using exploratory computerized environments in grades 1 to 8 mathematics: A meta-analysis of research. *International Journal of STEM Education*, *2*(1), 8. https://doi.org/10.1186/s40594-015-0022-z

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, *105*(4), 970–987. https://doi.org/10.1037/a0032447

Tingir, S., Cavlazoglu, B., Caliskan, O., Koklu, O., & Intepe-Tingir, S. (2017). Effects of mobile devices on K-12 students' achievement: A meta-analysis. *Journal of Computer Assisted Learning*, 1–15. https://doi.org/10.1111/jcal.12184

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 den Heuvel-Panhuizen, M., Kolovou, A., & Robitzsch, A. (2013). Primary school students' strategies in early algebra problem solving supported by an online game. *Educational Studies in Mathematics*, *84*, 281–307. https://doi.org/10.1007/s10649-013-9483-5

Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, *85*(4), 475–511. https://doi.org/10.3102/0034654314564881

Van der Ploeg, P. (2010). *Dalton Plan: Oorsprong en theorie van het daltononderwijs*. Deventer: Saxion Dalton University Press.

Van der Ven, F., Segers, E., Takashima, A., & Verhoeven, L. (2017). Effects of a tablet game intervention on simple addition and substraction fluency in first graders. *Computers in Human Behavior*, *72*, 200–207.

Van Weerden, J., Janssen, J., & Scheltens, F. (2016). Peilingsonderzoek rekenenwiskunde in het basisonderwijs – Balans van 25 jaar PPON. *Volgens Bartjens*, (35), 41–50.

Van Zanten, M., & van den Heuvel-Panhuizen, M. (n.d.). Primary School Mathematics in the Netherlands: The Perspective of the Curriculum Documents. In D. Thompson, C. Suurtamm, & M. Huntley (Eds.), *International Perspectives on Mathematics Curriculum*. harlotte, NC, USA: Information Age Publishing.

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.* Dordrecht: Springer.

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.

Walshaw, M., & Anthony, G. (2008). Creating productive learning communities in the mathematics classroom: An international literature review. *Pedagogies: An International Journal*, *3*(3), 133–149. https://doi.org/10.1080/15544800802026595

Walshaw, M., & Anthony, G. (2013). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, *78*(3), 516–551.

Wijnstra, J. M. (1988). *Balans van het rekenonderwijs in de basisschool. Uitkomsten van de eerste rekenpeiling medio en einde basisonderwijs*. Anrhem: Cito.

Young, M. F., Slota, S., Cutter, B., R., Jalette, G., Mullin, G., … Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research*, *82*(1), 61–89. https://doi.org/10.3102/0034654312436980

Zheng, B., Warschauer, M., Lin, C.-H., & Chang, C. (2016). Learning in one-to-one laptop environments: A meta-analysis and research synthesis. *Review of Educational Research*, *86*(4), 1–33. https://doi.org/10.3102/0034654316628645

Bijlage: Leerkrachtvariabelen in nadere analyses PPON-2011

Algemene factoren

- Groepsgrootte (20 of minder / 21-24 / 25-29 / 30 of meer)
- Leerkracht: aantal jaren ervaring in groep 8 (5 of minder / 6-9 / 10 of meer)
- Leerkracht: vakspecifieke bijscholing in de afgelopen vijf jaar (nee / ja)
- Leerkracht: geslacht (man / vrouw)
- Leerkracht: leeftijd in jaren

Rekenspecifieke schoolfactoren

- Rekenmethode (Wereld in Getallen / Pluspunt / Rekenrijk / Alles telt / overig)
- Aanwezigheid mogelijkheid extra individuele ondersteuning rekenen (nee / ja)
- Percentage leerlingen dat extra ondersteuning ontvangt
- Intensiteit ondersteuning leerlingen thuis (niet – altijd; schaal 1-5)
- Percentage leerlingen dat externe hulp krijgt
- Tevredenheid ondersteuning van school, thuis en extern (zeer tevreden – zeer ontevreden; schaal 1 – 5)

Rekenspecifieke leerkrachtfactoren

- Begrip van denkwijze matige leerling versus gemiddelde / goede leerling
- Gebruik van computersoftware bij rekenlessen (nee / ja)
- Passen van rekenmethode bij leerkracht (zeer slecht - heel goed; schaal 1-5)
- Belang klassikale instructie (erg onbelangrijk - erg belangrijk; schaal 1-5)
- Belang (extra) individuele instructie (erg onbelangrijk - erg belangrijk; schaal 1-5)
- Belang zelfstandig werken (erg onbelangrijk - erg belangrijk; schaal 1-5)
- Mate van differentiatie naar niveau en/of tempo (in het algemeen krijgen alle leerlingen tegelijkertijd dezelfde instructie en oefenstof / de instructie is in het algemeen voor alle leerlingen gelijk; bij de verwerking van de oefenstof wordt gedifferentieerd naar niveau en tempo / de instructie wordt per niveau- of tempogroep gegeven, eventueel met verdere differentiatie bij de verwerking van de oefenstof / de instructie wordt individueel gegeven en de oefenstof wordt per leerling bepaald)
- Frequentie stellen klassikale vragen (minder dan 1 keer per maand – minstens 1 keer per week; schaal 1-5)
- Frequentie leerlingen uitwerkingen op bord laten schrijven vragen (minder dan 1 keer per maand – minstens 1 keer per week; schaal 1-5)
- Frequentie vragen hoe leerlingen op antwoord zijn gekomen vragen (minder dan 1 keer per maand – minstens 1 keer per week; schaal 1-5)
- Frequentie veelgemaakte fouten klassikaal bespreken vragen (minder dan 1 keer per maand – minstens 1 keer per week; schaal 1-5)
- Aantal functies van de rekenmachine dat wordt geïnstrueerd

Onderwijstijd

- Uren per week besteed aan rekenen (4 / 5 / 6 / 7 of meer)
- Aantal minuten extra leertijd voor zwakste leerlingen
- Relatieve hoeveelheid lestijd besteed aan getallen en bewerkingen
- Relatieve hoeveelheid lestijd besteed aan verhoudingen, breuken en procenten
- Relatieve hoeveelheid lestijd besteed aan meten, meetkunde tijd en geld
- Relatieve hoeveelheid lestijd besteed aan klassikale instructie
- Relatieve hoeveelheid lestijd besteed aan (extra) individuele instructie
- Relatieve hoeveelheid lestijd besteed aan zelfstandig werk

Rekenspecifieke-domeinspecifieke leerkrachtfactoren: hoofd- en schattend rekenen

- Aantal strategieën voor hoofdrekenend optellen en aftrekken (beide één strategie / een van beide twee strategieën / beide twee strategieën)
- Aantal strategieën voor hoofdrekenend vermenigvuldigen en delen (beide één strategie / een van beide twee strategieën / beide twee strategieën)
- Gebruik aanvullend materiaal bij lessen over hoofdrekenen of schattend rekenen (nee / ja)
- Minuten per week besteed aan hoofdrekenen en schattend rekenen
- Frequentie aandacht aan basisvaardigheden hoofdrekenen en schattend rekenen
- Frequentie aandacht zoeken en hanteren handige oplossingsstrategieën
- Frequentie aandacht hanteren meerdere oplossingsstrategieën voor één type opgave
- Frequentie aandacht schattend rekenen
- Frequentie aandacht aan hoofdrekenen en schattend rekenen in het algemeen

Rekenspecifieke-domeinspecifieke leerkrachtfactoren: strategieën bewerkingen

- Aantal strategieën voor bewerkingen optellen en aftrekken (beide één strategie / een van beide twee strategieën / beide twee strategieën)
- Aantal strategieën voor bewerkingen vermenigvuldigen en delen (beide één strategie / een van beide twee strategieën / beide twee strategieën)
- Proportie cijferende strategieën van totale hoeveelheid strategieën voor bewerkingen optellen en aftrekken
- Proportie cijferende strategieën van totale hoeveelheid strategieën voor bewerkingen vermenigvuldigen en delen
- Persoonlijke voorkeur voor cijferende versus kolomsgewijze strategieën voor optellen en aftrekken
- Persoonlijke voorkeur voor cijferende versus kolomsgewijze strategieën voor vermenigvuldigen en delen