

# MASTER'S THESIS

## Het Effect van Peerdiscussies bij Formatieve Toetsing op de Rekenresultaten en Motivatie van Bovenbouwleerlingen in het Primair Onderwijs.

Peters, Driek

**Award date:**  
2024

**Awarding institution:**  
Faculty of Educational Sciences

[Link to publication](#)

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### Take down policy

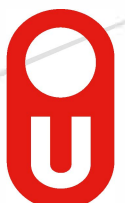
If you believe that this document breaches copyright please contact us at:

[pure-support@ou.nl](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 19. Jun. 2024

**Open Universiteit**  
[www.ou.nl](http://www.ou.nl)



**Het Effect van Peerdiscussies bij Formatieve Toetsing op de Rekenresultaten en  
Motivatie van Bovenbouwleerlingen in het Primair Onderwijs.**

**The Effect of Peer Discussions in Formative Assessment on Upper Primary School  
Students' Maths Achievement and Motivation.**

Driek Peters

Master Onderwijswetenschappen, Open Universiteit

E-mailadres: [driekpeters@hotmail.com](mailto:driekpeters@hotmail.com)

Cursuscode en cursusnaam: OM9906 Masterscriptie

Naam begeleider: Dr. François Molin

Woordenaantal: 8876

Datum: 29 - 1 - 2024

### Samenvatting

Rekenscores van leerlingen in Nederland op het eind van het primair onderwijs voldoen niet aan de maatschappelijke ambities. Een mogelijke oplossing om deze rekenresultaten te verbeteren ligt binnen het concept 'formatief toetsen'. Vaak worden hierbij peerdiscussies ingezet, waarbij leerlingen in kleine groepjes kort met elkaar bespreken waarom ze voor een bepaald antwoord hebben gekozen. Het voeren van peerdiscussies bij formatieve toetsing kan zorgen voor leerwinst en een hogere motivatie. Voor dit quasi-experimenteel onderzoek, waaraan 76 basisschoolleerlingen in de leeftijd van 10-12 jaar deel hebben genomen, is een interventie ontworpen waarin formatieve toetsing met peerdiscussies twee keer per week is ingezet tijdens de herhalingslessen van rekenen. De interventie duurde zeven weken. Er is onderzocht wat het effect is van het voeren van peerdiscussies op rekenresultaten en rekenmotivatie. Daarnaast is gekeken of rekenmotivatie een mediator is tussen het voeren van peerdiscussies en de rekenresultaten. De onderzoeksvragen zijn beantwoord middels de *Causal Steps* methode van Baron en Kenny. De resultaten tonen dat het voeren van peerdiscussies een significant positieve invloed heeft op de rekenresultaten en de rekenmotivatie. Er is sprake van een volledige mediatie van motivatie in de relatie tussen het voeren van peerdiscussies en de rekenresultaten: het voeren van peerdiscussies is een effectieve interventie om via het verhogen van de rekenmotivatie uiteindelijk de rekenresultaten te verbeteren. Binnen het primair onderwijs is het voeren van peerdiscussies bij rekenen aan te bevelen, zodat rekenproblemen in de basis aangepakt kunnen worden en aan de maatschappelijke ambitie kan worden voldaan.

*Keywords:* formatief toetsen, peerdiscussies, rekenresultaten, rekenmotivatie, student response systems (SRS)

### **Abstract**

Mathematics results of students in the Netherlands at the end of primary education do not meet society's ambitions. A possible solution to improve these mathematics results lies within the concept of 'formative testing'. This often involves peer discussions, in which students in small groups briefly discuss why they chose a certain answer. Peer discussions in formative assessment can lead to learning gains and higher motivation. For this quasi-experimental study, in which 76 primary school students aged 10-12 participated, an intervention was designed in which formative assessment with peer discussions was used twice a week during math lessons. The intervention lasted for seven weeks. The effect of using peer discussions on mathematics results and on math motivation was investigated. It was also investigated whether motivation is a mediator between peer discussions and mathematics results. The questions from this study were answered using Baron and Kenny's Causal Steps method. The results show that using peer discussions has a significant positive influence on both mathematics results and math motivation. Motivation completely mediates the relationship between conducting peer discussions and mathematics results: using peer discussions is an effective intervention to ultimately increase the mathematics results by increasing math motivation. In primary education, using peer discussions in mathematics is recommended so that math problems can be tackled at a young age and the social ambition can be met in the future.

*Keywords:* formative testing, peer discussions, mathematics results, math motivation, student response systems (SRS)

## Inhoud

Samenvatting .....	2
Abstract .....	3
Inhoud.....	4
1. Inleiding .....	5
1.1 Probleemschets en Doel.....	5
1.2 Theoretisch Kader.....	7
1.3 Huidige Studie .....	12
2. Methode.....	14
2.1 Deelnemers .....	14
2.2 Meetinstrumenten en Materialen .....	15
2.3 Procedure .....	21
2.4 Data-Analyse .....	23
2.5 Beschrijvende Statistiek van de Conditie.....	26
2.6 Interne Consistentie Meetinstrument.....	27
3. Resultaten .....	27
3.1 Statistische assumpties .....	27
3.2 Resultaten Statistische Analyses .....	28
4. Discussie.....	30
4.1 Antwoord Onderzoeksvragen .....	31
4.2 Beperkingen van het Onderzoek en Toekomstig Onderzoek .....	34
4.3 Maatschappelijke en Wetenschappelijke Significantie .....	35
4.4 Conclusie .....	36
Referenties.....	38
Bijlage .....	53

## Titel

Het Effect van Peerdiscussies bij Formatieve Toetsing op de Rekenresultaten en Motivatie van Bovenbouwleerlingen in het Primair Onderwijs

## 1. Inleiding

### 1.1 Probleemschets en Doel

In de huidige samenleving is het van belang om goed te kunnen rekenen. Dit blijft het hele leven belangrijk, omdat de meeste informatie wordt aangeboden en opgeslagen in numerieke vorm (Ruijsenaars et al., 2014). Basiskennis en -vaardigheden hebben van rekenen zijn daarom voorwaardelijk om volledig deel te kunnen nemen aan de huidige samenleving (van den Berg et al., 2017). Wanneer het rekenen niet vanzelf gaat, kan dit aanzienlijke beperkingen opleveren in het dagelijks functioneren (Ruijsenaars et al., 2014). Het leggen van een goede rekenbasis op vroege leeftijd is daarom cruciaal (Aubrey et al., 2006).

Uit een peilings-onderzoek van de Inspectie van het Onderwijs (2023) blijkt dat 42,2% van de leerlingen in Nederland op het eind van het primair onderwijs het streefniveau voor rekenen behaalt. Dit in tegenstelling tot de maatschappelijke ambitie van 65% vanuit de commissie Meijerink (van der Lubbe, 2021). De afgelopen jaren zijn leerlingen minder vooruitgegaan in rekenvaardigheid dan de jaren daarvoor (Inspectie van het Onderwijs, 2023). Omdat de rekenresultaten van Nederlandse basisschoolleerlingen niet aan de maatschappelijke ambities voldoen, moet er gezocht worden naar manieren om deze trend te buigen en de prestaties op het gebied van rekenen te verbeteren (Kiplagat, 2016).

Niet alleen de rekenresultaten dalen, ook de motivatie van leerlingen neemt af. Longitudinale studies tonen aan dat de motivatie van leerlingen afneemt naarmate leerlingen ouder worden (bijv. Gnams & Hanfstingl, 2016). Uit onderzoek blijkt dat de daling in motivatie voor het vak rekenen/wiskunde het grootst is (Gottfried et al., 2001, 2007). De

daling in motivatie is van invloed op de prestaties van leerlingen (Ryan et al., 2021; Vansteenkiste et al., 2006).

Een mogelijke oplossing voor de dalende rekenresultaten en de dalende motivatie ligt binnen het concept 'formatief toetsen' (van den Berg et al., 2017). Volgens Black en Wiliam (1998) omvat formatieve toetsing alle activiteiten die de leerkracht en leerlingen informatie opleveren over de mate waarin leerlingen de leerstof begrijpen. Deze informatie wordt vervolgens gebruikt als feedback om onderwijs- en leeractiviteiten aan te passen. Uit de literatuur blijkt dat formatief toetsen het potentieel heeft om de motivatie te verhogen (Cauley & McMillan, 2010; Leenknecht et al., 2020) en prestaties te verbeteren (Andersson & Palm, 2018; Cauley & McMillan, 2010). Akay (2011), Allison (2012) en Lee et al. (2020) hebben aangetoond dat formatief toetsen een positief effect heeft op rekenresultaten van leerlingen in het primair onderwijs.

Leerkrachten kunnen formatieve toetsing op verschillende manieren organiseren. Vaak wordt formatieve toetsing zo georganiseerd dat de leerling een actieve rol speelt (Schildkamp et al., 2014). Deze actieve rol wordt bijvoorbeeld gerealiseerd door peerdiscussies, waarbij de leerlingen kort hun antwoord bespreken met elkaar (Strijbos & Wichmann, 2017). Verschillende onderzoeken hebben aangetoond dat er leerwinst wordt behaald door het voeren van peerdiscussies (Egelandsdal & Krumsvik, 2017; Smith et al., 2011; Tullis & Goldstone, 2020). Dit effect is ook bij rekenen gevonden: het voeren van peerdiscussies stimuleert de motivatie om rekenen/wiskunde te leren en vergroot het begrip van wiskundeconcepten (Oloo et al., 2016).

Het vergroten van het rekenbegrip en het stimuleren van motivatie is nodig. Het doel van dit onderzoek is om na te gaan of het voeren van peerdiscussies bij formatieve toetsing ervoor kan zorgen dat rekenresultaten en motivatie van bovenbouwleerlingen uit het primair onderwijs verbeteren. De bijdrage van deze studie is tweeledig. Ten eerste wordt de studie

uitgevoerd in het primair onderwijs. Volgens Palm et al. (2017) is er een tekort aan studies gericht op formatieve toetsing in combinatie met peerdiscussies bij rekenen. Studies waarin wel peerdiscussies zijn ingezet bij het vak rekenen, hebben hoofdzakelijk plaatsgevonden in het voortgezet onderwijs en in een universitaire setting. Ten tweede wordt in deze studie een mediatieanalyse uitgevoerd, waarin het effect van motivatie tussen het voeren van peerdiscussies tijdens formatieve toetsing en leerprestaties wordt onderzocht. Zover we kunnen nagaan heeft er geen enkele studie plaatsgevonden die dit effect bij rekenen heeft onderzocht. Dit onderzoek tracht deze leemtes in de huidige literatuur te vullen.

## **1.2 Theoretisch Kader**

### ***1.2.1 Formatief Toetsen Middels Peerdiscussies***

Formatief toetsen is een cyclisch proces dat gericht is op het in kaart brengen van het leerproces om de voortgang van het leren te volgen en om feedback te genereren voor de leerkracht en de leerling (Black & Wiliam, 1998). De leerling ontdekt door formatieve toetsen of hij de leerstof beheerst en de leerkracht ziet welke fouten er worden gemaakt en kan vervolgens passende feedback geven (McCarthy, 2017; Sluijsmans & Kneyber, 2016). De kern van formatieve toetsing is dus het geven van feedback (Hattie & Timperley, 2007). Feedback over het begrip van lesinhoud is volgens González et al. (2016) onmisbaar in het leerproces. Volgens Konold et al. (2004) volgt feedback van de leerkracht op een actie van een leerling en geeft het vorm aan toekomstig gedrag. Het geven van feedback is daarom cruciaal voor het leerproces. Feedback hoeft niet per se door de leerkracht gegeven te worden. Leerlingen kunnen elkaar ook van feedback voorzien. Dit wordt peerfeedback genoemd (Liu & Carless, 2006).

Het geven van peerfeedback is cognitief veeleisend, omdat het verder gaat dan het ophalen van kennis. Het vereist meer verfijnde processen, zoals het geven van uitleg (McConlogue, 2015), het rechtvaardigen van antwoorden (Cho & Cho, 2011) en het



vergelijken en samen oplossen van problemen (Nicol et al., 2014). Het gaat daarnaast ook om metacognitieve processen zoals evaluatie (Nicol, 2010), monitoring (Nicol et al., 2014) en reflectie op het eigen leren (Evans, 2015). Onderzoek heeft aangetoond dat peerfeedback een positief effect heeft op leerresultaten (Topping, 1998). Een van de vormen van peerfeedback is het voeren van peerdiscussies.

Tijdens peerdiscussies bespreken leerlingen ideeën of antwoorden op vragen met elkaar (Knight & Brame, 2018) en ontdekken hoeveel ze weten door hun denkstappen hardop te verwoorden (Hoekstra, 2008). Door het voeren van peerdiscussies bouwen leerlingen voort op hun eigen begrip of op het idee van een medeleerling en kan er sprake zijn van onderhandeling en/of co-constructie van ideeën (Bojinova & Oigara, 2011; Dawson et al., 2018; Nicol & Boyle, 2003). Het voeren van peerdiscussies in het klaslokaal zorgt dus voor een verschuiving van eenrichtingsoverdracht van de leerkracht naar een meer collaboratieve, actieve leeromgeving (Porter et al., 2011). Leervoordelen komen voort uit het zien van andere voorbeelden of andere benaderingen om tot een antwoord te komen (McCarthy, 2017). Het resultaat van dit proces is versterking van het begrip (Barth-Cohen et al., 2016; Vickrey et al., 2015). Daarnaast blijkt uit onderzoek van Nicol en Boyle (2003) en Perez et al. (2010) dat het voor leerlingen veelal gemakkelijker is om een concept te begrijpen dat in eenvoudige taal wordt uitgelegd door een medeleerling, dan dat het wordt uitgelegd door de leerkracht die zich op een veel hoger kennisniveau bevindt. In verschillende onderzoeken, veelal uitgevoerd in een universitaire setting, is aangetoond dat er betere resultaten worden behaald door het voeren van peerdiscussies (Egelandsdal & Krumsvik, 2017; Smith et al., 2011; Tullis & Goldstone, 2020) en dat een combinatie van peerdiscussies, gevolgd door feedback van de leerkracht een positiever effect heeft op de resultaten dan alleen het voeren van peerdiscussies of alleen maar feedback krijgen van de leerkracht (Smith et al., 2011). Feedback van de leerkracht is vaak uitgebreider dan die van een klasgenoot, bevat uitleg over waarom andere

antwoorden onjuist zijn en is daarom een goede toevoeging op de peerdiscussie (Barth-Cohen et al., 2016; Levesque, 2011; Smith et al., 2011). Peerdiscussies voeren heeft niet alleen een positieve invloed op leerprestaties, ook de motivatie van leerlingen neemt hierdoor toe (Bojinova & Oigara, 2011).

Motivatie verwijst naar processen die doelgerichte activiteiten in gang zetten en in stand houden. Motiverende processen zijn interne invloeden die leiden tot inspanning, volharding en uiteindelijk tot betere leerprestaties (Schunk & DiBenedetto, 2020). Bij peerdiscussies is er sprake van interactie, waardoor leerlingen actief betrokken worden in het leerproces (Deci & Ryan, 1985; Wentzel et al., 2010). Daarnaast kunnen leerlingen relaties opbouwen en elkaar ondersteunen, wat hun gevoel van saamhorigheid en verbondenheid versterkt (Johnson & Johnson, 1999). Ook kan het voeren van peerdiscussies helpen om een gevoel van competitie te creëren (Topping, 1998). Deze aspecten hebben een positieve invloed op de motivatie van leerlingen (Deci & Ryan, 2004; Lewin et al., 2016) en verbeteren leerprestaties (Edwards et al., 2015). Als leerlingen gemotiveerd en betrokken zijn in de klas, leren ze met plezier, wat een positieve invloed kan hebben op de leerprestaties (Amrai et al., 2011). Een veelgebruikte strategie om peerdiscussies te voeren bij formatieve toetsing is het inzetten van de methode *Peer Instruction* (PI; Mazur, 1997) middels *student response systems* (SRS).

### ***1.2.2 Peerdiscussies Middels PI met Student Response Systems (SRS)***

PI is een werkwijze die gebruikt kan worden bij het inzetten van peerdiscussies (Mazur, 1997). De PI-strategie bestaat uit drie stappen (Duncan, 2005; Mazur 1997). In de eerste stap wordt voorin de klas op het bord een meerkeuzevraag getoond die de leerlingen individueel beantwoorden. Het individueel beantwoorden geeft leerlingen de kans eerst zelf na te denken over een mogelijke oplossing van een vraag. Dit is niet alleen belangrijk om de leerling voor te bereiden op de uitleg, maar ook noodzakelijk om de leerling voor te bereiden

op het bijdragen aan de peerdiscussie (Simon & Cutts, 2012). Na het beantwoorden van de meerkeuzevraag toont de leerkracht vooraan in de klas de histogramweergave van de gegeven antwoorden. Het voordeel hiervan is dat leerlingen hun eigen antwoorden gepositioneerd zien in relatie tot de antwoorden van hun medeleerlingen, waardoor ze hun eigen vooruitgang kunnen controleren of bevestiging krijgen dat ze niet de enige zijn met een misvatting (Kay & LeSage, 2009).

In de tweede stap wordt leerlingen gevraagd hun gegeven antwoorden kort (2 tot 3 minuten) te bediscussiëren met hun *peers*, waarbij de leerlingen de stappen die ze zetten om tot het antwoord te komen hardop verwoorden aan elkaar. De leerkracht loopt tijdens deze peerdiscussies rond en luistert naar de gesprekken. Zo kan de leerkracht eventuele misconcepties op het spoor komen, waarop later feedback kan worden gegeven (Simon & Cutts, 2012).

In de derde en laatste fase beantwoorden de leerlingen individueel de vraag opnieuw, waarbij ze moeten beslissen of ze hetzelfde antwoord of een ander antwoord kiezen (Jones et al., 2012). De leerkracht geeft daarna feedback aan de leerlingen over waarom de antwoordopties juist of onjuist zijn (Chang et al., 2015).

Wanneer PI in de klas wordt ingezet, verzamelt de leerkracht de antwoorden van de leerlingen. Dit kan via low-tech of high-tech methoden. Bij low-tech methoden geven leerlingen hun antwoordkeuze aan door bijvoorbeeld hun hand op te steken (Bartsch & Murphy, 2011), of antwoordkaarten op te steken (Fallon & Forrest, 2011). Volgens Caldwell (2007) hebben deze low-tech methoden een nadeel: het is voor een leerkracht moeilijk om het aantal stemmen in een beperkte tijd te schatten. Bovendien zullen leerlingen door een gebrek aan anonimiteit niet altijd eerlijk antwoorden (Kay & LeSage, 2009). Om formatieve toetsing via PI makkelijker en efficiënter in te zetten, worden daarom vaak high-tech methoden ingezet, zoals *Student Response Systems* (SRS; Hunsu et al., 2016). Een SRS (bijv. *Clickers*,

*Socrative* of *Kahoot*) is een digitale applicatie die het mogelijk maakt dat alle leerlingen in een klas tegelijk de meerkeuzevraag kunnen beantwoorden, waarbij iedere leerling een *clicker*, laptop of mobiele telefoon heeft waarmee anoniem een antwoord kan worden gekozen (Egelandsdal & Krumsvik, 2015). Het voordeel van het gebruik van SRS is dat het direct informatie presenteert aan de leerkracht en de leerlingen over het begripsniveau in de klas (Egelandsdal & Krumsvik, 2015; Irving, 2015). Het inzetten van SRS bij formatieve toetsing via de PI-methode wordt benoemd als krachtige formatieve toetspraktijk (Sluijsmans et al., 2013).

Er zijn verschillende onderzoeken uitgevoerd waarbij peerdiscussies middels PI zijn ingezet bij het vak rekenen/wiskunde. Kaymak et al. (2020) en Cronhjort et al. (2013) hebben dit in een universitaire setting onderzocht en tonen aan dat studenten in lessen waar peerdiscussies gevoerd worden hogere resultaten behalen. Ook Abdelkarim en Abuiyada (2016) prijzen het toepassen van peerdiscussies bij rekenen/wiskunde. Het is volgens hen een goede manier om wiskundige prestaties op universitair niveau te verbeteren. Daarnaast is gebleken dat het voeren van peerdiscussies binnen het hoger onderwijs het vermogen van leerlingen verbetert om wiskundige concepten toe te passen op reële problemen en om hun redeneringen effectief te communiceren (Henderson et al., 2011). Ook bij leerlingen op de middelbare school zorgt het voeren van peerdiscussies bij rekenen/wiskunde ervoor dat er een dieper begrip van wiskundige concepten wordt ontwikkeld en dat dit begrip na verloop van tijd vastgehouden wordt (Dawson & Venville, 2010). Enkele onderzoeken die plaats hebben gevonden in het primair onderwijs betreffen de onderzoeken van Akay (2011) en Allison (2012). Zij vonden beiden positieve effecten op de rekenresultaten na het inzetten van peerdiscussies middels PI bij het vak rekenen (Akay 2011; Allison, 2012). Allison (2012) toonde daarnaast ook aan dat de motivatie positief beïnvloed werd door het inzetten van PI.

### 1.3 Huidige Studie

Het doel van dit onderzoek is nagaan of peerdiscussies bij formatieve toetsing ervoor kunnen zorgen dat rekenresultaten en motivatie van bovenbouwleerlingen in het primair onderwijs verbeteren. In dit onderzoek staan drie onderzoeksvragen centraal:

1. ‘Wat is het effect van peerdiscussies bij formatieve toetsing op de rekenresultaten?’

Peerdiscussies voeren kan een significant positief effect hebben op de rekenresultaten (Akay, 2011; Abdelkarim & Abuiyada, 2016; Allison, 2012). Er wordt daarom verwacht dat leerlingen die peerdiscussies voeren bij het formatief toetsen significant betere rekenresultaten behalen dan leerlingen die formatief toetsen zonder peerdiscussies (H1).

2. ‘In hoeverre leiden peerdiscussies tijdens formatieve toetsing tot een hogere motivatie voor rekenen?’

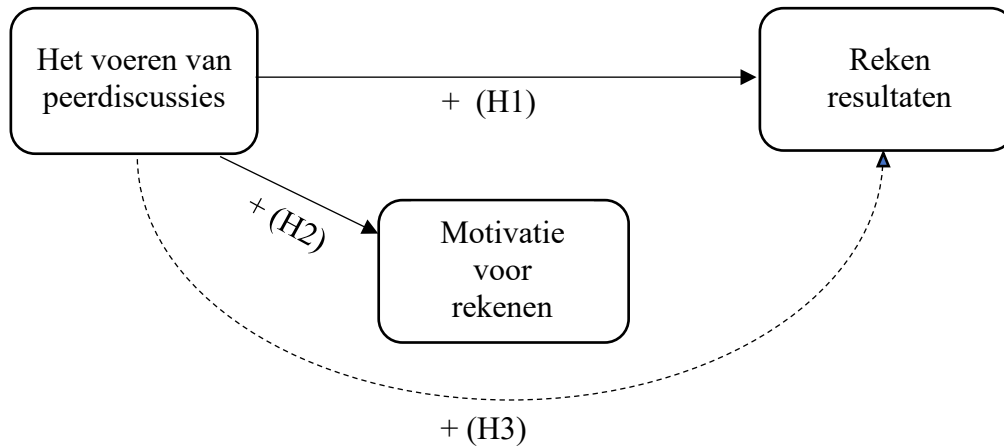
Het voeren van peerdiscussies bij formatieve toetsing kan ervoor zorgen dat de motivatie van leerlingen toeneemt (Akay, 2011; Allison, 2012; Edwards et al., 2015). Er wordt daarom verwacht dat het voeren van peerdiscussies de motivatie voor het vak rekenen verhoogt (H2).

3. ‘In welke mate beïnvloedt motivatie de relatie tussen het voeren van peerdiscussies tijdens formatieve toetsing en rekenresultaten?’

Omdat leerprestaties deels verklaard kunnen worden door motivatie (Amrai et al., 2011), wordt met deze onderzoeksvraag nagegaan of – en in welke mate – motivatie een rol speelt in de relatie tussen het voeren van peerdiscussies en de rekenresultaten. Er wordt verwacht dat motivatie deels of volledig van invloed is op de relatie tussen het voeren van peerdiscussies tijdens formatieve toetsing en de rekenresultaten (H3). In Figuur 1 zijn de variabelen schematisch opgenomen in een conceptueel model.

**Figuur 1**

*Conceptueel Model: Veronderstelde Relatie Variabelen*



Het betreft een quasi-experimenteel onderzoek (Creswell & Guetterman, 2021), gezien de participanten deelnemen vanuit hun eigen klassensituatie en dus niet random toegewezen worden aan de experimentele groep of de controlegroep. Het voeren van peerdiscussies wordt opgenomen als onafhankelijke variabele met twee condities: wel of geen peerdiscussies voeren bij de formatieve toetsing. De rekenresultaten worden als afhankelijke variabele opgenomen. Motivatie wordt als mediërende variabele opgenomen. Er is sprake van een *mixed design* (Creswell & Guetterman, 2021). Er vindt een meting plaats tussen de twee condities (*between-subjects design*), waarbij het effect van peerdiscussies op de rekenresultaten en op motivatie uit de experimentele groep wordt vergeleken met de rekenresultaten en de motivatie van de controlegroep waar geen peerdiscussies plaatsvinden. Ook vinden er metingen binnen de condities plaats (*within-subjects design*), waarbij de rekenresultaten en de motivatie voor rekenen kwantitatief in kaart worden gebracht middels een *pre-* en *posttest*. In Tabel 1 is het onderzoeksontwerp weergegeven.

**Tabel 1***Onderzoeksontwerp*

	<i>Pretest (T0)</i>	Interventie	<i>Posttest (T1)</i>
Controlegroep	O1		O2
(Formatieve toetsing zonder peerdiscussies)	O3		O3
Experimentele groep			
(Formatieve toetsing met peerdiscussies)	O1	X	O2
	O3		O3

*Noot:* O1 is de Cito rekenen E7, O2 is de Cito rekenen B8 en O3 betreft de motivatie-vragenlijst.

## 2. Methode

### 2.1 Deelnemers

Het voeren van peerdiscussies is een krachtige manier om het leren te verbeteren (Fisher et al., 2016). In meta-analyses noemt Hattie (2008) een effectgrootte  $d = .82$  en in een meta-analyse van Balta et al. (2017), waarin gekeken is naar het effect van peerdiscussies op leerwinst, wordt een effectgrootte van  $d = .70$  genoemd.

Uit de poweranalyse, a priori uitgevoerd met *G\*power* door middel van een T-test: *Means: difference between two independent means (two groups)*, volgde een aanbevolen minimale steekproef van 68 participanten. Bij deze analyse is gekozen voor *two tails* en een effectgrootte van Cohen's  $d = .70$ . Verder is gekozen voor een power van  $1 - \beta = .80$  en de meest gebruikte  $p$ -waarde van  $\alpha = .05$  (Field, 2018). Er wordt gecontroleerd voor andere variabelen, zoals geslacht en leeftijd, waardoor de power nog verder toe kan nemen. Een minimumpower van 0.80 wordt als adequaat beschouwd (Akobeng, 2016), waar met bovenstaande poweranalyse aan is voldaan.

Voor dit onderzoek zijn twee verschillende basisscholen benaderd in het zuiden van het land. Beide scholen liggen in hetzelfde dorp. De benaderde scholen vallen onder hetzelfde

bestuur en werken met dezelfde methoden. De scholen zijn ongeveer even groot<sup>1</sup> en hebben bovenbouwgroepen met gemiddeld 25 leerlingen in de leeftijd van 10 tot 12 jaar. Aan dit onderzoek hebben vier groepen 8 deelgenomen. De vier leerkrachten die deel hebben genomen aan het onderzoek hebben allen tussen de 5-10 jaar leservaring en geven meerdere jaren les aan groep 8. De deelnemende groepen 8 zijn als geheel aan de experimentele- of controlegroep toegewezen. Op iedere school is er één groep toegewezen aan de controleconditie en één groep aan de experimentele conditie. Dit is gedaan middels het trekken van lootjes.

## **2.2 Meetinstrumenten en Materialen**

### **2.2.1 Rekenresultaten**

In dit onderzoek zijn de Cito-toetsen eind groep 7 (E7) en begin groep 8 (B8) gebruikt. De data van de E7-toets is bekeken om te bepalen of de groepen bij aanvang van het onderzoek significant van elkaar verschilden. De data van de B8-toets werd meegenomen in de statistische analyses als afhankelijke variabele. Beide toetsen waren gelijk voor alle deelnemers, waarbij de vaardigheidsscore als rekenresultaat is meegenomen. De vaardigheidsscore is het rekenresultaat van individuele leerlingen en is een intervalvariabele met een score van 1 tot en met 380. Hoe hoger deze vaardigheidsscore, hoe beter de leerling kan rekenen. De toetsen zijn op papier afgenomen en de resultaten zijn, nadat deze handmatig werden ingevoerd, beschikbaar gesteld in het Cito leerlingvolgsysteem 'Leerling in Beeld'.

Om de betrouwbaarheid van de Cito-toetsen te meten, is het programma OPLAT gebruikt (Hop et al., 2017). Met dit programma is een betrouwbaarheidscoëfficiënt berekend volgens de klassieke testtheorie. De betrouwbaarheid wordt aangeduid als '*Accuracy of Measurement*' (MAcc) en wordt als goed beschouwd vanaf 0.80. Uit onderzoek blijkt dat

---

<sup>1</sup> Beide scholen hebben ongeveer 400 leerlingen en 30 leerkrachten. De leeftijd van de leerkrachten varieert van 23-65 jaar. In de groepen 8 op beide scholen staan leerkrachten jonger dan 30 jaar voor de groep.



zowel de E7-toets als de B8-toets een betrouwbaarheidscoëfficiënt heeft van  $MAcc = 0.96$  (Hop et al. 2017; 2019).

### 2.2.2 *Motivatie*

Om de motivatie voor het vak rekenen te meten is gebruikgemaakt van een subschaal uit de *Attitudes Toward Mathematics Inventory* (ATMI; Tapia & Marsh, 2004). De ATMI is ontworpen om verschillende dimensies te beoordelen ten opzichte van rekenen/wiskunde (Tapia, 1996). De ATMI is gebruikt bij verschillende populaties in verschillende landen om te testen of het instrument betrouwbaar en valide is. De ATMI werd steeds valide bevonden om de motivatie voor rekenen te meten (bijv. Castillo et al., 2022; Majeed et al., 2013; Romero & Angeles, 2023). In dit onderzoek is gekozen om de subschaal *Enjoyment* (plezier) te gebruiken voor het meten van motivatie. Volgens de *Self determination theory* (SDT) van Deci en Ryan (1985) wordt plezier vaak gekoppeld aan intrinsieke motivatie. Plezier ontstaat volgens hen wanneer individuen binnen de activiteit een gevoel van autonomie, competentie en verbondenheid ervaren. Dit zorgt er uiteindelijk voor dat leerlingen meer gemotiveerd worden. Ook Reeve (1989) en Pekrun (2006) geven in hun onderzoeken aan dat plezier bijdraagt aan de intrinsieke motivatie.

De tien items uit de subschaal *Enjoyment* gaan over het genieten van de rekenles, de uitdaging van het oplossen van nieuwe problemen, het comfortniveau van het deelnemen aan discussies en het gevoel van geluk in de rekenles (Majeed et al., 2013). Voorbeelditems betreffen: *'I like to solve new problems in mathematics'* en *'I really like mathematics'*. De volledige subschaal is opgenomen in de bijlage. De subschaal heeft volgens Tapia en Marsh (2004) een hoge interne consistentie, met een Cronbach's alpha ( $\alpha$ ) van .89. Daarnaast werd test-hertest betrouwbaarheid over een periode van vier maanden vastgesteld met een Pearson correlatie van .84 (Tapia & Marsh, 2004).

Omdat de subschaal *Enjoyment* uit een Engelstalige vragenlijst komt, is deze vertaald volgens de richtlijnen van Beaton et al. (2000): allereerst is er door twee *native speakers* een vertaling gemaakt naar de doeltaal (Nederlands). Deze twee vertalers hebben hun vertaling gesynthetiseerd, zodat er één gemeenschappelijke vertaling ontstond. Een andere vertaler heeft vervolgens de vragenlijst terugvertaald in de oorspronkelijke taal (Engels). Vervolgens is de vertaling bekeken door de drie *native speakers* samen en is er consensus bereikt over eventuele discrepanties. Tot slot is de vragenlijst in de praktijk getest bij 30 leerlingen van dezelfde leeftijd als die in het onderzoek zouden gaan deelnemen. Het doel hiervan was om te kijken of de vragen duidelijk waren voor de leerlingen. Alle leerlingen konden de vragen beantwoorden en gaven aan de vraagstellingen te begrijpen. Daarom zijn er na deze test geen aanpassingen meer verricht aan de items uit de vragenlijst. Na deze stappen is er volgens Beaton et al. (2000) een valide vertaling tot stand gekomen en kan de vertaalde vragenlijst gebruikt worden.

De items uit de subschaal *Enjoyment* werden beantwoord op een 5-punts Likertschaal, met de volgende antwoordmogelijkheden: (1) helemaal niet mee eens, (2) niet mee eens, (3) neutraal, (4) mee eens en (5) helemaal mee eens. De score wordt berekend door de scores op de vragen bij elkaar op te tellen (Tapia & Marsh, 2002, 2004), waarbij een hogere score betekent dat de leerling meer plezier heeft in rekenen en daardoor meer gemotiveerd is (Deci & Ryan, 1985; Pekrun, 2006; Reeve, 1989).

De vragenlijst is als pretest en als posttest afgenomen. De eerste keer is de vragenlijst afgenomen om na te gaan of de groepen bij aanvang van het onderzoek significant van elkaar verschilden in de motivatie voor het vak rekenen. De tweede keer is de vragenlijst afgenomen na de uitvoerende fase van het onderzoek. De data van de tweede afname (posttest) is in de statistische analyses meegenomen.

### 2.2.3 *Ontwerp van de Interventie*

In beide condities zijn formatieve toetsen ingezet. De formatieve toetsing werd in zowel de experimentele- als de controlegroep middels het SRS *Socratic* vormgegeven. *Socratic* is via een WiFi-verbinding op bijvoorbeeld een laptop of tablet toegankelijk voor leerlingen (Dervan, 2014). De leerlingen van beide scholen werkten al regelmatig met SRS-systemen en waren dus al gewend aan het gebruik van een SRS om vragen te beantwoorden. De formatieve toetsing werd ingezet tijdens de herhalingslessen van de rekenmethode *Snappet*, die gebruikt werd op de deelnemende scholen. De opgestelde meerkeuzevragen sloten aan bij de leerstof die in die week werd behandeld. Hierdoor misten de leerlingen geen instructie. Het probleem dat door Crouch et al. (2007) benoemd wordt, namelijk dat er door de meerkeuzevragen bij formatieve toetsing minder tijd overblijft voor het geven van instructies over de leerstof, werd hiermee ondervangen. De formatieve toetsing bestond uit acht meerkeuzevragen. Beide condities kregen dezelfde vragen. De interventie nam ongeveer 45 minuten van de les in beslag.

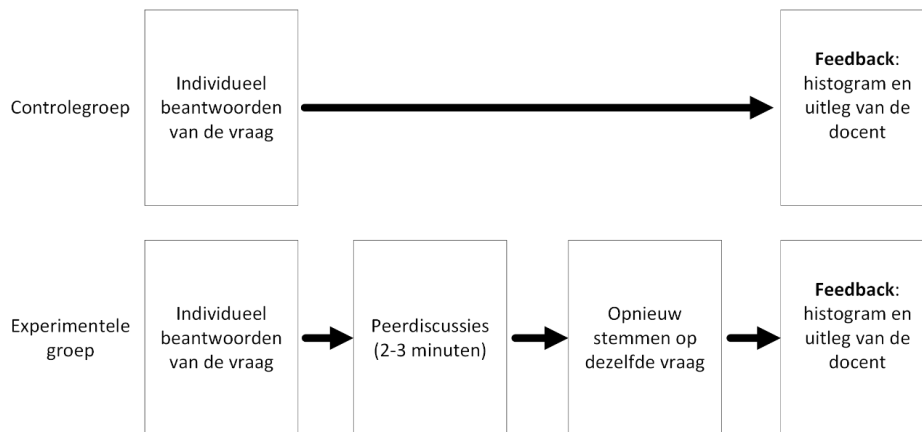
In de controlegroep gaf de leerkracht, nadat de leerlingen een vraag individueel hadden beantwoord, uitleg over de juiste redenering en beantwoordde eventuele vragen van leerlingen. In de experimentele groep werden, voordat de leerkracht feedback gaf, peerdiscussies gevoerd (de interventie). Hierbij werd gebruikgemaakt van de stappen van PI (Mazur, 1997). Zie voor het ontwerp van de interventie Figuur 2.

De formatieve toetsing vond in beide condities plaats in het laatste drie kwartier van de rekenles op woensdag- en vrijdagochtend (9:15-10:00 uur). De uitvoeringsfase bestond uit zeven weken, omdat in vergelijkbare onderzoeken met dit aantal weken positieve effecten zijn gevonden op leerprestaties (bijv. Austin, 2008; Shamir et al., 2006). Ook meta-analyses van Leung (2014) en Alegre-Ansuategui et al. (2017) geven aan dat met korte periodes waarin

peerdiscussies worden ingezet al effectieve resultaten kunnen worden behaald. Een periode van zeven weken was passend bij de toetskalender van de deelnemende scholen.

## Figuur 2

### *Beschrijving van het Ontwerp*



### **2.2.4 Meerkeuzevragen (ConceptTests)**

Tijdens de interventie werd er formatief getoetst. Hiervoor werden door de onderzoeker meerkeuzevragen opgesteld, ook wel *ConceptTests* genoemd. *ConceptTests* zijn meerkeuzevragen die gericht zijn op belangrijke concepten uit de les (McConnell et al., 2006) en gemeenschappelijke moeilijkheden omtrent het materiaal bloot kunnen leggen (Crouch & Mazur, 2001).

Bij het ontwikkelen en/of kiezen van *ConceptTests* wordt gebruik gemaakt van de taxonomie van Bloom (Bloom & Krathwohl, 1956). Dit raamwerk kan volgens Schell en Butler (2018) helpen bij de analyse van de cognitieve processen die een bepaalde vraag met zich meebrengt. Om ervoor te zorgen dat leerlingen tijdens dit onderzoek oefenen met vragen van eenzelfde niveau als bij de posttest, is vóór het zoeken van *ConceptTests* voor de formatieve toetsen eerst gekeken naar het niveau van de vragen uit de posttest (Cito B8 rekenen). De vragen uit de posttest zijn, afzonderlijk van elkaar, door de onderzoeker en door de rekencoördinator van de school ingedeeld in de niveaus uit de taxonomie van Bloom (Bloom & Krathwohl, 1956). Daarna zijn de antwoorden vergeleken en besproken. Zo is er

consensus bereikt over alle vragen. In Tabel 2 worden de resultaten weergegeven. Daarna is er in oude Cito toetsen van rekenen gezocht naar opdrachten die als *ConcepTest* gebruikt konden worden voor de formatieve toetsen. Hierbij is geprobeerd dezelfde procentuele verdeling van de niveaus uit de taxonomie van Bloom (Bloom & Krathwohl, 1956) aan te houden als bij de posttest. De uiteindelijke procentuele verdeling van vragen die gevonden zijn voor de formatieve toetsen is te zien in Tabel 2. Daarnaast is in Figuur 3 is een voorbeeld opgenomen van een lesdoel en een passende *ConcepTest* uit een oude Citotoets.

**Tabel 2**

*Verdeling Vragen in Niveaus van de Taxonomie van Bloom (Bloom & Krathwohl, 1956).*

		<b>B8- toets (Posttest)</b>	<b>Formatieve toetsen</b>
<i>Lage orde denkvaardigheden</i>	<i>Begrijpen</i>	4.1%	5.4%
	<i>Toepassen</i>	80.2%	82.1%
<i>Hoge orde denkvaardigheden</i>	<i>Analyseren</i>	16.7%	12.5%

De meeste vragen uit de oude rekentoetsen van Cito zijn volgens de onderzoeker en de rekencoördinator in te delen in de lagere niveaus uit de taxonomie van Bloom. Volgens Knight et al. (2013) heeft dit echter geen invloed op de kwaliteit van de peerdiscussies. Het cognitieve niveau van een vraag hangt namelijk niet noodzakelijkerwijs samen met hoe eenvoudig of moeilijk de leerlingen de vraag ervaren (Lemons & Lemons, 2013).

### Figuur 3

Voorbeeld Lesdoel en bijpassende ConceptTest groep 8



Noot: Links een lesdoel uit *Snappet* voor groep 8. Rechts een *conceptTest* die past bij het concept verhoudingen.

### 2.3 Procedure

De betreffende scholen zijn in juni 2023 benaderd om mee te doen aan het onderzoek. De scholen werden goed geïnformeerd en vervolgens werd er om toestemming gevraagd (zie Figuur 4). Na goedkeuring van het bestuur en de scholen, werden ouders/verzorgers middels een informatiebrief geïnformeerd over het doel van het onderzoek, de werkwijze en wat er van het kind werd verwacht (zie Figuur 4). Bij deze informatiebrief zat ook een toestemmingsformulier waarin ouders/verzorgers werden bevraagd op toestemming voor deelname aan het onderzoek. Er werd toestemming gevraagd voor het gebruik van de rekenresultaten van de Cito E7 en B8 en voor het opvragen van controlevariabelen uit het leerlingvolgsysteem, zoals geslacht en leeftijd. Ouders werden ervan verzekerd dat de gegevens vertrouwelijk behandeld werden en dat hun kind anoniem zou blijven in de rapportage van de uitkomsten. Er werd duidelijk vermeld dat leerlingen en ouders te allen tijde, zonder opgaaf van redenen, konden afzien van deelname aan het onderzoek. Leerlingen van ouders die geen toestemming verlenen voor deelname, maakten wel de Cito E7 en B8 toetsen, maar werden niet meegenomen in de data. De vragenlijst werd niet afgenomen bij deze leerlingen. Deze leerlingen namen wel deel aan de formatieve toetsing, omdat deze deel uit ging maken van de reguliere rekenles.

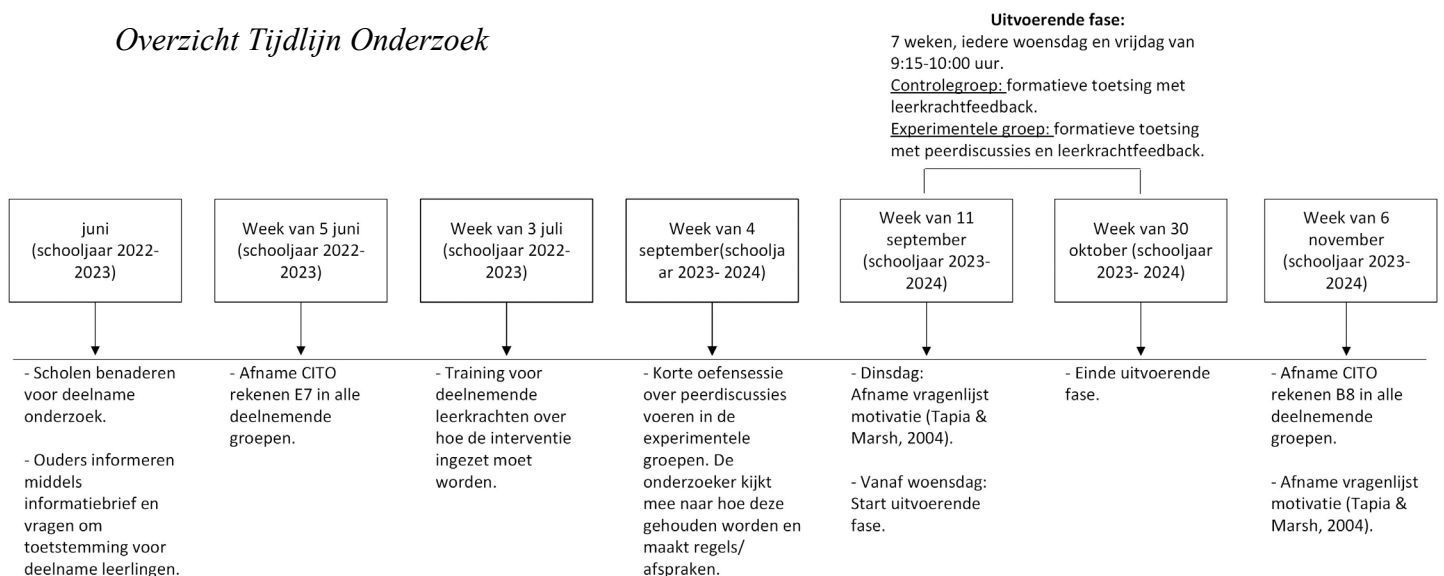
In de week van 5 juni 2023 werd de Cito E7 van rekenen afgenomen. Deze toets diende in dit onderzoek als startmeting om te kijken of de rekenresultaten van de twee condities voor de start van het onderzoek gelijkwaardig waren (zie Figuur 4). De leerkrachten die deel hebben genomen kregen eind schooljaar 2022-2023 een training over hoe de interventie moest worden vormgegeven. Dit betrof twee momenten: één training met de leerkrachten van de leerlingen uit de controlegroep, waarin de formatieve toetsing werd uitgelegd en één training met de leerkrachten die naast de formatieve toetsing de peerdiscussies gingen inzetten (zie Figuur 4). Deze leerkrachten kregen in de training uitleg over hoe formatieve toetsing werkte en hoe ze de peerdiscussies het best konden begeleiden. De rest van het onderzoek vond plaats in schooljaar 2023-2024.

In de derde week van het nieuwe schooljaar (week van 11 september) werd de vertaalde subschaal *Enjoyment* uit de ATMI-vragenlijst (Tapia & Marsh, 2004) afgenomen via *Limesurvey* (zie Figuur 4). Deze vragenlijst werd klassikaal afgenomen door de eigen leerkracht. De leerkrachten kregen een korte schriftelijke instructie over de afname van de vragenlijst. Leerlingen die de vragenlijst niet in hoefden te vullen, omdat ze geen toestemming hadden gekregen om deel te nemen aan het onderzoek, kregen een andere taak van de leerkracht. Verder hielden de leerlingen en de leerkrachten uit de experimentele conditie in de tweede week van het nieuwe schooljaar, onder leiding van de onderzoeker, een korte oefensessie ( $\pm 30$  minuten). In deze oefensessie keek de onderzoeker of de leerlingen wel beiden hun denkstappen aan elkaar verwoordden en ze niet na de redentatie van één leerling al stopten of alleen maar hun gekozen antwoord aan elkaar gaven zonder de stappen om tot het antwoord te komen te bespreken. Na deze oefensessie werden regels/afspraken gemaakt over hoe de peerdiscussies gevoerd diende te worden en wat er van de leerlingen verwacht werd.

De uitvoerende fase duurde vanaf van de derde week van schooljaar 2023-2024 (week van 11 september 2023) tot aan de week voor de *posttest* (week van 30 oktober 2023). In de week van 6 november vond de afname van de Cito rekenen B8 plaats (posttest) en werd nogmaals de vragenlijst afgenomen om de motivatie rondom het vak rekenen te meten. Een overzicht van de tijdlijn voor dit onderzoek is schematisch weergegeven in Figuur 4.

**Figuur 4**

*Overzicht Tijdlijn Onderzoek*



## 2.4 Data-Analyse

De gegevens uit de motivatievragenlijst zijn vanuit *Limesurvey* geëxporteerd naar Microsoft Excel. In Excel is de tekstuele data omgezet naar kwantitatieve data en zijn negatief geformuleerde vragen andersom gecodeerd. Vervolgens is deze data geïmporteerd naar *IBM Statistics (SPSS)*, versie 29. De resterende data werd handmatig ingevoerd in SPSS. Alle data werd gecontroleerd op (invoer)fouten en *missing values*. Daarna werden de waardes van de variabelen gestandaardiseerd, zodat de verschillende variabelen met elkaar te vergelijken waren (Creswell & Guetterman, 2021). Tot slot werd de betrouwbaarheid van de motivatievragenlijst gemeten middels Cronbach's Alpha. In dit onderzoek zijn de covariaten



leeftijd en geslacht meegenomen. Er is besloten de rekenscore van groep 7 en de pretestscore van motivatie niet als covariaten mee te nemen in de statistische analyses, omdat de correlatie tussen de rekenscore groep 7 en de rekenscore uit groep 8 en tussen de pretest en posttest motivatie boven de .80 was.

Voor de analyses is gebruik gemaakt van meerdere (multiple) lineaire regressieanalyses, aangezien er meerdere onafhankelijke variabelen (o.a. covariaten) zijn gebruikt om het effect op de afhankelijke variabele te meten (Nimon & Oswald, 2013). Om de hypothesen te toetsen middels de multiple regressieanalyses, werd gebruik gemaakt van de *Casual Steps* methode van Baron en Kenny (1986) Deze methode bestaat uit 4 stappen:

De eerste stap van de *Casual Steps* methode van Baron en Kenny (1986), waarin onderzocht wordt of er een significante relatie bestaat tussen de onafhankelijke variabele en de afhankelijke variabele, werd gebruikt om antwoord te geven op de eerste onderzoeksvraag: ‘Wat is het effect van peerdiscussies bij formatieve toetsing op de rekenresultaten?’. Hiervoor is een regressieanalyse uitgevoerd, waarbij het voeren van peerdiscussies als onafhankelijke variabele werd meegenomen en de rekenresultaten als afhankelijke variabele (pad a in Figuur 5). In een tweede run van deze regressieanalyse werden verklarende variabelen (geslacht en leeftijd) toegevoegd als onafhankelijke variabelen om het effect van de interventie nauwkeuriger te schatten.

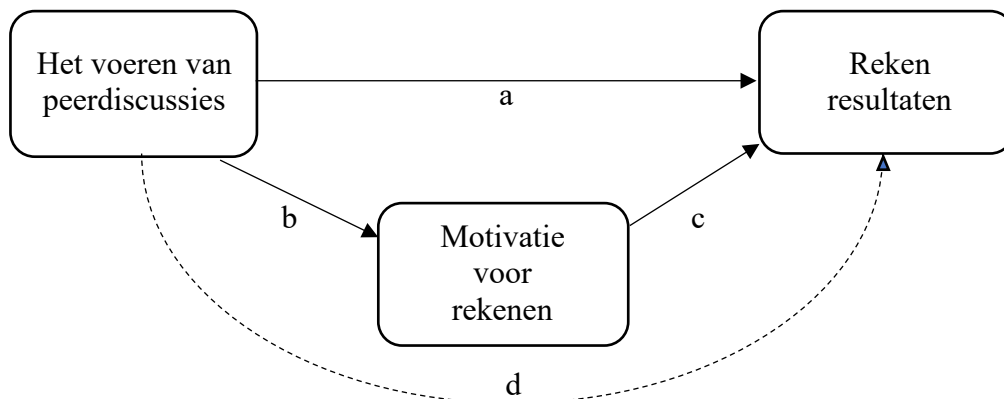
De tweede stap uit de *Casual Steps* methode (Baron & Kenny, 1986), waarin wordt onderzocht of er een significante relatie is tussen de onafhankelijke variabele en de mediator, werd gebruikt om de tweede deelvraag ‘In hoeverre leiden peerdiscussies tijdens formatieve toetsing tot een hogere motivatie voor rekenen?’ te beantwoorden. Bij deze regressieanalyse werd het voeren van peerdiscussies opgenomen als onafhankelijke variabele en werd motivatie opgenomen als afhankelijke variabele (pad b in Figuur 5). Ook deze

regressieanalyse werd een tweede keer uitgevoerd met verklarende variabelen leeftijd en geslacht.

Om de derde onderzoeksvraag ‘In welke mate beïnvloedt motivatie de relatie tussen het voeren van peerdiscussies tijdens formatieve toetsing en de rekenresultaten?’ te beantwoorden en te testen of er sprake was van een mediërend effect, werden de resterende stappen uit de *Casual Steps* methode uitgevoerd (Baron & Kenny, 1986). In de derde stap wordt er onderzocht of er een significante relatie is tussen de mediator en de afhankelijke variabele, nadat er gecontroleerd is voor de invloed van de onafhankelijke variabele. Voor deze stap werd een regressieanalyse uitgevoerd waarbij rekenresultaten als afhankelijke variabele werd opgenomen en de motivatie als onafhankelijke variabele (pad c in Figuur 5). In deze analyse zijn de covariaten de eerste keer meteen meegenomen. Tot slot is er voor stap vier, de mediatieanalyse, een regressieanalyse uitgevoerd waarbij het voeren van peerdiscussies en de motivatie als onafhankelijke variabelen zijn opgenomen en de rekenresultaten als afhankelijke variabele. Bij deze analyse zijn de covariaten meteen meegenomen in de eerste run. Omdat Baron en Kenny (1986) suggereren om bijkomend een Sobel-test (Sobel, 1982) uit te voeren om de significantie van de mediatie na te gaan, is deze ook uitgevoerd.

### Figuur 5

*Grafische Weergave Data-Analyse*



## 2.5 Beschrijvende Statistiek van de Conditie

Om na te gaan of de experimentele groep wat betreft geslacht, leeftijd, motivatie en rekenresultaten voor de start van het onderzoek vergelijkbaar was met de controlegroep, zijn er onafhankelijke t-testen uitgevoerd. Middels Levene's Test is gekeken of er voor het experiment significante verschillen aanwezig waren tussen beide condities. In Tabel 3 zijn de resultaten van de twee condities naast elkaar weergegeven.

Het geslacht bleek bij de start van dit onderzoek significant te verschillen tussen de condities ( $p = .04$ ). In de experimentele groep zaten dus significant meer jongens dan in de controlegroep. Daarnaast bleek leeftijd voor aanvang van het onderzoek ook significant verschillend tussen de controlegroep en de experimentele groep ( $p = .09$ ). In de experimentele groep zijn de leerlingen significant ouder dan in de controlegroep.

**Tabel 3**

*Vergelijking van de Conditie Voor de Start van het Onderzoek*

	<i>Conditie</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Levene's Test (sig.)</i>
Leeftijd	Controlegroep	40	10.88	.34	.09*
	Experimenteel	36	10.89	.52	
Geslacht	Controlegroep	40	.53	.50	.04**
	Experimenteel	36	.67	.50	
Pretest motivatie	Controlegroep	40	31.73	5.11	.11
	Experimenteel	36	32.17	7.19	
Rekenscore groep 7	Controlegroep	40	273.48	22.99	.83
	Experimenteel	36	279.86	22.98	

*Noot.* \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .001$

## 2.6 Interne Consistentie Meetinstrument

In versie 29 van *IBM SPSS Statistics* werd aan het begin van dit onderzoek gekeken naar de interne consistentie van de vertaalde *Enjoyment*-schaal uit de ATMI (Tapia & Marsh, 2004) die werd gebruikt om de motivatie voor het vak rekenen te meten. Het derde item uit de vragenlijst ('Rekenen is suf en saai om te doen') werd andersom gecodeerd, aangezien dit een negatief gestelde vraag betrof. Vervolgens werd de berekening uitgevoerd met data uit de pretest ( $N = 76$ ). De Cronbach's Alpha ( $\alpha$ ) voor de tien vertaalde items uit de *Enjoyment*-schaal van de ATMI (Tapia & Marsh, 2004) was .81, wat volgens Bland en Altman (1997) als goed wordt beschouwd voor een onderzoek waarin groepen met elkaar worden vergeleken.

## 3. Resultaten

### 3.1 Statistische assumpties

Om de hypothesen uit dit onderzoek te toetsen zijn er (multiple) regressieanalyses uitgevoerd, zoals voorgesteld door Baron en Kenny (1986). Aangezien een mediatieanalyse op lineaire regressieanalyses terugvalt, moest met bepaalde assumpties rekening gehouden worden (Abu-Bader & Jones, 2021): (a) de residuen moeten ongeveer normaal verdeeld zijn (normaliteit), (b) er moet sprake zijn van homoscedasticiteit en (c) er mag geen sprake zijn van multicollineariteit.

De P-P plot liet zien dat de punten van de verschillende variabelen ongeveer op de lijn lagen en in de histogrammen volgden de variabelen ongeveer de normaalverdeling. De assumptie dat de data normaal verdeeld moest zijn, werd dus niet geschonden. Daarnaast werd in de scatterplots gekeken of de data verdeeld was. Dit was het geval; er was dus sprake van homoscedasticiteit. Tot slot bleek dat er bij de variabelen die zijn meegenomen in de statistische analyses geen correlaties hoger dan .80 waren en bleken alle VIF-waardes onder de 4 te zijn. Hierdoor is er geen sprake van multicollineariteit. In dit onderzoek is aan alle vereiste assumpties voldaan.

### 3.2 Resultaten Statistische Analyses

De multiple regressieanalyses die uitgevoerd werden via de *Casual Steps* methode van Baron en Kenny (1986) hielden verband met de drie onderzoeksvragen in dit onderzoek. Om het risico op een type I fout te verminderen, werd voor alle schattingen in de modellen de Bonferroni correctie gebruikt (Abdi, 2007), zoals ook gebruikt in Molin et al. (2019). De resultaten worden gerapporteerd in Tabel 4.

**Tabel 4**

*Resultaten Analyses Weergegeven in Gestandaardiseerde Regressiecoëfficiënt ( $\beta$ )*

Afhankelijke variabele:	Pad a		Pad b		Pad c	Pad d
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	<b>Rekenresultaten</b>	<b>Rekenresultaten</b>	<b>Rekenmotivatie</b>	<b>Rekenmotivatie</b>	<b>Rekenresultaten</b>	<b>Rekenresultaten</b>
<b>Conditie<sup>a</sup></b>	.25*	.23*	.47***	.44***	-	.003
<b>Motivatie</b>	-	-	-	-	.51***	.51***
Leeftijd	-	-.17	-	-.21*	-.06	-.06
Geslacht <sup>b</sup>	-	-.14	-	.23**	.03	.03

*Noot.* \* $p < .05$ , \*\* $p < .025$ , \*\*\* $p < .01$ , <sup>a</sup>0 = controlegroep, 1 = experimentele groep, <sup>b</sup>0 = meisje, 1 = jongen

In de eerste stap uit de *Casual Steps* methode (Baron & Kenny, 1986) wordt gekeken naar pad a (Figuur 5). De eerste twee modellen uit Tabel 4 presenteren de resultaten van de (multiple) regressieanalyses met ‘rekenresultaten’ als uitkomstmaat. In Model 1 is alleen de conditie als onafhankelijke variabelen meegenomen. De resultaten van Model 1 tonen een positief significant effect van het voeren van peerdiscussies op rekenresultaten ( $\beta = .25$ ,  $p < .05$ ), wat aangeeft dat peerdiscussies voeren tijdens formatieve toetsing leidt tot betere rekenresultaten. Vervolgens zijn voor Model 2 de covariaten leeftijd en geslacht toegevoegd. Model 2 laat zien dat het effect van het voeren van peerdiscussies op de rekenresultaten significant blijft ( $\beta = .23$ ,  $p < .05$ ).

In de tweede stap uit de *Casual Steps* methode (Baron & Kenny, 1986) wordt gekeken naar pad b (Figuur 5). Model 3 betreft de regressieanalyse waarin ‘rekenmotivatie’ de

uitkomstmaat is en waarbij alleen de conditie als onafhankelijke variabele is meegenomen.

Model 3 toont aan dat het voeren van peerdiscussies een positief significant effect heeft op de motivatie voor het vak rekenen ( $\beta = .47, p < .01$ ): leerlingen die peerdiscussies hebben gevoerd zijn meer gemotiveerd voor rekenen dan de leerlingen die geen peerdiscussies hebben gevoerd. In Model 4 zijn de covariaten leeftijd en geslacht toegevoegd om het effect van het voeren van peerdiscussies nauwkeuriger te schatten. De resultaten uit Model 4 laten zien dat het effect van de interventie op de rekenmotivatie significant blijft ( $\beta = .44, p < .01$ ). Ook leeftijd is een significante voorspeller voor de rekenmotivatie ( $\beta = -.21, p < .05$ ): leerlingen die ouder zijn, zijn minder gemotiveerd voor rekenen. Daarnaast is geslacht in Model 4 een significante voorspeller voor de rekenmotivatie ( $\beta = .23, p < .025$ ). Dit betekent dat in dit onderzoek jongens meer gemotiveerd zijn voor rekenen dan meisjes.

Model 5 uit Tabel 4 hoort bij de derde stap uit de *Casual Steps* methode (Baron & Kenny, 1986). Model 5 toont de regressieanalyse, waarin 'rekenresultaten' de uitkomstmaat is en waarbij de motivatie voor rekenen als onafhankelijke variabele is meegenomen (pad c in Figuur 5). In deze analyse zijn alle covariaten meegenomen. De resultaten van Model 5 tonen een significant effect van de motivatie voor rekenen op de rekenresultaten ( $\beta = .51, p < .01$ ), wat aangeeft dat een hogere motivatie voor rekenen leidt tot betere rekenresultaten.

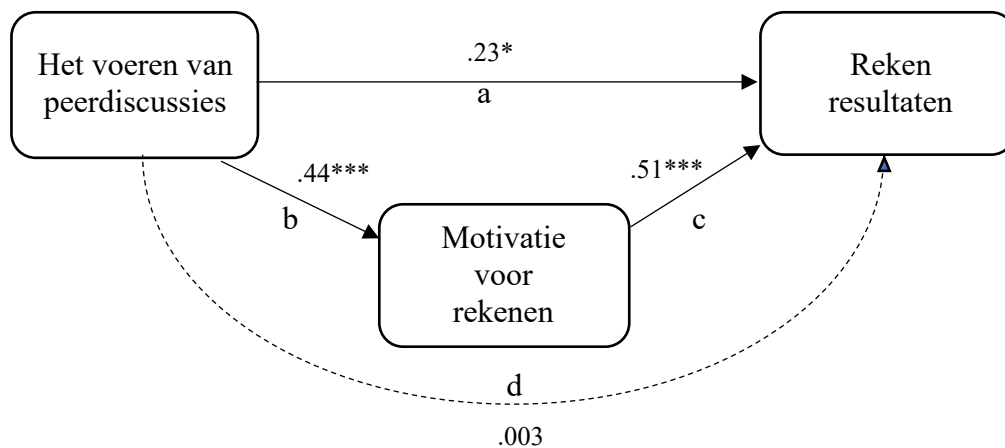
Omdat alle drie de stappen uit de *Casual Steps* methode van Baron en Kenny (1986) significant zijn, is de mediatieanalyse uitgevoerd. Bij deze analyse is het rekenresultaat de uitkomstmaat en werden de conditie en de motivatie als onafhankelijke variabelen ingevoerd, aangevuld met de covariaten leeftijd en geslacht (pad d in Figuur 5). De resultaten van Model 6 tonen aan dat er geen significant effect meer is van het voeren van peerdiscussies op de rekenresultaten ( $\beta = .003, p = .98$ ). De motivatie voor rekenen blijft in Model 6 wel een significante voorspeller van de rekenresultaten ( $\beta = .51, p < .01$ ), wat betekent dat de motivatie voor rekenen de relatie tussen het voeren van peerdiscussies en de rekenresultaten

geheel medieert. In Figuur 6 worden de resultaten in het mediatiemodel schematisch weergegeven.

Als aanvulling op de methode van Baron en Kenny (1986) is een Sobel-test (Sobel, 1982) uitgevoerd om na te gaan of motivatie voor rekenen inderdaad een mediator is in de relatie tussen het voeren van peerdiscussies en rekenresultaten. De Sobel-test (Sobel, 1982) toont ook aan dat deze mediatie significant is ( $p < .001$  waarbij  $t_a = 4.5624$  en  $t_b = 4.7880$ ).

### ***Figuur 6***

#### *Resultaten Data-Analyse*



## **4. Discussie**

Het doel van dit onderzoek was nagaan of het voeren van peerdiscussies bij formatieve toetsing ervoor kon zorgen dat de rekenresultaten en de rekenmotivatie van bovenbouwleerlingen uit het primair onderwijs verbeterden. Een quasi-experiment onder 76 basisschoolleerlingen uit groep 8 werd uitgevoerd over een periode van zeven weken. De veronderstelling was dat het voeren van peerdiscussies een positief significant effect zou hebben op de rekenresultaten en op de motivatie voor het vak rekenen. Er werd ook onderzocht of de motivatie voor rekenen een mediërend effect had op de relatie tussen het

voeren van peerdiscussies en de rekenresultaten. Dit onderzoek heeft deze drie onderzoeksvragen beantwoord, hieronder per paragraaf weergegeven.

## **4.1 Antwoord Onderzoeksvragen**

### ***4.1.1 De Relatie tussen het Voeren van Peerdiscussies en de Rekenresultaten***

De eerste onderzoeksvraag was: ‘Wat is het effect van peerdiscussies bij formatieve toetsing op de rekenresultaten?’ Uit de resultaten blijkt dat de experimentele groep, waar peerdiscussies zijn gevoerd tijdens de formatieve toetsing, significant hogere rekenresultaten behaalt dan de controlegroep, waarin de formatieve toetsing gegeven is zonder het voeren van peerdiscussies. Er wordt een significante effectgrootte gevonden van  $d = .23$ . Deze Cohen’s  $d$  wordt gezien als een klein positief effect (Cohen, 1969). De alternatieve hypothese (H1) wordt dus aangenomen: het voeren van peerdiscussies heeft een significant positief effect op de rekenresultaten. De bevinding dat het voeren van peerdiscussies zorgt voor betere leerprestaties komt overeen met de resultaten uit de meta-analyse van Balta et al. (2017). Deze meta-analyse is gebaseerd op 29 studies ( $N = 4738$ ). Er werd geconcludeerd dat PI effectiever is dan lesgeven via traditionele onderwijsmethoden. Ook wanneer we specifiek kijken naar het vak rekenen komt deze bevinding overeen met de literatuur: Akay (2011), Allison (2012) en Kaymak et al. (2020) geven allen aan dat het inzetten van peerdiscussies middels PI (Mazur, 1997) er voor zorgt dat leerlingen binnen het primair onderwijs significant hoger scoren op de rekentoetsen dan wanneer er geen gebruik wordt gemaakt van een pedagogiek zoals PI (Mazur, 1997).

Het positieve effect van het voeren van peerdiscussies op de rekenresultaten kan verklaard worden door het feit dat leerlingen concepten aan elkaar uit kunnen leggen in eenvoudige taal, waardoor ze het beter te begrijpen (Nicol en Boyle, 2003; Perez et al., 2010). Leerlingen reageren op elkaars uitleg, vullen elkaar aan en kunnen zo samen kennis opbouwen die niet beschikbaar is voordat ze begonnen aan de peerdiscussie (Hogan et al.,



2016). Volgens Barth-Cohen et al. (2016) en Vickrey et al. (2015) zorgt dit proces voor versterking van het begrip, waardoor de leerlingen dus beter leren rekenen door toedoen van de gesprekken die ze voeren tijdens de peerdiscussies.

#### ***4.1.2 De Relatie tussen het Voeren van Peerdiscussies en de Motivatie voor Rekenen***

De tweede onderzoeksvraag was: ‘In hoeverre leiden peerdiscussies tijdens formatieve toetsing tot een hogere motivatie voor rekenen?’ Uit de resultaten blijkt dat de experimentele groep, waarin peerdiscussies werden gevoerd tijdens de formatieve toetsing, significant hoger scoort op motivatie voor het vak rekenen dan de controlegroep, waarin de formatieve toetsing gegeven werd zonder het voeren van peerdiscussies. Er wordt een significante effectgrootte gevonden van  $d = .44$ . Deze Cohen’s  $d$  nadert een medium effect (Cohen, 1969). De hypothese (H2) wordt dus ondersteund: het voeren van peerdiscussies heeft een significant positief effect op de motivatie voor het vak rekenen.

Dit resultaat is in lijn met vergelijkbare onderzoeken die uitgevoerd zijn in de bovenbouw van het primair onderwijs (Akay, 2011; Allison 2012). Allison (2012) rapporteert een significante uitkomst van het voeren van peerdiscussies op motivatie en ook Akay (2011) geeft aan dat het hoofdeffect van de conditie op motivatie statistisch significant is. De effectgrootte uit dit onderzoek is vergelijkbaar is met de effectgrootte van .52 uit het onderzoek van Akay (2011). Mogelijke verklaringen voor de significant hogere motivatie in de experimentele groep zijn in de literatuur te vinden. De interacties tijdens het voeren van peerdiscussies dragen bij aan het vervullen van de psychologische basisbehoeften zoals autonomie, competentie en verbondenheid (Deci & Ryan, 1985). Uit het onderzoek van Cronhjort (2013) blijkt daarnaast dat peerdiscussies zorgde voor een hogere motivatie voor rekenen, omdat leerlingen actiever bezig zijn dan bij een rekenles waarin er enkel sprake is van kennisoverdracht van de leraar richting de leerlingen. Volgens Abdelkarim en Abuiyada (2016) is het de interactie met leeftijdsgenoten die positief van invloed is op de motivatie.

De analyses die horen bij de tweede deelvraag tonen daarnaast aan dat leeftijd een significante voorspeller is voor de rekenmotivatie: de rekenmotivatie van leerlingen neemt af naarmate leerlingen ouder worden. Dit komt volgens Gottfried et al. (2007) en Fredricks en Eccles (2002) doordat het competentiegevoel en de waarde die leerlingen aan het vak rekenen/wiskunde hangen in de loop van de tijd afnemen. Dit onderzoek levert wat betreft de relatie tussen leeftijd en de motivatie voor het vak rekenen geen verrassende resultaten op.

Ook geslacht is een significante voorspeller van motivatie: in dit onderzoek zijn jongens significant meer gemotiveerd voor rekenen dan meisjes. Dit significante effect wordt in andere studies ook gevonden. Zo geven Ganley en Lubienski (2016) en Martínez et al. (2020) in hun onderzoeken ook aan dat binnen het primair onderwijs jongens een grotere intrinsieke motivatie voor rekenen rapporteren dan meisjes. De motivatie van jongens zou hoger zijn omdat zij meer vertrouwen hebben in hun capaciteiten en meer gemotiveerd zijn om zich te onderscheiden. Meisjes zouden een hogere mate van angst hebben tegenover rekenen/wiskunde, wat de motivatie negatief beïnvloedt (Ganley & Lubienski, 2016).

#### ***4.1.3 De Mediërende Rol van Motivatie***

De derde onderzoeksvraag was ‘In welke mate beïnvloedt motivatie de relatie tussen het voeren van peerdiscussies tijdens formatieve toetsing en de rekenresultaten?’ Volgens de *Casual Steps* methode van Baron en Kenny (1986) is er in dit onderzoek sprake van volledige mediatie. Dit houdt in dat door het toevoegen van motivatie als verklarende variabele het directe effect van peerdiscussies op de rekenresultaten bijna verdwijnt. In eerste instantie lijkt de verbetering van de rekenresultaten in dit onderzoek voort te komen uit het voeren van peerdiscussies. Echter, wanneer motivatie als verklarende variabele wordt meegenomen, is te zien dat het effect van het voeren van peerdiscussies verdwijnt en dat motivatie de verklarende factor is voor de betere rekenresultaten.

Ook zijn de resultaten van de Sobel test (Sobel, 1982) significant, wat betekent dat motivatie de relatie tussen het voeren van peerdiscussies en de rekenresultaten in dit onderzoek geheel medieert: hypothese (H3) wordt dus aangenomen. In dit onderzoek verbetert voeren van peerdiscussies de motivatie voor het vak rekenen. De verhoogde motivatie voor rekenen zorgt voor betere rekenresultaten.

Omdat dit zover we weten het eerste onderzoek betreft waarin het effect van motivatie onderzocht is tussen peerdiscussies tijdens formatieve toetsing en leerprestaties kunnen deze resultaten niet getoetst worden aan de reeds bestaande literatuur. Wel vullen deze resultaten dus een leemte in de literatuur. Vervolgonderzoek met dezelfde opzet is nodig om te kijken of deze effecten vaker gevonden worden.

#### **4.2 Beperkingen van het Onderzoek en Toekomstig Onderzoek**

Ondanks dat dit onderzoek zorgvuldig is uitgevoerd, zijn er toch beperkingen te noemen. Omdat er gebruik is gemaakt van een quasi-experimenteel design, is er geen sprake geweest van randomisatie op leerlingniveau, wat van invloed kan zijn op de interne validiteit (Creswell & Guetterman, 2021). De resultaten dienen met enige voorzichtigheid geïnterpreteerd te worden.

Een tweede beperking is dat in dit onderzoek geen multilevel analyse is uitgevoerd om te corrigeren voor kenmerken van de groep waarin de leerlingen zaten. De regressieanalyses die zijn uitgevoerd hebben namelijk geen rekening gehouden met de verschillende levels: leerlingen die genest zijn in groepen en groepen die genest zijn in scholen. Met een multilevel analyse hadden variabelen kunnen worden opgenomen die op verschillende niveaus van de hiërarchie worden gemeten. Dit had kunnen resulteren in nauwkeurigere schattingen van regressiecoëfficiënten en statistische significantietests (Austin et al., 2001). Hierdoor was het wellicht mogelijk geweest om rijkere conclusies te trekken uit de gegevens die tijdens dit onderzoek zijn verzameld.

Dit onderzoek heeft veelbelovende resultaten gevonden: door het voeren van peerdiscussies zijn de rekenresultaten significant verbeterd en is de motivatie voor het vak rekenen significant verhoogd. Motivatie medieert de relatie tussen het voeren van peerdiscussies en de rekenresultaten volledig. Het is interessant om te onderzoeken of het voeren van peerdiscussies in leerjaren waar de leerlingen jonger zijn ook zorgt voor een hogere motivatie voor het vak rekenen en daardoor zorgt voor betere rekenresultaten. Als dit zo is, kan het voeren van peerdiscussies in de toekomst wellicht bijdragen aan het behalen van de ambitie die gesteld is door de commissie Meijerink (van der Lubbe, 2021).

#### **4.3 Maatschappelijke en Wetenschappelijke Significantie**

Kunnen rekenen blijft het hele leven belangrijk, omdat de meeste informatie wordt aangeboden en opgeslagen in numerieke vorm (Ruijssenaars et al., 2014). Basiskennis en -vaardigheden hebben van rekenen zijn voorwaardelijk om volledig deel te kunnen nemen aan de huidige samenleving (van den Berg et al., 2017). Wanneer het rekenen niet vanzelf gaat, kan dit aanzienlijke beperkingen opleveren in het dagelijks functioneren (Ruijssenaars et al., 2014). Ook voldoen de rekenprestaties van leerlingen op het eind van de basisschoolperiode niet aan de gestelde ambities (van der Lubbe, 2021). Het leggen van een goede rekenbasis op vroege leeftijd is daarom cruciaal (Aubrey et al., 2006). Niet alleen de rekenresultaten dalen, ook de motivatie van leerlingen neemt af naarmate leerlingen ouder worden (bijv. Gnambs & Hanfstingl, 2016). De resultaten van dit onderzoek tonen aan dat het voeren van peerdiscussies de rekenresultaten significant verbetert en dat de motivatie van leerlingen significant verhoogd wordt wanneer leerlingen peerdiscussies voeren. Leerkrachten moeten gaan inzien dat het voeren van peerdiscussies geen verspilde tijd is, maar juist zorgt voor meer begrip van de leerstof. Om dit te bewerkstelligen, zal er op de pedagogische academie voor het basisonderwijs (PABO) meer aandacht besteed moeten worden aan wat het voeren van peerdiscussies inhoudt en wat effect hiervan is op de leerlingen. Op deze manier nemen

alle toekomstige leerkrachten gedurende de opleiding kennis van deze interventie en de bewezen effectiviteit hiervan. Hierdoor hopen we dat toekomstige leerkrachten het voeren van peerdiscussies sneller zullen implementeren in hun onderwijs.

De bijdrage van deze studie aan de wetenschappelijke literatuur is tweeledig. Volgens Palm et al. (2017) zijn er op het niveau van het primair onderwijs te weinig studies uitgevoerd die formatieve toetsing in combinatie met peerdiscussies bij rekenen hebben onderzocht. De meeste onderzoeken waarin het voeren van peerdiscussies als interventie is ingezet bij rekenen, zijn uitgevoerd in het middelbaar onderwijs, het hoger onderwijs of in een universitaire setting. Deze studie toont, net als de studies van Akay (2011) en Allison (2012) aan dat het voeren van peerdiscussies significant positieve effecten heeft op de rekenresultaten en de motivatie voor het vak rekenen op het niveau van het primair onderwijs. De studie is daarmee een toevoeging aan de huidige kennisbasis.

Ten tweede is dit, zover bekend, de eerste studie die de mediërende rol van motivatie heeft onderzocht tussen het voeren van peerdiscussies en de rekenresultaten binnen het primair onderwijs. Daarmee is deze studie een toevoeging aan de bestaande wetenschappelijke kennisbasis rondom het voeren van peerdiscussies en de rol van motivatie.

#### **4.4 Conclusie**

In dit onderzoek is onderzocht of het voeren van peerdiscussies een positieve invloed heeft op de rekenresultaten en de motivatie voor rekenen van leerlingen uit groep 8 uit het primair onderwijs. Daarnaast is er gekeken of motivatie een mediator was tussen het voeren van peerdiscussies en de rekenresultaten. Er wordt een positief significante relatie gevonden tussen het voeren van peerdiscussies en de rekenresultaten. Het voeren van peerdiscussies bij rekenen zorgt in dit onderzoek dus voor aantoonbare voordelen voor het leerproces. Het bespreken van rekenkundige problemen met een medeleerling biedt de mogelijkheid om

diverse denkstrategieën te verkennen en te begrijpen (Bojinova & Oigara, 2011; Dawson et al., 2018; Nicol & Boyle, 2003).

Daarnaast wordt een positief significante relatie gevonden tussen het voeren van peerdiscussies en motivatie voor het vak rekenen. De interacties tijdens het voeren van peerdiscussies dragen bij aan het vervullen van de psychologische basisbehoeften zoals autonomie, competentie en verbondenheid (Deci & Ryan, 1985). Tevens zorgt het voeren van peerdiscussies voor een actievere houding tijdens de rekenlessen (Cronhjort, 2013). De combinatie van bovenstaande zorgt voor een hogere motivatie voor rekenen. Motivatie medieert de relatie tussen het voeren van peerdiscussies en de rekenresultaten zelfs helemaal.

Kortom, het integreren van peerdiscussies in het rekenonderwijs bevordert niet alleen de rekenvaardigheden van de leerlingen, maar heeft ook een aantoonbare positieve impact op de rekenmotivatie. Integratie van peerdiscussies in rekenlessen op het niveau van het primair onderwijs dient daarom te worden beschouwd als een essentiële strategie om zowel rekenvaardigheden als de motivatie voor rekenen te bevorderen.

### Referenties

- Abdelkarim, R., & Abuiyada, R. (2016). The Effect of Peer Teaching on Mathematics Academic Achievement of the Undergraduate Students in Oman. *International Education Studies*, 9(5), 124. <https://doi.org/10.5539/ies.v9n5p124>
- Abdi, H. (2007), Bonferroni and Sidak corrections for multiple comparisons, in N.J. Salkind (ed.): *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage
- Abu-Bader, S., & Jones, T. V. (2021). Statistical mediation analysis using the Sobel test and Hayes SPSS Process Macro. *International Journal of Quantitative and Qualitative Research Methods*, 9(1), 42-61.
- Akobeng, A. K. (2016). Understanding type I and type II errors, statistical power and sample size. *Acta Paediatrica*, 105(6), 605–609. <https://doi.org/10.1111/apa.13384>
- Alegre-Ansuategui, F. J., Miravet, L. M., & Maroto, A. D. (2017). Peer Tutoring and Academic Achievement in Mathematics: A Meta-Analysis. *Eurasia journal of mathematics, science and technology education*, 14(1).  
<https://doi.org/10.12973/ejmste/79805>
- Allison, T. H. (2012). *The impact of classroom performance system-based instruction with peer instruction upon student achievement and motivation in eighth grade math students*. Geraadpleegd op 27 februari 2023, van <https://core.ac.uk/download/pdf/58824403.pdf>
- Akay, G. (2011). The effect of peer instruction method on the 8th grade students' mathematics achievement in transformation geometry and attitudes towards mathematics. *Unpublished master thesis, METU, Ankara*.
- Amrai, K., Motlagh, S. E., Zalani, H. A., & Parhon, H. (2011). The relationship between academic motivation and academic achievement students. *Procedia - Social and Behavioral Sciences*, 15, 399–402. <https://doi.org/10.1016/j.sbspro.2011.03.111>

- Andersson, C., & Palm, T. (2018). Reasons for teachers' successful development of a formative assessment practice through professional development – a motivation perspective. *Assessment in Education: Principles, Policy & Practice*, 25(6), 576–597. <https://doi.org/10.1080/0969594x.2018.1430685>
- Aubrey, C., Godfrey, R. C., & Dahl, S. (2006). Early mathematics development and later achievement: Further evidence. *Mathematics Education Research Journal*, 18(1), 27–46. <https://doi.org/10.1007/bf03217428>
- Austin, J. (2008). *The Effects of Peer Tutoring on Fifth-Grade Students' Motivation and Learning in Math* [Master's theses]. State University of New York.
- Austin, P. C., Goel, V., & Van Walraven, C. (2001). An introduction to multilevel regression models. *Canadian Journal of Public Health*, 92(2), 150–154. <https://doi.org/10.1007/bf03404950>
- Balta, N., Michinov, N., Balyimez, S., & Ayaz, M. F. (2017). A meta-analysis of the effect of Peer Instruction on learning gain: Identification of informational and cultural moderators. *International Journal of Educational Research*, 86, 66–77. <https://doi.org/10.1016/j.ijer.2017.08.009>
- Baron, R. M., & Kenny, D. A. (1986). The Moderator–Mediator Variable Distinction in Social Psychological Research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bartsch, R. A., & Murphy, W. (2011). Examining the Effects of an Electronic Classroom Response System on Student Engagement and Performance. *Journal of Educational Computing Research*, 44(1), 25–33. <https://doi.org/10.2190/ec.44.1.b>
- Barth-Cohen, L. A., Smith, M. K., Capps, D. K., Lewin, J. D., Shemwell, J. T., & Stetzer, M. R. (2016). What are Middle School Students Talking About During Clicker



- Questions? Characterizing Small-Group Conversations Mediated by Classroom Response Systems. *Journal of Science Education and Technology*, 25(1), 50–61.  
<https://doi.org/10.1007/s10956-015-9576-2>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.  
<https://doi.org/10.1080/0969595980050102>
- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's Alpha. *BMJ*, 314(7080), 572. <https://doi.org/10.1136/bmj.314.7080.572>
- Bloom, B. S., & Krathwohl, D. R. (1956). Taxonomy of educational objectives; the classification of educational goals by a committee of college and university examiners. *Handbook I: Cognitive Domain*. New York, NY; Longmans, Green.
- Bojinova, E., & Oigara, J. (2011). Teaching and Learning with Clickers: Are Clickers Good for Students? *Interdisciplinary journal of e-skills and lifelong learning*, 7, 169–184.  
<https://doi.org/10.28945/1506>
- Caldwell, J. E. (2007). Clickers in the Large Classroom: Current Research and Best-Practice Tips. *CBE- Life Sciences Education*, 6(1), 9–20. <https://doi.org/10.1187/cbe.06-12-0205>
- Castillo, A. J., Durán, P. A., Fuller, E., Watson, C., Potvin, G., & Kramer, L. (2022). Student attitudes in an innovative active learning approach in calculus. *International Journal of Mathematical Education in Science and Technology*, 1–29.  
<https://doi.org/10.1080/0020739x.2021.2023771>

- Cauley, K. M., & McMillan, J. H. (2010). Formative Assessment Techniques to Support Student Motivation and Achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 83(1), 1–6. <https://doi.org/10.1080/00098650903267784>
- Chang, C. Y., Chien, Y. T., Lee, Y. H., & Li, T. Y. (2015). Examining the Effects of Displaying Clicker Voting Results on High School Students' Voting Behaviors, Discussion Processes, and Learning Outcomes. *EURASIA Journal of Mathematics, Science and Technology Education*, 11(5).  
<https://doi.org/10.12973/eurasia.2015.1414a>
- Cho, Y. M., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science*, 39(5), 629–643. <https://doi.org/10.1007/s11251-010-9146-1>
- Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York: Academic Press.
- Creswell, J. W., & Guetterman, T. C. (2021). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6e ed.). Pearson.
- Cronhjort, M., Filipsson, L. & Weurlander, M. (2013) Can peer instruction in calculus improve student learning? *Proceedings of the 9th International CDIO Conference*. Cambridge, Massachusetts: Massachusetts Institute of Technology and Harvard University School of Engineering and Applied Sciences, 1–9.
- Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. <https://doi.org/10.1119/1.1374249>
- Crouch, C., Watkins, J., Fagen, A., & Mazur, E. (2007). Peer Instruction: Engaging Students One-on-One, All at Once. In Redish, E., & Cooney, P. (Eds.), *Research-Based Reform of University Physics* (1). College Park: American Association of Physics Teachers.
- Geraadpleegd op 6 maart 2023, van

<https://www.compadre.org/Repository/document/ServeFile.cfm?ID=4990&DocID=24>

1

- Dawson, V., & Venville, G. (2010). Teaching Strategies for Developing Students' Argumentation Skills About Socioscientific Issues in High School Genetics. *Research in Science Education*, 40(2), 133–148. <https://doi.org/10.1007/s11165-008-9104-y>
- Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2018). What makes for effective feedback: staff and student perspectives. *Assessment & Evaluation in Higher Education*, 44(1), 25–36. <https://doi.org/10.1080/02602938.2018.1467877>
- Deci, E. L., & Ryan, R. M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. *Springer eBooks*. <https://doi.org/10.1007/978-1-4899-2271-7>
- Deci, E. L., & Ryan, R. M. (2004). *Handbook of Self-determination Research*. Amsterdam University Press.
- Dervan, P. (2014). Increasing in-class student engagement using Socrative (an online Student Response System). *AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education*, 6(3). Geraadpleegd op 8 april 2023 van: <http://eprints.teachingandlearning.ie/2182/>
- Duncan, D., & Mazur, E. (2005). *Clickers in the Classroom: How to Enhance Science Teaching Using Classroom Response Systems*. Pearson Education.
- Edwards, B. I., Aris, B., Harun, J., & Mohammed, H. (2015). Using response system through voting in peer instruction for learning sustainability. *Jurnal Teknologi*. <https://doi.org/10.11113/jt.v77.6371>
- Egelandsdal, K., & Krumsvik, R. J. (2015). Clickers and formative feedback at university lectures. *Education and Information Technologies*, 22(1), 55–74. <https://doi.org/10.1007/s10639-015-9437-x>

- Egelandsdal, K., & Krumsvik, R. J. (2017). Peer discussions and response technology: short interventions, considerable gains. *Nordic Journal of Digital Literacy*, *12*(1–2), 19–30. <https://doi.org/10.18261/issn.1891-943x-2017-01-02-03>
- Evans, C. (2015). Students' Perspectives on the Role of Peer Feedback in Supporting Learning. *Journal of Cognitive Education and Psychology*. <https://doi.org/10.1891/1945-8959.14.1.110>
- Fallon, M., & Forrest, S. L. (2011). High-Tech Versus Low-Tech Instructional Strategies. *Teaching of Psychology*, *38*(3), 194–198. <https://doi.org/10.1177/0098628311411896>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage.
- Fisher, D. B., Frey, N., & Hattie, J. A. (2016). *Visible Learning for Literacy, Grades K-12: Implementing the Practices That Work Best to Accelerate Student Learning*. Corwin Publishers.
- Fredricks, J. A., & Eccles, J. S. (2002). Children's Competence and Value Beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Developmental Psychology*, *38*(4), 519–533. <https://doi.org/10.1037/0012-1649.38.4.519>
- Ganley, C. M., & Lubienski, S. T. (2016). Mathematics Confidence, Interest, and Performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, *47*, 182–193. <https://doi.org/10.1016/j.lindif.2016.01.002>
- Gnambs, T., & Hanfstingl, B. (2015). The Decline of Academic motivation during Adolescence: An Accelerated longitudinal cohort analysis on the effect of Psychological need satisfaction. *Educational Psychology*, *36*(9), 1691–1705. <https://doi.org/10.1080/01443410.2015.1113236>
- González, A., Fernández, M. V. C., & Paoloni, P. V. (2016). Hope and anxiety in physics class: Exploring their motivational antecedents and influence on metacognition and

performance. *Journal of Research in Science Teaching*, 54(5), 558–585.

<https://doi.org/10.1002/tea.21377>

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: a longitudinal study. *Journal of Educational Psychology*, 93(1), 3–13. <https://doi.org/10.1037/0022-0663.93.1.3>

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2007). Multivariate Latent Change Modeling of developmental decline in academic intrinsic math motivation and achievement: Childhood through Adolescence. *International Journal of Behavioral Development*, 31(4), 317–327.

<https://doi.org/10.1177/0165025407077752>

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>

Hattie, J. (2008). Visible Learning. In *Routledge eBooks*. Informa.

<https://doi.org/10.4324/9780203887332>

Henderson, C., Beach, A., & Finkelstein, N. D. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, 48(8), 952–984. <https://doi.org/10.1002/tea.20439>

Hoekstra, A. (2008). Vibrant student voices: exploring effects of the use of clickers in large college courses. *Learning, Media and Technology*, 33(4), 329–341.

<https://doi.org/10.1080/17439880802497081>

Hogan, K., Nastasi, B.K., & Pressley, M. (1999). Discourse Patterns and Collaborative Scientific Reasoning in Peer and Teacher-Guided Discussions. *Cognition and Instruction*, 17, 379-432.

Hop, M., Janssen, J., & Engelen, R. (2017). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 7*. Cito. Geraadpleegd op 4 maart 2023, van:

<https://www.cito.nl/kennis-en-innovatie/kennisbank/006-wetenschappelijke-verantwoording-lvs-toetsen-rekenen-wiskunde-3-0-voor-groep-7>

Hop, M., Scheltens, F., & Engelen, R. (2019). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 8*. Cito. Geraadpleegd op 4 maart 2023, van:

<https://www.cito.nl/kennis-en-innovatie/kennisbank/166-1-wetenschappelijke-verantwoording-lvs-toetsen-rekenen-wiskunde-3-0-voor-groep-8>

Hunsu, N. J., Adesope, O., & Bayly, D. J. (2016). A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Computers & Education*, 94, 102–119. <https://doi.org/10.1016/j.compedu.2015.11.013>

Inspectie van het Onderwijs (2023). *De Staat van het Onderwijs 2023*. Inspectie van het Onderwijs. Geraadpleegd op 8 december 2023, van

<https://www.onderwijsinspectie.nl/documenten/rapporten/2023/05/10/rapport-de-staat-van-het-onderwijs-2023>

Irving, K. (2015). Technology-assisted formative assessment. In M. J. Urban & D. A. Falvo (Eds.), *Improving K-12 STEM education outcomes through technological integration* (380 – 398). <http://doi.org/10.4018/978-1-4666-9616-7.ch017>

Johnson, D. W., & Johnson, R. T. (1999). *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning*. Pearson.

Jones, M., Antonenko, P., & Greenwood, C. (2012). The impact of collaborative and individualized student response system strategies on learner motivation, metacognition, and knowledge transfer. *Journal of Computer Assisted Learning*, 28(5), 477–487. <https://doi.org/10.1111/j.1365-2729.2011.00470.x>

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3), 819–827. <https://doi.org/10.1016/j.compedu.2009.05.001>

- Kaymak, S., Balta, N., Almas, A., Kazmagambet, B., & Mbala, O. (2020). The Impact of Peer Instruction on Students' Achievement in Mathematics Analysis. *Journal of Psychology Research*, 10(12). <https://doi.org/10.17265/2159-5542/2020.12.003>
- Kiplagat, P.K. (2016). *Rethinking primary school mathematics teaching: a formative assessment approach*. Geraadpleegd op 5 maart 2023, van [https://www.researchgate.net/publication/313521889\\_RETHINKING\\_PRIMARY\\_SCHOOL\\_MATHEMATICS\\_TEACHING\\_A\\_FORMATIVE\\_ASSESSMENT\\_APPROACH](https://www.researchgate.net/publication/313521889_RETHINKING_PRIMARY_SCHOOL_MATHEMATICS_TEACHING_A_FORMATIVE_ASSESSMENT_APPROACH)
- Knight, J. K., Wise, S., & Southard, K. M. (2013). Understanding Clicker Discussions: Student reasoning and the impact of instructional cues. *CBE- Life Sciences Education*, 12(4), 645–654. <https://doi.org/10.1187/cbe.13-05-0090>
- Knight, J. K., & Brame, C. J. (2018). Peer Instruction. *CBE—Life Sciences Education*, 17(2), fe5. <https://doi.org/10.1187/cbe.18-02-0025>
- Konold, K. B., Miller, S. C., & Konold, K. B. (2004). Using Teacher Feedback to Enhance Student Learning. *Teaching Exceptional Children*, 36(6), 64–69. <https://doi.org/10.1177/004005990403600608>
- Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The Effectiveness and Features of Formative Assessment in US K-12 Education: A Systematic Review. *Applied Measurement in Education*, 33(2), 124–140. <https://doi.org/10.1080/08957347.2020.1732383>
- Leenknecht, M., Wijnia, L., Köhler, M., Fryer, L., Rikers, R., & Loyens, S. (2020). Formative assessment as practice: the role of students' motivation. *Assessment & Evaluation in Higher Education*, 46(2), 236–255. <https://doi.org/10.1080/02602938.2020.1765228>

- Lemons, P. P., & Lemons, J. D. (2013). Questions for Assessing Higher-Order Cognitive Skills: It's not just Bloom's. *CBE- Life Sciences Education*, *12*(1), 47–58.  
<https://doi.org/10.1187/cbe.12-03-0024>
- Leung, K. C. C. (2014). 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>
- Levesque, A. A. (2011). Using Clickers to Facilitate Development of Problem-Solving Skills. *CBE- Life Sciences Education*, *10*(4), 406–417. <https://doi.org/10.1187/cbe.11-03-0024>
- Lewin, J. D., Vinson, E. L., Stetzer, M. R., & Smith, M. K. (2016). A Campus-Wide Investigation of Clicker Implementation: The Status of Peer Discussion in STEM Classes. *CBE—Life Sciences Education*, *15*(1), ar6. <https://doi.org/10.1187/cbe.15-10-0224>
- Liu, N., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290.  
<https://doi.org/10.1080/13562510600680582>
- Majeed, A., Darmawan, I. G. N., & Lynch, P. E. (2013). A confirmatory factor analysis of Attitudes Toward Mathematics Inventory (ATMI). *The Mathematics Educator*, *15*(1), 121–135.
- Martínez, S. R., Regueiro, B., Piñeiro, I., Estévez, I., & Arias, A. V. (2020). Gender Differences in Mathematics Motivation: Differential effects on performance in Primary education. *Frontiers in Psychology*, *10*.  
<https://doi.org/10.3389/fpsyg.2019.03050>
- Mazur, E. (1997). Peer instruction: Getting students to think in class. *AIP Conference Proceedings*. <https://doi.org/10.1063/1.53199>



- McCarthy, J. (2017). Enhancing feedback in higher education: Students' attitudes towards online and in-class formative assessment feedback models. *Active Learning in Higher Education*, 18(2), 127–141. <https://doi.org/10.1177/1469787417707615>
- McConlogue, T. (2015). Making judgements: investigating the process of composing and receiving peer feedback. *Studies in Higher Education*, 40(9), 1495–1506. <https://doi.org/10.1080/03075079.2013.868878>
- McConnell, D., Steer, D., Owens, K. A., Knott, J. R., Van Horn, S. E., Borowski, W. S., Dick, J. E., Foos, A., Malone, M., McGrew, H. S., Greer, L., & Heaney, P. J. (2006). Using Conceptests to Assess and Improve Student Conceptual Understanding in Introductory Geoscience Courses. *Journal of geoscience education*, 54(1), 61–68. <https://doi.org/10.5408/1089-9995-54.1.61>
- Molin, F., Cabus, S., Haelermans, C., & Groot, W. (2019). Toward reducing anxiety and increasing performance in physics education: evidence from a randomized experiment. *Research in Science Education*, 51(S1), 233–249. <https://doi.org/10.1007/s11165-019-9845-9>
- Nicol, D. M. (2010). From monologue to dialogue: improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, 35(5), 501–517. <https://doi.org/10.1080/02602931003786559>
- Nicol, D. M., & Boyle, J. (2003). Peer Instruction versus Class-wide Discussion in Large Classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, 28(4), 457–473. <https://doi.org/10.1080/0307507032000122297>
- Nicol, D. M., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: a peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102–122. <https://doi.org/10.1080/02602938.2013.795518>

- Nimon, K., & Oswald, F. L. (2013). Understanding the Results of Multiple Linear Regression. *Organizational Research Methods*, 16(4), 650–674.  
<https://doi.org/10.1177/1094428113493929>
- Oloo, E., Mutsotso, S., & Masibo, E. (2016). Effect of Peer Teaching among Students on their Performance in Mathematics. *International Journal of Scientific Research and Innovative Technology*, 3(12), 10–24. Geraadpleegd op 5 april 2023, van  
<http://erepository.kibu.ac.ke/handle/123456789/1092>
- Palm, T., Andersson, C., Boström, E., & Vingsle, C. (2017). A review of the impact of formative assessment on student achievement in mathematics. *Nordic Studies in Mathematics Education*, 22(3), 25-50.
- Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., & Cooper, S. (2010). Does Displaying the Class Results Affect Student Discussion during Peer Instruction? *CBE—Life Sciences Education*, 9(2), 133–140. <https://doi.org/10.1187/cbe.09-11-0080>
- Porter, L., Bailey Lee, C., Simon, B., Cutts, Q., & Zingaro, D. (2011). Experience report: A Multi-classroom Report on the Value of Peer Instruction. *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education*. <https://doi.org/10.1145/1999747.1999788>
- Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion*, 13(2), 83–103. <https://doi.org/10.1007/bf00992956>
- Romero, A. A., & Angeles, E. D. (2023). Flipped Classroom in a Digital Learning Space: Its Effect on the Students' Attitude Toward Mathematics. *International Journal of*

*Learning, Teaching and Educational Research*, 22(1), 210–227.

<https://doi.org/10.26803/ijlter.22.1.12>

- Ruijsenaars, W., Minnaert, A., & Ghesquière, P. (2014). Leerproblemen en leerstoornissen. In P. Prins, & C. Braet (Eds.), *Handboek klinische ontwikkelingspsychologie* (tweede, geheel herziene druk ed., pp. 349-371). Bohn, Stafleu, Van Loghum.
- Ryan, R. M., Deci, E. L., Vansteenkiste, M., & Soenens, B. (2021). Building a Science of Motivated Persons: Self-determination Theory's empirical approach to human experience and the regulation of behavior. *Motivation science*, 7(2), 97–110.
- <https://doi.org/10.1037/mot0000194>
- Schell, J., & Butler, A. C. (2018). Insights from the science of learning can inform Evidence-Based implementation of peer instruction. *Frontiers in Education*, 3.
- <https://doi.org/10.3389/feduc.2018.00033>
- Schildkamp, K., Heitinka, M., van der Kleij, F., Hoogland, I., Dijkstra, A., Kippers, W., & Veldkamp, B. (2014). *Voorwaarden voor effectieve formatieve toetsing: Een praktische review* (Nr. 405-14–534). Universiteit Twente. Geraadpleegd op 9 maart 2023, van <https://www.nro.nl/sites/nro/files/migrate/Eindrapport-formatief-toetsen-Utwente-revisie.pdf>
- Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology*, 60, 101832.
- <https://doi.org/10.1016/j.cedpsych.2019.101832>
- Shamir, A., Tzuriel, D., & Rozen, M. (2006). Peer Mediation. *School Psychology International*, 27(2), 209–231. <https://doi.org/10.1177/0143034306064548>
- Simon, B., & Cutts, Q. (2012). Peer instruction. *Communications of the ACM*, 55(2), 27–29.
- <https://doi.org/10.1145/2076450.2076459>

- Sluijsmans, D. M. A., Joosten-ten Brinke, D., van der Vleuten, C. (2013). *Toetsen met leerwaarde: een reviewstudie naar de effectieve kenmerken van formatief toetsen*. Den Haag: NWO.
- Sluijsmans, D., & Kneyber, R. (2016). *Toetsrevolutie: Naar een feedbackcultuur in het voortgezet onderwijs*. Phronese. <https://toetsrevolutie.nl/wp-content/uploads/2016/11/Toetsrevolutie-WEB.pdf>
- Smith, L. A., Shon, H., & Santiago, R. (2011). Audience Response Systems: Using “Clickers” to Enhance BSW Education. *Journal of Technology in Human Services*, 29(2), 120–132. <https://doi.org/10.1080/15228835.2011.587737>
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equations models. In S. Leinhardt (Ed.), *Sociological Methodology*, 290–312. San Francisco: Jossey-Bass.
- Strijbos, J. W., & Wichmann, A. (2017). Promoting learning by leveraging the collaborative nature of formative peer assessment with instructional scaffolds. *European Journal of Psychology of Education*, 33(1), 1–9. <https://doi.org/10.1007/s10212-017-0353-x>
- Tapia, M. (1996). *The Attitudes toward Mathematics Instrument*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Tuscaloosa, AL, November 6-8, 1996)
- Tapia, M., & Marsh, G. E., II. (2002). *Confirmatory factor analysis of the Attitudes Toward Mathematics Inventory*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Chattanooga, TN. (ERIC Document Reproduction Service No. ED 471301).
- Tapia, M., & Marsh, G. E., II. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, 8(2), 16-21.

- Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68(3), 249. <https://doi.org/10.2307/1170598>
- Tullis, J. G., & Goldstone, R. L. (2020). Why does peer instruction benefit student learning? *Cognitive Research: Principles and Implications*, 5(1).  
<https://doi.org/10.1186/s41235-020-00218-5>
- van den Berg, M., Bosker, R. J., & Suhre, C. J. (2017). Testing the effectiveness of classroom formative assessment in Dutch primary mathematics education. *School Effectiveness and School Improvement*, 29(3), 339–361.  
<https://doi.org/10.1080/09243453.2017.1406376>
- van der Lubbe, M. (2021). Veel leerlingen leren niet zo goed rekenen als ze zouden kunnen. *Volgens Bartjens*, 41(2), 4-7. Geraadpleegd op 18 februari 2023, van  
[https://www.volgens-bartjens.nl/media/3/vb41\\_2\\_peil.rekenen\\_vanderlubbe.pdf](https://www.volgens-bartjens.nl/media/3/vb41_2_peil.rekenen_vanderlubbe.pdf)
- Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic Goal Contents in Self-Determination Theory: Another look at the quality of academic motivation. *Educational Psychologist*, 41(1), 19–31. [https://doi.org/10.1207/s15326985ep4101\\_4](https://doi.org/10.1207/s15326985ep4101_4)
- Vickrey, T., Rosploch, K., Rahmanian, R., Pilarz, M., & Stains, M. (2015). Research-Based Implementation of Peer Instruction: A Literature Review. *CBE—Life Sciences Education*, 14(1), es3. <https://doi.org/10.1187/cbe.14-11-0198>
- Wentzel, K. R., Battle, A., Russell, S. L., & Looney, L. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35(3), 193–202. <https://doi.org/10.1016/j.cedpsych.2010.03.002>

### **Bijlage**

#### ***Enjoyment-subschaal uit de ATMI (Tapia & Marsh, 2004)***

1. I get a great deal of satisfaction out of solving a mathematics problem.
2. I have usually enjoyed studying mathematics in school.
3. Mathematics is dull and boring.
4. I like to solve new problems in mathematics.
5. I would prefer to do an assignment in mathematics than to write an essay.
6. I really like mathematics.
7. I am happier in a mathematics class than in any other class.
8. Mathematics is a very interesting subject
9. I am comfortable expressing my own ideas on how to look for solutions to a difficult problem in mathematics.
10. I am comfortable answering questions in mathematics class.

#### ***Vertaalde Enjoyment-Subschaal uit de ATMI (Tapia & Marsh, 2004)***

1. Ik haal veel voldoening uit het oplossen van een rekensom.
2. Ik heb rekenen meestal leuk gevonden op school.
3. Wiskunde is suf en saai.
4. Ik vind het leuk om nieuwe rekenopdrachten op te lossen.
5. Ik maak liever rekenopdrachten dan dat ik een verhaal schrijf.
6. Ik vind rekenen erg leuk.
7. Ik ben gelukkiger in de rekenles dan bij iedere andere les.
8. Rekenen is een erg interessant vak.
9. Ik voel me comfortabel om de stappen die ik zet om een moeilijke rekensom op te lossen te vertellen.
10. Ik voel me comfortabel bij het beantwoorden van vragen tijdens de rekenles.