

ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

Bachelor Thesis Economics and Business Economics

**Het Effect van Formatieve Feedback op Leerprestaties bij
Basisschoolkinderen: Een Kwantitatieve Studie**

Louise Stoop (617803)



Supervisor: J. Korpershoek

Datum: 10 July 2024

Voorwoord

Voor u ligt de bachelorscriptie “Het Effect van Formatieve Feedback op Leerprestaties bij Basisschoolkinderen: Een Kwantitatieve Studie”. Deze scriptie onderzoekt de effecten van beoordelingscriteria (summatief versus formatief) op de prestaties van leerlingen op de basisschool tussen de zes en acht jaar oud. Met het afronden van mijn bachelor scriptie is het einde van mijn leerzame bacheloropleiding economie en bedrijfseconomie aan de Erasmus Universiteit Rotterdam bereikt. Dit onderzoek zou niet mogelijk zijn geweest zonder de hulp en steun die ik heb gekregen vanuit mijn omgeving.

Om te beginnen zou ik graag mijn scriptiebegeleider J. Korpershoek willen bedanken voor zijn geduld en waardevolle feedback tijdens het onderzoeksproject. Zijn ondersteuning heeft mij enorm geholpen met het voltooien van mijn scriptie.

Daarnaast wil ik de basisscholen en hun leraren die hebben deelgenomen aan dit experiment hartelijk danken. Het enthousiasme van de leraren en de openheid van de scholen hebben een cruciale rol gespeeld voor de totstandkoming van dit onderzoek.

Tot slot gaat er een speciaal dankwoord uit naar mijn familie. Zij hebben mij gedurende het onderzoek ondersteund. De aanmoediging heeft mij door deze periode heen geholpen en bijgedragen aan mijn persoonlijke en academische groei.

Ik hoop dat mijn onderzoek een waardevolle bijdrage kan leveren aan de ontwikkeling van het onderwijs en dat mijn onderzoek zal inspireren tot verdere studies over de impact van beoordelingsmethoden op de leerprestaties van jonge kinderen.

Met dankbaarheid,

Louise Stoop

Barendrecht, 14 juni 2024

Abstract

Dit onderzoek gaat in op de invloed van beoordelingscriteria (summatief versus formatief), op de academische prestaties van basisschoolleerlingen van zes tot en met acht jaar in Nederland. Summatieve beoordeling, in de vorm van cijfers, markeert het einde van een leerperiode en kan hiermee het leerproces van jonge kinderen verhinderen. Daarentegen kan formatieve feedback, die gericht is op voortgang en het verdere leerproces, effectiever zijn. De onderzoeksvraag luidt: “Hoe beïnvloeden de beoordelingscriteria (summatief versus formatief) de academische prestaties van Nederlandse basisschoolleerlingen van zes tot en met acht jaar?”. De summatieve feedback is gegeven in de vorm van een cijfer en de formatieve feedback is gegeven aan de hand van de parel en puzzelmethode.

In een kwantitatief experiment zijn de prestaties van leerlingen die zijn beoordeeld aan de hand van cijfers vergeleken met de prestaties van leerlingen die zijn beoordeeld met formatieve feedback. Uit de resultaten blijkt dat de groep met formatieve feedback gemiddeld 0.87 punt hoger scoort op de eindtoets. Bovendien blijkt formatieve feedback bij te dragen aan een positiever zelfbeeld, hogere intrinsieke motivatie en een meer ondersteunende leeromgeving.

Deze studie suggereert dat formatieve feedback een waardevol instrument is in het basisonderwijs in Nederland. Echter, vanwege de geringe tijdspanne, is vervolgonderzoek naar de langetermijneffecten en de toepasbaarheid van de parel en puzzelmethode aanbevolen.

Inhoudsopgave

Inleiding	5
Theoretisch kader	7
Historische Achtergrond van het Cijfermatig Beoordelingssysteem	7
Kritiek op het Huidige Beoordelingssysteem	7
De Impact van Summatieve Beoordeling op Leerlingen	8
Formatieve Beoordeling als Alternatief	9
De Parel en Puzzel Methode	10
Empirische Ondersteuning van Formatieve Feedback	11
Doelstelling en Relevantie	12
Methodologie	13
Dataverzamelmethode	13
Onderzoeksverloop	14
Datakenmerken	16
Mogelijke Problemen en Oplossingen	18
Data-analyse	18
Validiteit en Betrouwbaarheid	19
Statistische Analyse	19
Resultaten	21
Discussie	29
Conclusie	31
Referenties	32
Bijlage A Parel en Puzzel Formulier	34
Bijlage B Contact Scholen	37
Bijlage B1 Verzoek deelnamen experiment scholen	37
Bijlage B2 Mail Docenten Controlegroep	39
Bijlage B3 Mail aan Docenten Behandelgroep	40
Bijlage C: Rekenlessen Experiment	42
Bijlage D Eindtoets Rekenen	64
Appendix A: Poweranalyse Berekening	67

Inleiding

Cijfers markeren het einde van een leerperiode en kunnen een negatief effect op het leer- en ontwikkelingsproces van jonge kinderen hebben. Dit kan ertoe leiden dat jonge kinderen een *fixed mindset* ontwikkelen (Steffens, 2022). Ze zijn dan overtuigd dat ze geboren zijn met bepaalde capaciteiten en competenties, waardoor ze denken dat extra inspanning om te leren geen zin heeft. Om de verschuiving van een *fixed mindset* naar een *growth mindset* te bevorderen, wordt er gekeken naar een nieuwe beoordelingsmethode waarbij de focus ligt op het leerproces.

Formatieve feedback is een beoordelingsmethode waarbij de sterke punten van een leerling worden benadrukt en de punten waar verbetering mogelijk is vroegtijdig worden geïdentificeerd. In tegenstelling tot de summatieve beoordelingen, die het eindresultaat van de leerperiode vaststellen, richten formatieve beoordelingen zich op het leerproces. Waar summatieve beoordelingen terugkijken op de prestaties, kijken formatieve beoordelingen juist vooruit om prestaties te verbeteren. Met het invoeren van formatieve feedback wordt er naar verbetering van de leercultuur gestreefd.

Dit onderzoek biedt inzicht over de korte termijneffecten van formatieve feedback op de leerprestaties van leerlingen op de basisschool tussen de zes en acht jaar oud in Nederland. Uit eerder onderzoek van Butler en Nissan (1986) is gebleken dat formatieve feedback een significant positief effect heeft op de intrinsieke motivatie en prestaties van leerlingen. Echter, het onderzoek dateert uit 1986 en is uitgevoerd in een ander land en in een andere leeftijdscategorie. Het onderzoek is daarom niet representatief voor basisschoolleerlingen in Nederland tussen de zes en acht jaar oud. Ook Black en William (1988) stellen vast dat scholen substantiële leeropbrengsten ervaren wanneer zij gebruik maken van formatieve feedback. Echter zijn deze bevindingen gebaseerd op een literatuurstudie en is er geen experimenteel onderzoek gedaan naar deze bevindingen.

Dit onderzoek draagt bij aan de bestaande literatuur door een beter begrip te creëren van hoe verschillende beoordelingsmethodes bijdragen aan de verbetering van academische prestaties van leerlingen op de basisschool. Dit onderzoek is maatschappelijk relevant omdat het onderzoek een bijdrage levert aan het verbeteren van effectieve beoordelingsmethodes om de kwaliteit van het onderwijs te verbeteren.

In dit onderzoek is het effect van de beoordelingsmethode (summatief versus formatief) op prestaties van leerlingen op de basisschool van zes tot en met acht jaar in Nederland onderzocht. De onderzoeksvraag van dit onderzoek luidt dan ook: “Hoe beïnvloedt de beoordelingscriteria (summatief versus formatief) de academische prestaties van Nederlandse basisschoolkinderen van zes tot en met acht jaar?”.

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLOPPRESTATIES

Om antwoord te geven op deze onderzoeksvraag is er een kwantitatief experimenteel onderzoek uitgevoerd. De data zijn verzameld aan de hand van toetsresultaten van leerlingen van de deelnemende scholen aan het experiment. In totaal hebben er 233 leerlingen deelgenomen aan dit experiment. Deze leerlingen, verdeeld over groep 4 en groep 5, zijn willekeurig toegewezen aan de controlegroep of de behandelgroep.

Het experiment is als volgt opgezet: elke leerling volgt negen rekenlessen. Na elke rekenles ontvangen de leerlingen afhankelijk van de experimentele conditie summatieve feedback of formatieve feedback. De summatieve feedback wordt gegeven in de vorm van een cijfer, dat onder de rekenopdrachten in het schrift van de leerling wordt geschreven. De formatieve feedback wordt gegeven aan de hand van de parel- en puzzelmethode, een methode die intensief gebruik maakt van formatieve feedback gedurende het leerproces. Hierbij ontvangen de leerlingen een parel- en puzzelformulier met gedetailleerde feedback van de docent.

Gedurende het experiment worden diverse rekenvaardigheden getest. Op de laatste dag van het experiment maken alle deelnemende leerlingen de eindtoets. De resultaten van de controlegroep en behandelgroep worden vervolgens met elkaar vergeleken. Uit de resultaten blijkt dat de groep die formatieve feedback ontvangt, gemiddeld 0.87 punt hoger scoort op het eindcijfer dan de groep die summatieve feedback ontvangt. Ook worden er procentueel gezien meer voldoende behaald in de groep die formatieve feedback ontvangt (82%) in vergelijking met de groep die summatieve feedback ontvangt (75%).

De volgende sectie van deze scriptie bevat een overzicht van de literatuur over de voor- en nadelen van de summatieve- en formatieve beoordelingsmethoden. Sectie 3 gaat in op de hypothese en de onderzoeksmethode. Vervolgens beschrijft sectie 4 de kwantitatieve data-analyse en de resultaten hiervan. Sectie 5 bevat de conclusie. Tot slot wordt in sectie 6 de discussie gepresenteerd.

Theoretisch kader

Historische Achtergrond van het Cijfermatig Beoordelingssysteem

De oorsprong van het huidige numerieke beoordelingssysteem dat op Nederlandse basisscholen wordt gebruikt, is niet precies bekend (Peeters, 2023). Wel is bekend dat cijfers al in 1880 werden toegepast op de zogenaamde Latijnse scholen, de voorlopers van het gymnasium. Cijfers worden gezien als een objectieve maatstaf. Ze zijn meetbaar en gemakkelijk te interpreteren voor buitenstaanders. Door de objectiviteit van cijfers wordt de besluitvorming over kinderen vereenvoudigd, bijvoorbeeld of kinderen overgaan naar het volgende leerjaar of niet, is eenvoudig en objectief te bepalen aan de hand van cijfers. Bovendien stimuleren cijfers de betrokkenheid van kinderen. De betrokkenheid van kinderen neemt toe wanneer ze beoordeeld worden op hun prestaties (Emons et al., 2021). Tot slot dienen cijfers als graadmeter voor de onderwijskwaliteit en zijn daarom zeer waardevol voor docenten.

Kritiek op het Huidige Beoordelingssysteem

Desondanks is er door de jaren heen veel kritiek op het huidige numerieke beoordelingssysteem dat in Nederland wordt gebruikt. In 1972 schreef hoogleraar Adriaan Groot een boek waarin hij zijn kritiek uitte op het huidige numerieke beoordelingssysteem. In zijn boek classificeert hij cijfers als ondemocratisch, willekeurig en inefficiënt. Groot betoogt dat de interpretatie van cijfers kan variëren tussen leraren waardoor er verschillende beoordelingscriteria ontstaan. Zo stelt hij dat het cijfer 5 of het cijfer 6 voor de ene leraar een andere betekenis heeft dan voor de andere leraar. Bovendien geven deze cijfers slechts een rangorde aan (een 8 is hoger dan een 4), maar is het intervalniveau van de cijfers niet proportioneel (een 8 is niet twee keer zo hoog als een 4).

Uit de analyse van Groot blijkt dat ongeveer een kwart van de studenten de toets niet haalt. Dit fenomeen staat bekend als de Wet van Posthumus, genoemd naar de geleerde Kees Posthumus. Deze wet beschrijft dat leraren hun beoordelingen altijd aanpassen aan het niveau van de groep. Volgens Groot is dit tegenstrijdig met de eerlijkheid van cijfers. Onderwijs is bedoeld om kinderen iets te leren en als een kind de stof niet voldoende beheerst, vindt hij dat de verantwoordelijkheid van de leerkracht (Groot, 1972). Gordon en Fay (2010) bevestigen de bevindingen van de Groot in recenter onderzoek. Uit de studie van Gordon en Fay (2010) blijkt dat leraren worden beïnvloed door diverse contextuele factoren en de samenstellingen van de klas, wat kan resulteren tot inconsistenties in cijfers.

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

In 2016 schreef Dylan William een boek waarin hij betoogt dat de traditionele numerieke beoordelingscriteria niet effectief zijn in het bevorderen van prestaties en de ontwikkeling van kinderen. Hij stelt dat cijfers te veel druk leggen op het kind, wat ten koste zou kunnen gaan van de ontwikkeling van de leerling. William (2016) pleit voor verschuiving van de focus van numerieke beoordeling naar het geven van feedback om leerprocessen te verbeteren.

De Impact van Summatieve Beoordeling op Leerlingen

In Nederland heerst een cultuur van prestatiegerichtheid die vaak omschreven wordt als een ‘prestatiecultuur.’ De maatschappij legt veel nadruk op resultaat en prestaties, wat zich weerspiegelt in de indeling van het onderwijssysteem dat gericht is op summatieve toetsing. Summatief toetsen is prestatiegericht toetsen, verbonden aan een norm. In Nederland is dit het cijfer. Leerlingen zijn hierdoor geneigd zich vooral te concentreren op wat nodig is om een voldoende te behalen, zonder zich volledig te verdiepen in het onderwerp of het leerproces (Phrones et al., 2016).

De summatieve beoordeling dient als eindpunt van het oordeel en markeert het einde van de leerperiode. Een veel voorkomend fenomeen dat voortkomt uit de summatieve beoordeling, in de vorm van een cijfer, is de fixed mindset. Studenten geloven dat hun intelligentie en vaardigheden vaste kenmerken zijn die niet kunnen veranderen. Dit kan leiden tot het vermijden van uitdagingen en het opgeven na tegenslagen uit angst om te falen (Haimovitz & Dweck, 2017).

Daarnaast kan een summatieve beoordeling in de vorm van cijfers een negatief effect hebben op de intrinsieke motivatie van de leerling (Van der Kooij & Kennisrotonde, 2016). Intrinsieke motivatie, het innerlijke verlangen om te leren en te groeien, gedijt wanneer kinderen nieuwsgierigheid en plezier ervaren in hun leeractiviteiten (Vansteenkiste et al., 2007). Wanneer een leerling verwacht beoordeeld te worden op prestaties, kan dit het gevoel van autonomie en intrinsieke motivatie ondermijnen. De motivatie van de leerling verschuift naar het vermijden van fouten in plaats van het streven naar begrip van de stof en groei. De druk om te presteren op basis van cijfers kan leiden tot angst, wat een negatief effect kan hebben op het leren (Van der Kooij & Kennisrotonde, 2016).

Onderzoek door Moneva (2019) toonde een significante relatie aan tussen het zelfvertrouwen van leerlingen en de uitvoeringstaken. Leerlingen met veel zelfvertrouwen kunnen taken gemakkelijker voltooien en zijn niet bang om aan activiteiten deel te nemen. Aan de andere kant vertoonden leerlingen met een laag zelfvertrouwen taken van mindere kwaliteit en waren ze terughoudender om deel te nemen aan activiteiten. Bovendien vertoonden studenten met een prominent niveau van zelfvertrouwen een

hoger motivatieniveau. Een verbetering van het zelfvertrouwen zal leiden tot betere prestaties (Moneva, 2019).

Formatieve Beoordeling als Alternatief

De tegenpool van summatieve beoordeling is de formatieve beoordeling. Formatieve feedback wordt gedefinieerd als feedback die een leerling ontvangt tijdens het leerproces om het leren te verbeteren (Gipps, 1994). Dit komt overeen met de definitie die Shute (2008) aan formatieve feedback geeft; *“Formative feedback is information communicated to the learner that is intended to modify his or her thinking or behavior to improve learning”*. In eerste instantie is het doel van formatieve feedback om mensen iets te leren. In dit onderzoek wordt formatieve feedback gebruikt als manier van beoordelen.

Formative feedback is één van de meest invloedrijke factoren op leerprestaties (Hattie & Timperley, 2007; Shute, 2008). Helaas wordt formatieve feedback nog te weinig in het onderwijs toegepast wegens het tijdsintensieve karakter. De formatieve beoordelingsmethode richt zich op het verzamelen van informatie over de prestaties van leerlingen met als doel om die prestaties te verbeteren en het leerproces te optimaliseren. Wanneer formatieve feedback wordt toegepast in het onderwijs, wordt er tijdens de lessen veel gebruik gemaakt van interactie. De interactieve aanpak geeft de docent de mogelijkheid om de prestaties van de leerlingen te observeren, meten en analyseren. Eventuele knelpunten kunnen zo tijdig worden gesignaleerd. De directe feedback op de prestaties van de leerlingen draagt bij aan een effectief leerproces. De gegeven feedback is constructief van aard, benadrukt de sterke punten en identificeert gebieden waar verbetering mogelijk is.

Volgens Dumont et al. (2010) vormt formatieve feedback de bouwstenen voor een innovatieve leeromgeving en ontwikkeling van de 21^e eeuw. Black en William (1988) menen dat formatieve feedback het hart van goed onderwijs vormt. Tevens stellen zij vast dat scholen substantiële leeropbrengsten tonen wanneer zij gebruik maken van formatieve feedback als beoordelingsmethode.

Ondanks dat diverse onderzoeken hebben vastgesteld dat het geven van formatieve feedback een effectieve manier kan zijn om de resultaten van leerlingen te laten stijgen, wordt formatieve feedback niet beschouwd als wondermiddel (Visscher & Ehren, 2011). Visscher en Ehren (2011) stellen dat formatieve feedback niet altijd, overal en in elke vorm werkt. De effecten van formatieve feedback zijn sterk afhankelijk van de mate waarin de ontvanger van de feedback leert over zijn handelen. Wanneer de mogelijkheden om te achterhalen wat de leerling verkeerd doet beperkt zijn, is de kans op prestatieverbetering klein, gepaard met een stijging in demotivatie (Visscher & Ehren, 2011). Simons (1996) stelt dat de prestatieverbetering na het geven van formatieve feedback soms kan tegenvallen,

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

omdat leerlingen slecht gebruik maken van de geboden feedback. Wel is hij van mening dat het geven van formatieve feedback leidt tot een hogere prestatieverbetering dan de summatieve beoordelingsmethode.

In tegenstelling tot de summatieve beoordelingsmethode, waarbij het gaat om het vaststellen van het eindresultaat van de leerperiode, richt formatieve toetsing zich op het leerproces. Waar summatieve beoordelingen terugkijken op de prestaties, kijken formatieve beoordelingen juist vooruit om prestaties te verbeteren. Met het invoeren van formatieve feedback wordt er gestreefd naar een leercultuur.

De Parel en Puzzel Methode

Volgens de Wet Primair Onderwijs (WPO) heeft het basisonderwijs als doel om kinderen een solide basis te bieden in hun leer- en ontwikkelingsproces. Uit eerder genoemd onderzoek blijkt echter dat het geven van een summatieve beoordeling in de vorm van cijfers een negatief effect heeft op het leer- en ontwikkelingsproces. Het leren stopt na het ontvangen van een cijfer. Daarnaast zien we een dalende trend in de prestaties als gevolg van het negatieve effect van cijfers op het zelfvertrouwen. Door de huidige beoordelingsmaatstaf ontwikkelen kinderen een fixed mindset. Ze zijn ervan overtuigd dat ze geboren zijn met bepaalde capaciteiten en competenties, waardoor ze denken dat extra inspanning om te leren geen zin heeft. Het geven van cijfers gaat de doelstelling van de WPO volledig voorbij. Het is van belang dat we kijken naar een methode waarbij het leerproces van het kind geoptimaliseerd wordt.

In dit onderzoek wordt een zelfontwikkelde aanpak geïntroduceerd: de parel en puzzel methode. Deze methode maakt intensief gebruik van formatieve feedback gedurende het leerproces. De parel en puzzel methode is gericht op het verzamelen van informatie over de prestaties van studenten, met als doel de prestaties te verbeteren en het leerproces te optimaliseren. Deze methode beschouwt de sterke punten van de leerling als “parels”. Het gaat hierbij om de waardevolle aspecten van het werk van de leerling, zoals creatieve ideeën, correcte antwoorden of een sterke argumentatie. De “parel” dient als metafoor voor de sterke punten van een leerling. De leraar dient de “parels” te benadrukken en zo de leerling intrinsiek te motiveren. Het doel van de “parels” is de verschuiving van een fixed mindset naar een growth mindset. Het trainen van de growth mindset kan leiden tot betere prestaties (Blackwell et al., 2007; Good et al., 2003; Yeager & Dweck, 2012) en het verbeteren van de *deep learning skills* (kritisch denken, samenwerken; softkills, 21st century skills) van de leerling (Grant & Dweck, 2003).

De metafoor “puzzels” symboliseert de uitdagingen en verbeterpunten van de leerling. De docent heeft het doel om deze “puzzels” gezamenlijk met de leerling op te lossen. Hierbij is het van groot belang om deze feedback op een constructieve manier over te brengen. De feedback wordt geacht negatieve kritiek te

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

vermijden en moet gericht zijn op suggesties voor verbetering. Von Bergen et al. (2014) hebben in het onderzoek de ineffectiviteit van de sandwich methode aangetoond. De sandwich methode is een manier van feedback geven waarbij positieve feedback wordt gegeven, gevolgd door kritiek, en dan wordt afgesloten met positieve feedback. Volgens het onderzoek van Von Bergen et al. (2014) heeft het uitspreken van vertrouwen in de verbetering van het resultaat de hoogste impact.

Het is van belang dat er tijdens de les veel interactie plaatsvindt. De interactieve aanpak geeft de docent de mogelijkheid om de prestaties van de leerlingen te observeren, meten en analyseren. Eventuele puzzels moeten vroegtijdig worden geïdentificeerd. De gegeven feedback gaat over het stimuleren van de groei van een leerling. De directe feedback op de prestaties van de leerlingen draagt bij aan een effectief leerproces.

Empirische Ondersteuning van Formatieve Feedback

Butler en Nisan (1986) hebben eerder onderzoek verricht naar het effect van verschillende feedbackcondities op intrinsieke motivatie en prestaties. Butler en Nisan (1986) hebben geconcludeerd dat feedback met taakgerichte opmerkingen leidt tot versterking van intrinsieke motivatie. Daarnaast gaven de resultaten aan dat het geven van numerieke cijfers de extrinsieke motivatie bevordert, maar dit ging ten koste van de intrinsieke motivatie. Tot slot leidt de afwezigheid van feedback tot een afname van prestaties. Deelnemers gaven aan geschreven feedback meer te waarderen dan numerieke cijfers.

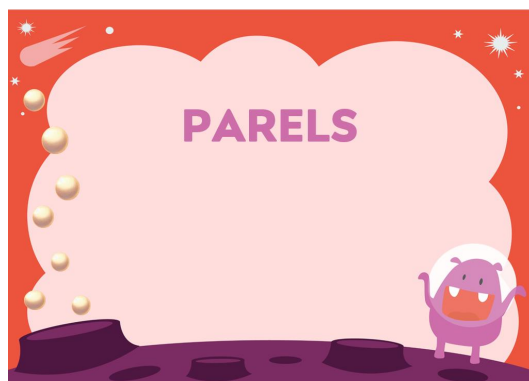
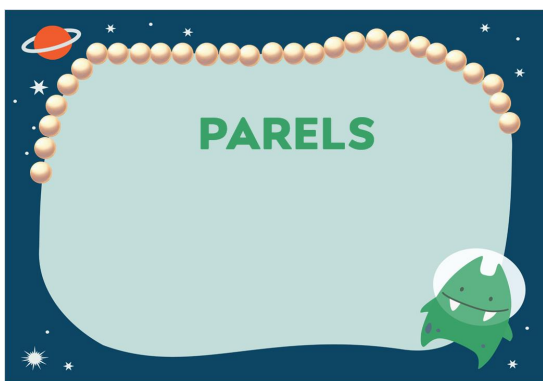
De omvang van het effect is als volgt: een ANOVA-analyse toonde significante verschillen tussen de feedbackcondities. Specifiek was de Cohen's d voor formatieve feedback versus geen feedback 0.5 (middelgroot effect), voor formatieve feedback versus numerieke beoordeling 0.65 (groot effect), en voor numerieke beoordeling versus geen feedback 0.3 (klein effect).

Het onderzoek van Butler en Nisan bestond uit 261 leerlingen van de zesde klas, die willekeurig werden toegewezen aan drie feedbackcondities: formatieve feedback, geen feedback of een numerieke beoordeling. De leerlingen hebben verschillende taken uitgevoerd, verdeeld over drie sessies. De resultaten bevestigen de prestatieverbetering wanneer de leerling formatieve feedback ontvangt. Dit experiment is uitgevoerd in Israël. Het onderwijs van Israël kent een geschiedenis waarbij de beoordelingscriteria gebaseerd zijn op een formatieve beoordelingscriteria als maatstaf van de resultaten (Butler & Nisan, 1986).

Uit de theorie is naar voren gekomen dat formatieve feedback een effectieve manier kan zijn om het leerproces te bevorderen. Dit heeft een positief effect op de prestaties. Echter is hier nog geen experimenteel onderzoek naar gedaan voor de leeftijdscategorie zes tot en met acht jaar in Nederland.

Doelstelling en Relevantie

Het doel van dit onderzoek is om te onderzoeken of het gebruik van formatieve feedback als beoordelingsmethode een positief effect heeft op de prestaties in vergelijking met het enkel toekennen van een cijfer. Om te meten of het effect positief is, worden de prestaties na het geven van formatieve feedback vergeleken met de prestaties van de leerlingen die alleen cijfers ontvangen. Om dit te testen, is er een experiment uitgevoerd. Hierbij worden de groepen 4 en 5 van de deelnemende scholen toegewezen aan een beoordelingsmethode. De groep die formatieve feedback ontvangt naar aanleiding van de parel en puzzel methode, ontvangt na elke gemaakte opdracht een parel en puzzel formulier zoals hieronder getoond (of zie Bijlage A). Het parel en puzzel formulier is een formulier speciaal ontworpen voor dit experiment. De leraar schrijft de formatieve feedback op het ontworpen parel en puzzel formulier. De voorkant van het formulier bestaat uit de parels, de sterke punten van de leerling, en de achterkant bestaan uit de puzzels, de uitdagingen van de leerling. De leerlingen in controlegroep ontvangen geen gedetailleerde feedback over hun sterke punten of verdere uitdagingen waar zij aan kunnen werken. In plaats daarvan wordt hun prestatie enkel samengevat in een numeriek cijfer wat onder hun werk wordt genoteerd.



Methodologie

Binnen dit onderzoek is een kwantitatief onderzoek uitgevoerd. Aan de hand van een experiment is het effect van de beoordelingscriteria (summatief versus formatief) op de academische prestaties van leerlingen op de basisschool in de leeftijdscategorie zes tot en met acht jaar in Nederland gemeten.

Dataverzamelmethode

Binnen dit onderzoek wordt een kwantitatief onderzoek uitgevoerd om de invloed van beoordelingscriteria op de academische prestaties van leerlingen op de basisschool van zes tot en met acht jaar in Nederland te onderzoeken. Er is gekozen voor een experimenteel ontwerp waarbij leerlingen willekeurig werden toegewezen aan twee verschillende groepen: Een groep die summatieve feedback ontving en een groep die formatieve feedback ontving. De leeftijdsgroep van zes tot en met acht jaar is gekozen omdat leerlingen in deze periode op school leren lezen, schrijven en rekenen. Dit zijn essentiële basisvaardigheden die een kind op de basisschool verwerft. Het optimaliseren van het leerproces van een kind tijdens deze periode is daarom van groot belang.

In totaal hebben 233 leerlingen van groepen 4 en 5 van verschillende basisscholen deelgenomen aan dit experiment. Er zijn 96 leerlingen toegewezen aan de summatieve feedbackgroep en 137 leerlingen zijn toegewezen aan de formatieve feedbackgroep. De toewijzing vond plaats door middel van de *flip the coin* methode. Door de leerlingen willekeurig toe te wijzen aan een groep wordt de bias geminimaliseerd.

De hypothese in dit onderzoek luidt als volgt “De beoordelingscriteria (summatief versus formatief) hebben een positief effect op de academische prestaties van basisschoolleerlingen in de leeftijdscategorie 6 tot en met 8 jaar in Nederland”.

Er wordt verwacht dat de academische prestaties van basisschoolkinderen in Nederland in deze leeftijdscategorie hoger zullen zijn wanneer formatieve beoordelingscriteria worden gebruikt in vergelijking tot summatieve beoordelingscriteria. Deze verwachting is gebaseerd op de bevindingen van onder andere het onderzoek van Butler en Nissan (1986). Butler en Nissan (1986) vonden een positief significant effect van het geven van formatieve feedback op prestaties. Echter is het onderzoek uitgevoerd in Israël. Het Israëliëse onderwijs is niet bekend met cijfers als beoordelingsmaatstaf. De context zou van invloed kunnen zijn op de resultaten. Het onderzoek is daarom niet representatief voor Nederland. Daarnaast is het experiment uitgevoerd met leerlingen tussen de 10 en 13 jaar oud.

Om de hypothese te toetsen, is veldonderzoek gedaan in de vorm van een experiment. Het experiment heeft plaats gevonden in mei 2024, gedurende de derde en vierde week van deze maand. Er is bewust

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

voor deze periode gekozen omdat deze periode geen bijzondere feest- of vakantiedagen bevat die de academische routine zou kunnen verstoren.

Onderzoeksverloop

De populatie binnen het onderzoek telt ongeveer een half miljoen leerlingen van zes tot en met acht jaar. Deze leerlingen staan ingeschreven op verschillende basisscholen. Deze leerlingen bevinden zich in groep 4 en groep 5 van de basisschool. Om de steekproefgrootte te berekenen, is een poweranalyse uitgevoerd waarbij rekening is gehouden met een aantal statistische parameters: een foutmarge van 5%, een betrouwbaarheidsinterval van 95%, en een verwachte effectgrootte van 0.5. Deze verwachting is gebaseerd op het onderzoek van Butler en Nissan (1986), waar een soortgelijke effectgrootte is gemeten.

De power-analyse gaf aan dat een steekproef van 63 leerlingen voldoende zou zijn om een statistisch significant effect te kunnen detecteren, rekening houdend met de verwachte variantie binnen de uitkomsten (zie Appendix A). Door het zorgvuldig bepalen van de steekproefgrootte wordt de kans op type II fouten geminimaliseerd, wat essentieel is voor de validiteit en betrouwbaarheid van de onderzoeksresultaten.

Om de gewenste steekproefgrootte te bereiken, zijn willekeurig 10 scholen binnen de omgeving Rijnmond benaderd om deel te nemen aan het experiment aan de hand van een e-mail (zie Bijlage B1). Van de benaderde scholen heeft 30% toegezegd te willen deelnemen aan het experiment. Er zijn enkele opvallende verschillen gedetecteerd tussen de scholen die wel en niet hebben deelgenomen aan het experiment. Over het algemeen zijn de scholen die hebben gekozen om deel te nemen aan het experiment kleiner van omvang, gemiddeld rond de 150 leerlingen, vergeleken met een gemiddelde van 275 leerlingen. Daarnaast bleken deelnemende scholen vaker gelegen te zijn in woonwijken met een hoger gemiddeld inkomensniveau. Ten slotte waren de deelnemende scholen vaker betrokken bij onderwijsvernieuwingen, zoals een groen schoolplein of het gebruik van tablets.

Vervolgens zijn de desbetreffende docenten van de groepen 4 en 5 van de deelnemende scholen geïnformeerd aan de hand van een e-mail (zie Bijlage B2 en B3). Tot slot is er een mail verstuurd naar de ouders/verzorgers van de leerlingen die geselecteerd zijn om mee te doen aan het experiment, om toestemming te vragen voor de deelname van hun kind.

Er is toestemming gevraagd aan 284 ouders/verzorgers. Hiervan hebben 16 ouders toestemming geweigerd en 20 ouders hebben niet gereageerd op de e-mail. De leerlingen waarvan de ouders geen toestemming hebben verleend of niet tijdig op de e-mail hebben gereageerd, zijn uitgesloten van het experiment. Er zijn leerlingen uit de steekproef gehaald wanneer zij één of meerdere keren niet aanwezig

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

waren in de weken waarin het experiment plaats vond. Wanneer een deelnemer niet aanwezig is kan dit leiden tot ontbrekende gegevens. Dit kan de betrouwbaarheid van de resultaten in gevaar brengen. Dit omvatte zes leerlingen. In totaal zijn er 42 leerlingen uitgesloten van het experiment.

Gedurende de reguliere lessen, zonder interventie, ontvangen leerlingen gewoonlijk geen summatieve feedback in de vorm van cijfers. Normaal gesproken krijgen zij enkel een cijfer na een toets. Formatieve feedback, zoals het parel en puzzel formulier, wordt ook niet gegeven. De gebruikelijke gang van zaken bestaat uit het geven van instructies voorafgaand aan de les en een klassikale bespreking waarin veelgemaakte fouten worden belicht. Bij deze vorm van lesgeven is de directe terugkoppeling minimaal.

Het onderzoek bestond uit twee fasen: een interventiefase en een evaluatiefase.

Interventiefase

De desbetreffende docenten van de groepen 4 en 5 van de deelnemende scholen zijn geïnformeerd via een e-mail (zie Bijlage B2 en B3). De docenten van de controlegroep hebben de instructie gekregen om een cijfer te noteren onderaan de rekenles, gebaseerd op de hoeveelheid goed beantwoorde vragen. Het cijfer werd berekend met de volgende formule: $(\text{Aantal goed} / \text{Totaal aantal opgaven}) \times 9 + 1$. Dit cijfer moest onderaan de pagina van de rekenles worden genoteerd, zonder verdere toelichting over de totstandkoming van het cijfer.

Tijdens de interventie zijn er twee contactmomenten geweest. Na de eerste week van de interventie is een check-in uitgevoerd om te kijken of de docenten problemen ondervonden. Zo was er onverwacht een docent ziek en was er een sportdag gepland, waardoor er die week geen rekenles werd gegeven. Daarom werd besloten deze groep uit te sluiten van het experiment om vertekening in de resultaten te voorkomen. Na het maken van de eindtoets vond een tweede check-in plaats om de ervaringen van de docenten te achterhalen. De docenten van de controlegroep rapporteerden meer onrust in de klas. Een docent gaf aan: “Kinderen zijn gewend cijfers te ontvangen voor hun toetsen, maar niet tussen de lessen door. Wanneer de cijfers werden uitgedeeld, creëerde dit veel onrust in de klas. Er ontstond competitie tussen de leerlingen. Sommige leerlingen bedekten hun cijfer of sloegen de bladzijde om, zodat het cijfer niet zichtbaar was voor anderen. Daarentegen waren er ook kinderen die hun cijfer door de groep heen riepen. De sfeer werd gespannener gedurende het experiment en er heerste meer angst om fouten te maken. Het was geen prettige leeromgeving”.

De leerlingen uit de behandelgroep kregen daarentegen een parel en puzzel formulier (formatieve feedback). Dit formulier, speciaal voor dit experiment ontworpen, bevat op de voorkant de parels, de

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

sterke punten van de leerling, en op de achterkant de puzzels, de verbeterpunten van de leerling. De docenten van de behandelgroep kregen de instructie om het parel en puzzel formulier na elke rekenles voor elke leerling in te vullen. Bij het invullen van de parels moest de nadruk liggen op specifieke, gedetailleerde en motiverende feedback, waarbij de sterke punten van de leerling werden belicht. Bij het invullen van de puzzels moest het verbeterpunt van de leerling worden vastgesteld, met de nadruk op groei en het vertrouwen van de docent in de leerling om het resultaat te verbeteren.

Evaluatiefase

Na de interventiefase wordt de eindtoets afgenomen. De resultaten van de eindtoets van de controlegroep worden vergeleken met de resultaten van de behandelgroep. De resultaten van de eindtoets worden vergeleken om de effecten van de verschillende feedback vormen te analyseren.

Datakenmerken

Om potentiële vertekening in de resultaten te minimaliseren en om interventies op groepsniveau te meten, wordt er gebruik gemaakt van clusterrandomisatie. Hierbij worden groepen willekeurig toegewezen aan de behandelgroep (die formatieve feedback ontvangt) of de controlegroep (die reguliere cijfers ontvangt). Binnen een school is er altijd een controlegroep en een behandelgroep aangewezen, waardoor clusterrandomisatie mogelijk is. Er is voor randomisatie op groepsniveau in plaats van individueel niveau gekozen omdat het zo mogelijk wordt om het effect van formatieve feedback tussen scholen te meten, en tegelijkertijd het spill-over effect te beperken. Het spill-over effect kan bijvoorbeeld optreden wanneer leerlingen uit de behandelgroep interactie hebben met de leerlingen uit de controlegroep, waardoor elementen van de formatieve feedback doorsijpelen naar de controlegroep. Door formatieve feedback binnen clusters te houden kan de invloed van enkele demografische factoren, zoals leeftijd, geslacht en sociaaleconomische status beter worden gecontroleerd. Door deze factoren constant te houden binnen de clusters, kan er met meer zekerheid gezegd worden dat verschillen in leerprestaties kunnen worden toegeschreven aan de beoordelingscriteria.

Het toewijzen van groepen aan de controlegroep of de behandelgroep binnen de scholen is via de “Flip the coin” methode gebeurd. Hierbij wordt er een munt opgegooid. Aan de hand van kop of munt worden de groepen toegewezen aan de controlegroep of de behandelgroep. In totaal zijn er 233 leerlingen, in de leeftijdscategorie van zes tot en met acht jaar. Deze zijn verdeeld over negen groepen die willekeurig zijn toegewezen aan de controlegroep of de behandelgroep. Hierbij zijn vier groepen toegewezen aan de controlegroep bestaande uit 100 leerlingen en vijf groepen zijn toegewezen aan de behandelgroep, bestaande uit 133 leerlingen (zie Tabel 1).

Tabel 1*Overzicht Datakenmerken*

Cluster	Aantal leerlingen	Percentage van totaal	Experimentele conditie
1	28	12.0	Behandeling
2	29	12.4	Behandeling
3	22	9.4	Controle
4	28	12.0	Behandeling
5	27	11.6	Behandeling
6	26	11.2	Controle
7	30	12.9	Controle
8	21	9.0	Behandeling
9	22	9.4	Controle
Totaal	233	100	

Het experiment is als volgt uitgevoerd: Elke leerling heeft negen lessen voor het vak rekenen gevolgd. Na elke rekenles is de controlegroep voorzien van een cijfer en de behandelgroep is voorzien van formatieve feedback. Het cijfer wordt genoteerd in het schrift van de leerling onder de opdrachten die de leerling heeft gemaakt. De formatieve feedback wordt uitgedeeld aan de leerlingen in de behandelgroep met een puzzel en parel formulier (zie Bijlage A). In deze twee weken zijn diverse rekenvaardigheden getest. Na de negen rekenlessen vond er een toets plaats waarbij de resultaten van de leerlingen zijn gemeten. Het is niet onwaarschijnlijk om te denken dat de subjecten van de controlegroep en de behandelgroep met elkaar in contact zijn gekomen. Echter ligt het niet in de lijn der verwachting dat contact tussen de leerlingen invloed heeft op de resultaten van het onderzoek omdat de individuele leerervaringen van de leerlingen niet eenvoudig kunnen worden overgedragen door enkel met elkaar te praten.

Mogelijke Problemen en Oplossingen

Eén van de hoofdproblemen bij de uitvoering van het experiment is het risico op verschillen tussen leerlingen die de resultaten van het experiment verstoren. Individuele verschillen tussen leerlingen zoals motivatie, thuissituatie of eerdere academische prestaties kunnen de resultaten beïnvloeden wanneer deze niet gelijkmatig verdeeld zijn over de controlegroep en behandelgroep. Door clusterrandomisatie worden deze verschillen geminimaliseerd door hele klassen of groepen te randomiseren, waardoor individuele verschillen binnen een groep worden gecompenseerd. Hiermee wordt de interne validiteit van het experiment verhoogd.

Een te kleine steekproef kan de betrouwbaarheid van de resultaten aantasten. Door gebrek aan statistische power wordt het lastig om een significant effect te detecteren. Door het uitvoeren van een poweranalyse en het willekeurig selecteren van basisscholen, wordt de externe validiteit van dit onderzoek gewaarborgd.

Wanneer er onverwachts deelnemers uitvallen tijdens het experiment, kan dit de betrouwbaarheid van de resultaten aantasten. Om de impact van onverwachte uitval te minimaliseren, worden er regelmatige check-ins met de leerkrachten ingepland om problemen die tot uitval kunnen leiden vroegtijdig te monitoren. Door een sensitiviteitsanalyse uit te voeren, wordt er onderzocht in welke mate de resultaten robuust zijn tegen uitval.

Tot slot, kan een andere beperking van het experiment de duur van het experiment zijn. Het is mogelijk dat het effect van formatieve feedback enkel op korte termijn wordt gemeten. Het experiment zou moeten worden uitgebreid met een langere follow-up periode, zodat zowel het korte- als langetermijneffect van formatieve feedback kan worden beoordeeld. Daarnaast is het van belang om rekening te houden met het spillover-effect. Dit effect kan bijvoorbeeld optreden wanneer de behandelgroep interactie heeft met de controlegroep. De elementen van de formatieve feedback kunnen zo doorsijpelen naar de controlegroep. Door dit effect te minimaliseren, kan worden vastgesteld of de waargenomen effecten daadwerkelijk kunnen worden toegeschreven aan de formatieve feedbackmethode.

Data-analyse

Het doel van de data-analyse is om effecten van formatieve feedback op academische prestaties van basisschoolleerlingen te onderzoeken en te kwantificeren. De behandel effecten zullen worden geïdentificeerd aan de hand van de scores van de gestandaardiseerde toetsen. De controlegroep (die cijfers ontvangt) en behandelgroep (die formatieve feedback ontvangt), worden vergeleken op basis van

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOHPRESTATIES

gemiddelde prestaties. Er is onderzocht of eventuele covariantie, zoals leeftijd, de effecten van formatieve feedback kunnen beïnvloeden.

Om de consistentie en spreiding van de academische prestaties binnen de controlegroep en de behandelgroep te beoordelen, wordt de variantie van de prestaties berekend en vergeleken tussen de controlegroep en de behandelingsgroep.

Validiteit en Betrouwbaarheid

In dit experiment wordt gebruik gemaakt van een gestandaardiseerde methodetoets van het vak rekenen.

De antwoorden van een rekentoets zijn eenvoudig objectief te beoordelen op basis van correctheid.

Binnen een rekentoets is geen ruimte voor subjectieve beoordeling. Ondanks dat de rekentoets hoofdzakelijke objectieve evaluatiecriteria hanteert, blijft het van belang de onafhankelijkheid van beoordelaars te waarborgen. Daarom is er in dit experiment gekozen om de beoordelaars niet op de hoogte te stellen van de experimentele conditie van elke leerling om bias in de resultaten te voorkomen.

Dit betekent dat de beoordelaars alleen de prestaties van de kinderen zullen zien, zonder informatie over de gebruikte beoordelingscriteria. Om onbewuste voorkeuren voor bepaalde leerlingen of groepen te filteren, worden de leerlingen volledig geanonimiseerd. Deze maatregel draagt bij aan de betrouwbaarheid en validiteit van de onderzoeksresultaten. Belangrijk is dat de beoordelaar van de eindtoets niet de leerkracht van de groep is. Bijvoorbeeld, de leerkracht van groep 4 kijkt de rekenlessen van groep 4 na maar beoordeelt niet de eindtoets van groep 4. In plaats daarvan krijgt zij de eindtoets van een andere groep toegewezen om na te kijken. Deze scheiding zorgt ervoor dat het onderzoeksteam en de beoordelaars van de prestaties van de leerlingen onafhankelijk blijven, wat cruciaal is voor een objectieve evaluatie van de effecten van formatieve feedback.

De beoordelaars zijn vrouwelijke docenten tussen de 45 en de 55 jaar van Nederlandse komaf. Alle docenten hebben de Pedagogische Academie voor het Basisonderwijs (PABO) afgerond. De leeftijd van de docenten suggereert een zekere mate van professionele ervaring en expertise. Het onderzoek is niet gerelateerd op een specifiek curriculum, maar bredere aspecten van academische prestaties in relatie tot beoordelingscriteria. Buiten de reguliere lessen worden er geen extra ondersteunende onderwijsmaatregelen getroffen. Dit benadrukt de natuurlijke omgeving van het onderzoek.

Statistische Analyse

Om de statistische significantie van het behandelingseffect te testen, is gebruik gemaakt van de randomisatie-inferentie methode. De randomisatie-inferentie methode houdt in dat de subjecten willekeurig worden toegewezen tot de behandelgroep of de controlegroep. Dit creëert homogeniteit tussen de groepen.

Eventuele verschillen tussen groepen zijn te wijten aan toeval en niet aan systematische verschillen tussen

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

groepen. Hiermee wordt de kans dat de twee groepen gelijkwaardig zijn gemaximaliseerd en wordt het effect van eventuele confounders verkleind. Door randomisatie-inferentie wordt de nulhypothese op een betrouwbare manier getest.

Resultaten

In dit hoofdstuk worden de resultaten van het uitgevoerde experiment weergegeven. De data zijn verzameld door middel van de toetsresultaten. De resultaten zijn verwerkt in het statistische programma SPSS¹. De resultaten worden gepresenteerd op basis van de hypothese zoals opgesteld in sectie twee: De beoordelingscriteria (summatief versus formatief) hebben een positief effect op de academische prestaties van basisschoolkinderen in Nederland, in de leeftijdscategorie 6 tot en met 8 jaar.

Beschrijvende Statistieken

De onderzochte groep bestaat uit 211 leerlingen, waarvan 96 leerlingen zijn toegewezen aan de controlegroep en 115 leerlingen zijn toegewezen aan de behandelgroep. De toetsresultaten lager dan 1 of hoger dan 10 worden aangemerkt als outliers en verwijderd uit de steekproef². Er zijn geen subjecten verwijderd uit de steekproef. Tabel 2 geeft de gemiddelde scores en de standaarddeviaties voor beide groepen weer.

Tabel 2

Beschrijvende Statistieken

Groep	N	Gemiddelde score	Standaarddeviatie
Controlegroep	96	7,3	2,5
Behandelgroep	115	8,2	1,6

In totaal zijn er 112 subjecten jongen (53%) en 99 subjecten meisje (47%). Om te bepalen hoeveel jongens en meisjes zich binnen de controlegroep en de behandelgroep bevinden, wordt er een kruistabelanalyse uitgevoerd. Tabel 3 toont de verdeling van geslacht binnen de controlegroep en de behandelgroep. De controlegroep bestaat uit 51 jongens (53%) en 45 meisjes (47%). De behandelgroep bestaat uit 61 jongens (53%) en 54 meisjes (47%).

¹ IBM SPSS Statistics versie 29

² De drempelwaarden voor outliers zijn gebaseerd op standaardpraktijken voor gegevensanalyse zoals beschreven in Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.

Tabel 3
Verdeling Geslacht per Groep

Type Beoordeling	Jongen	Percentage	Meisje	Percentage
Controlegroep	51	53%	45	47%
Behandelgroep	61	53%	54	47%
Totaal	112	53%	99	47%

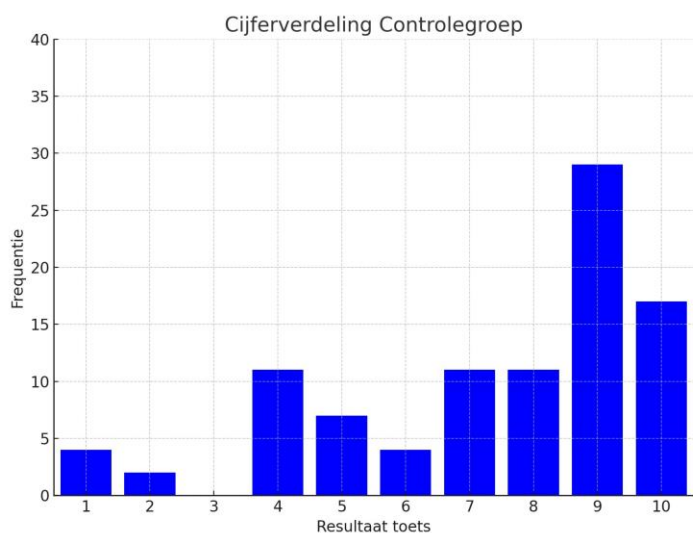
Om inzicht te krijgen van de cijferverdeling in de controlegroep en de behandelgroep zijn de toetsresultaten gevisualiseerd met behulp van histogrammen. De verdeling van de toetsresultaten in de controlegroep wordt weergegeven in Figuur 1. De resultaten variëren van 1.0 tot 10.0, met een gemiddelde score is 7.3 en de standaarddeviatie is 2.5 (zie tabel 2). De verdeling van de toetsresultaten in de behandelgroep wordt weergegeven in Figuur 2. De toetsresultaten variëren van 4.0 tot 10.0, met een gemiddelde score van 8.2 en een standaarddeviatie van 1.6 (zie tabel 2).

Figuur 1 toont de behaalde resultaten van de controlegroep. De cijfers variëren van 1 tot 10. De verdeling van de cijfers lijkt bimodaal met twee duidelijke pieken. De figuur laat zien dat de scores tussen de 8 en 10 de hoogste frequentie hebben. De controlegroep heeft een hoge spreiding vergeleken met de behandelgroep.

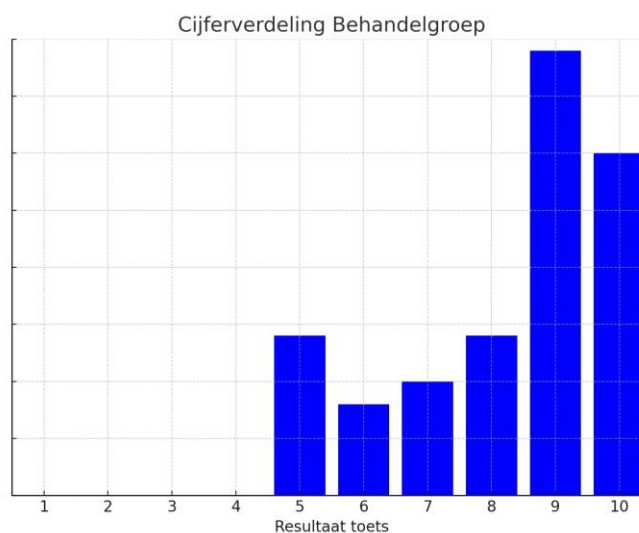
Figuur 2 toont de behaalde resultaten van de behandelgroep. De cijfers variëren van 4 tot 10. De verdeling van de cijfers in de behandelgroep lijkt meer unimodaal en rechtsscheef. Er zijn minder lage scores behaald in de behandelgroep in vergelijking met de controlegroep. Dit zou er op kunnen duiden dat de leerlingen die normaal lage scores halen er aanzienlijk op vooruit gaan. In de behandelgroep zijn de resultaten meer geconcentreerd rondom de hogere scores. Er is een duidelijke piek te zien in de resultaten tussen de 9 en 10. Er lijkt minder spreiding te zijn in vergelijking met de controlegroep.

De visualisatie van de toetsresultaten toont nog meer verschillen tussen de controlegroep en de behandelgroep. De gemiddelde score van de behandelgroep (8.19) is hoger dan de gemiddelde score van de controlegroep (7.32). De standaarddeviatie in de behandelgroep (1.632) is lager dan de standaarddeviatie in de controlegroep (2.502). Dit wijst op meer variabiliteit in de resultaten van de controlegroep.

Figuur 1
Resultaten Controlegroep



Figuur 2
Resultaten Behandelgroep



Hypothese Toetsen

De hypothese wordt getest aan de hand van een ongepaarde *t*-toets. Er wordt onderzocht of er een significant verschil is in de gemiddelde toetsscores tussen de controlegroep ($M=7.3$, $SD= 2.5$) en de behandelgroep ($M=8.2$, $SD=1.6$). Tabel 3 toont de resultaten van de ongepaarde *t*-toets. De resultaten wijzen op een significant verschil, $t(157.916) = -2.93$, $p = .040$. Dit suggereert dat de behandelgroep gemiddeld beter scoort voor de toets na het krijgen van formatieve feedback dan de controlegroep na het krijgen van cijfers. Het verschil is gemiddeld 0.87 punten.

Tabel 3
Resultaten van de Onafhankelijke T-Toets

Maatregel	Waarde
<i>T</i> -waarde	-2.93
Vrijheidsgraden (df)	157.92
<i>P</i> -waarde (tweezijdig)	0.004
Gemiddeld verschil	-0.87
Standaard verschil	0.30
95% CI Laag	-1.46
95% CI Hoog	-0.28

**Voor deze analyse zijn ongelijke varianties verondersteld*

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLOPPRESTATIES

Levene's test wordt uitgevoerd om de gelijkheid in varianties van beide groepen te controleren. Levene's test voor gelijkheid van variantie is significant ($F = 17.72, p < .001$). De nulhypothese van gelijke variantie wordt verworpen. De resultaten voor de ongepaarde t -toets wordt daarom gerapporteerd aan de hand van ongelijke varianties.

Overige Analyses

Naast het effect van de beoordelingscriteria op de resultaten, is er ook gekeken naar het verschil tussen jongens en meisjes. Om de gemiddelde scores van jongens en meisjes binnen de controlegroep en de behandelgroep te bepalen, wordt een analyse in SPSS uitgevoerd waarbij gemiddelde scores per geslacht worden berekend en vergeleken. Zoals weergegeven in Tabel 5, tonen de resultaten aan dat in de controlegroep jongens een gemiddelde score van 7.2 ($SD = 2.6$) behalen en de meisjes een score van 7.4 ($SD = 2.4$). In de behandelgroep scoren de jongens opnieuw lager dan de meisjes. De gemiddelde score van een man in de behandelgroep is 8.0 ($SD = 1.6$) terwijl een vrouw gemiddeld 8.4 ($SD = 1.6$) scoort.

Tabel 5

Beschrijvende Statistieken per Geslacht en Groep

Type Beoordeling	Geslacht subject	Gemiddelde	N	SD
Controlegroep	Man	7.21	51	2.57
	Vrouw	7.44	45	2.45
Behandelgroep	Man	8.00	61	1.63
	Vrouw	8.40	54	1.63
Totaal	Man	7.64	112	2.14
	Vrouw	7.97	99	2.09

Er wordt een tweerichtings-ANOVA uitgevoerd om het interactie-effect van geslacht en groep te meten en het effect van geslacht en de groep op de scores te onderzoeken. Tabel 6 toont de resultaten van de ANOVA test. De ANOVA test toont aan dat er geen significant effect van geslacht op scores bestaat, $F(1, 207) = 1.22, p = .271$. Dit resultaat suggereert dat het geslacht van het subject geen invloed heeft gehad op de toetsscore.

Tabel 6*Resultaten ANOVA test*

Variabele	<i>F</i>	<i>p</i> -waarde	Partial Eta Squared
Groep	9.29	0.00	0.04
Geslacht	1.22	0.27	0.01
Groep*Geslacht	0.08	0.77	0.00

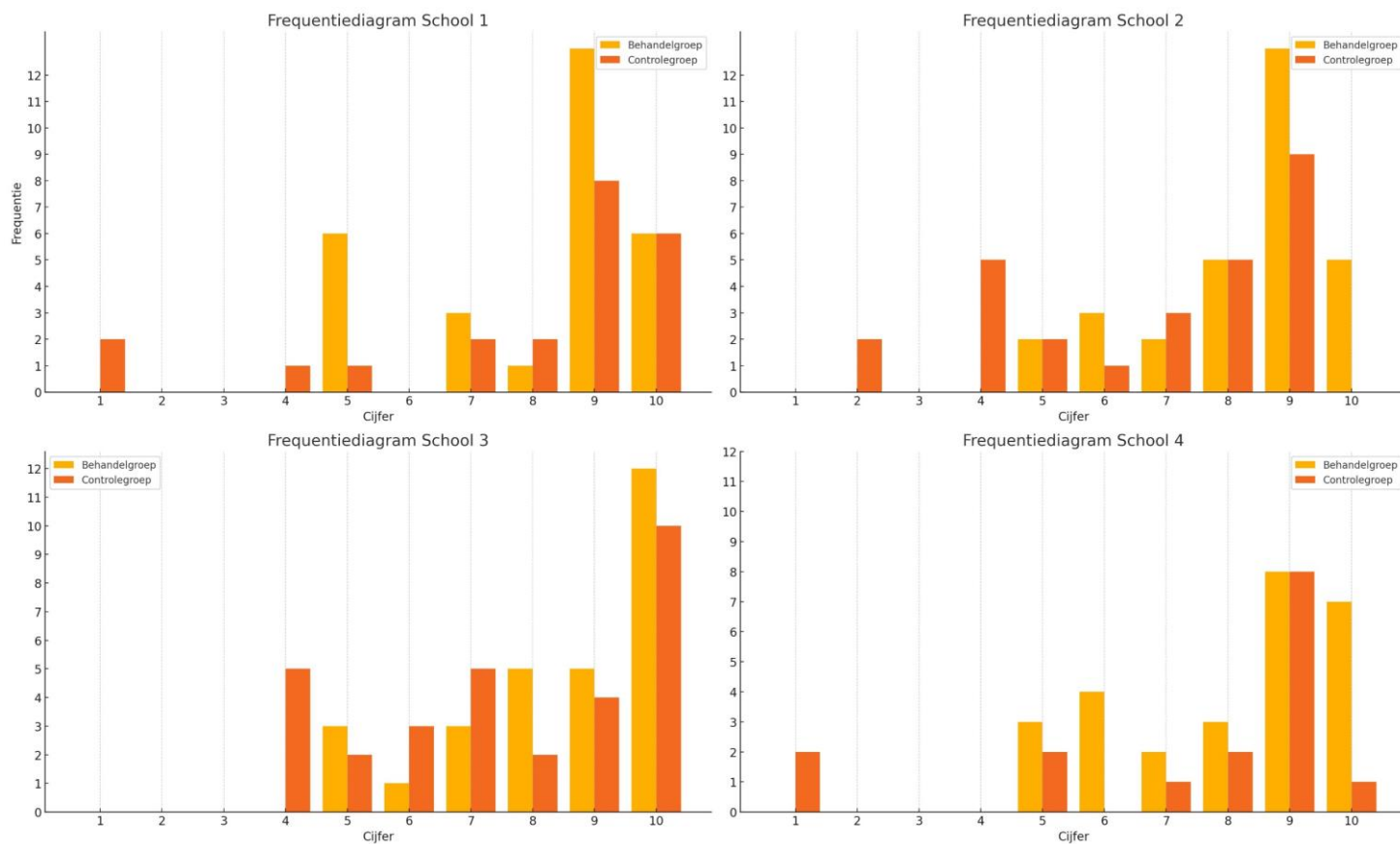
Tot slot wordt er gekeken naar het aantal leerlingen dat de toets heeft gehaald. De toets wordt gezien als gehaald, wanneer de leerling een 5.5 of hoger scoort. In de controlegroep heeft 75% een voldoende voor de toets behaald terwijl 88% van de leerlingen uit de behandelgroep de toets met een voldoende heeft afgerond. In totaal heeft 82% van de leerlingen een voldoende behaald voor de toets. (zie Tabel 7)

Tabel 7*Percentage Leerlingen met Voldoende Scores per Groep*

Groep	Aantal leerlingen	Aantal leerlingen met voldoende	Percentage leerlingen met voldoende
Controlegroep	96	72	75%
Behandelgroep	115	101	88%
Totaal	211	173	82%

Er is een analyse uitgevoerd waarbij de scholen met elkaar worden vergeleken. Figuur 3 toont het verschil in resultaten tussen de controlegroep en de behandelgroep binnen een school. De vergelijking tussen de scholen toont over het algemeen een positief effect van de interventie op de resultaten. Er zijn echter variaties in de effectiviteit van de interventie tussen de scholen.

Figuur 3
Vergelijking Controlegroep en Behandelgroep Binnen Scholen

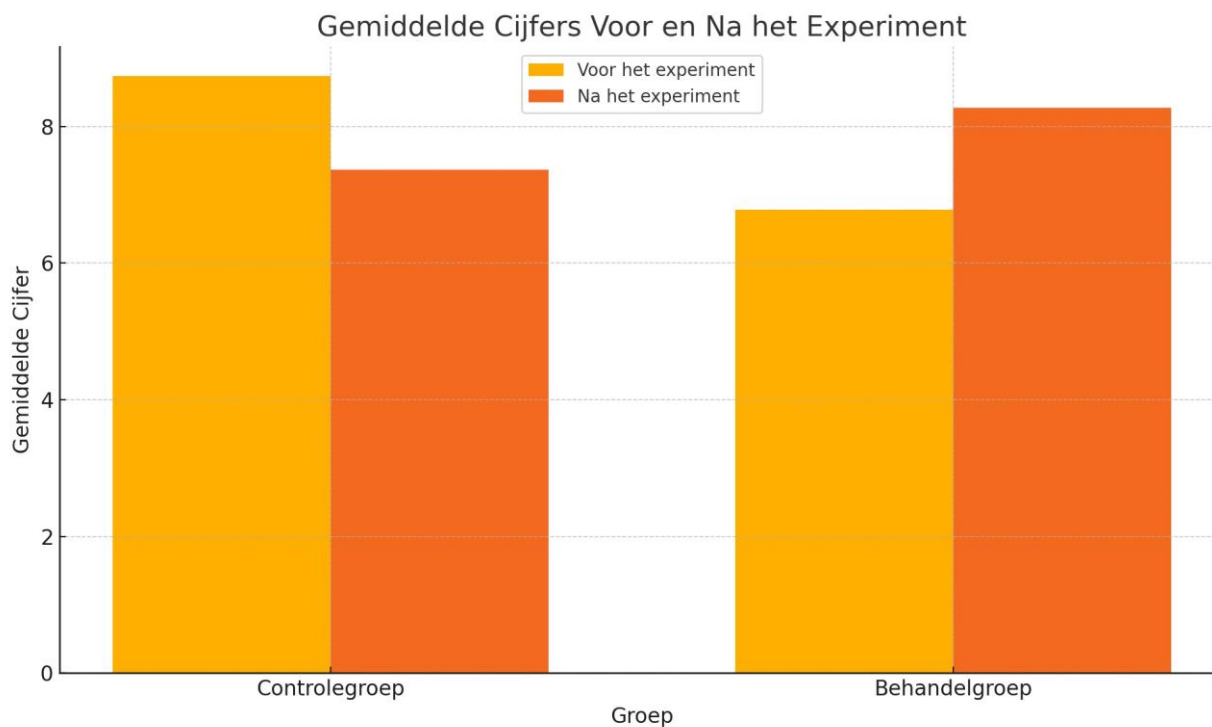


HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLOPRESTATIES

Tot slot is er een analyse uitgevoerd naar de score van de eindtoets ten opzichte van het gemiddelde cijfer van de leerling voor het vak rekenen. In Figuur 4 is een staafdiagram te zien met het gemiddelde cijfer van de leerling voor het vak rekenen voor deelname aan het experiment en het eindresultaat dat behaald is tijdens het experiment.

Figuur 4

Gemiddelde cijfer rekenen in vergelijking met gemiddelde eindcijfer experiment



HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

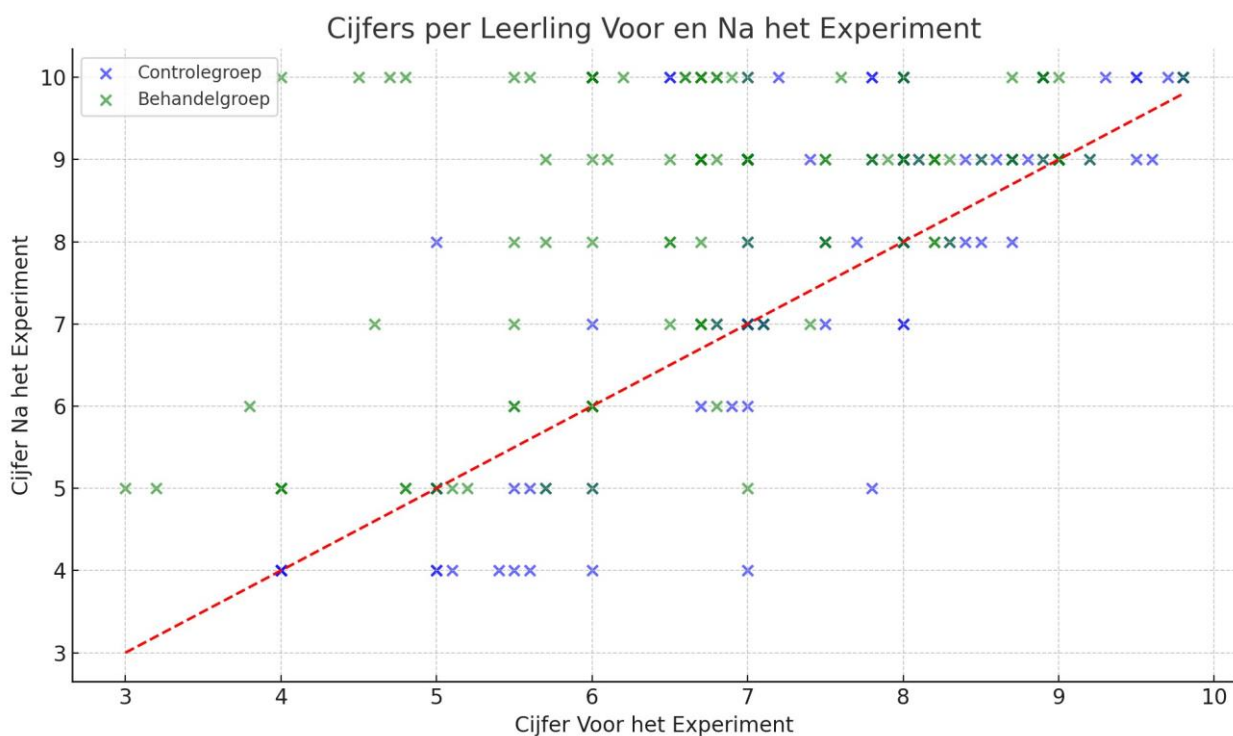
Om het verschil tussen het gemiddelde cijfer en het eindcijfer van het experiment te visualiseren, is er een scatterplot gemaakt (zie figuur 5). Deze scatterplot weergeeft het gemiddelde cijfer van elke leerling en het behaalde eindcijfer, verdeeld over de controlegroep en de behandelgroep.

De X-as van de scatterplot representeert de gemiddelde score van elke leerling vóór het experiment, terwijl de Y-as de gemiddelde eindscore na het experiment weergeeft. De blauwe punten weergeven de controlegroep en de groene punten weergeven de behandelgroep. De rode stippellijn, geeft aan waar de punten zouden liggen als de gemiddelde cijfers gelijk zijn aan de eindcijfers van de leerlingen. Punten boven de rode lijn geven een verbetering in de scores weer, terwijl punten onder de rode lijn een verslechtering van de scores weergeven.

De blauwe punten, die de controlegroep vertegenwoordigen, zijn verspreid rondom de rode lijn. Dit geeft aan dat sommige leerlingen hun scores verbeteren en sommige leerlingen hun scores verlaagden. De controlegroep toont variabiliteit zonder duidelijk trend.

De groene punten, die de behandelgroep vertegenwoordigen, vertonen een vergelijkbare spreiding, maar er lijkt een grotere concentratie van punten zich boven de rode lijn te bevinden. Het is mogelijk dat dit erop wijst dat het geven van formatieve feedback een positief effect heeft gehad op de eindscores.

Figuur 5
Scatterplot Gemiddelde Scores vs Toetsscores Eindtoets



Discussie

Uit de resultaten is gebleken dat formatieve feedback een positieve invloed kan hebben op de leerprestaties van basisschoolkinderen in Nederland. Oorzaken voor deze resultaten kunnen worden gevonden in de aard van formatieve feedback. Formatieve feedback biedt een meer constructieve en ontwikkelingsgerichte benadering dan de cijfermatige beoordeling. Formatieve feedback helpt kinderen begrip te krijgen voor waar hun talenten liggen en op welk vlak zij zich nog kunnen ontwikkelen, zonder de druk en angst te ervaren die vaak gepaard gaat met de cijfermatige beoordeling. Deze bevindingen komen overeen met de bevindingen uit het onderzoek van Butler en Nisan (1986), waarin een soortgelijk onderzoek werd uitgevoerd met Israëliëse kinderen tussen de 10 en 13 jaar.

De resultaten zijn veelbelovend hoewel er rekening moet worden gehouden met enkele beperkingen van het onderzoek. Het onderzoek is gespecificeerd tot een leeftijdsgroep van zes tot en met acht jaar. Het is niet met zekerheid te zeggen of de resultaten generaliseerbaar zijn voor andere leeftijdsgroepen. Daarnaast heeft het onderzoek zich beperkt tot enkel scholen in Nederland. Dit roept de vraag op of de resultaten generaliseerbaar zijn voor andere landen. Ook valt niet met zekerheid te zeggen of de resultaten zijn beïnvloed door het spillover-effect.

Een overweging in het onderzoek is de mate waarin docenten enthousiast waren over de toegewezen beoordelingsmethode. Zo lieten enkele docenten uit de controlegroep blijken niet enthousiast te zijn over de cijfermatige beoordelingsmethode.

In tegenstelling tot de docenten uit de controlegroep waren de docenten in de behandelgroep erg enthousiast over de beoordelingsmethode die aan hen toegewezen is: “De kinderen waren erg enthousiast en trots na het ontvangen van het parel en puzzel formulier. De kinderen mochten het formulier mee naar huis nemen en konden het aan de ouders laten zien. Ook creëerde deze methode cohesie binnen de groep. De kinderen waren liever voor elkaar. Helaas kost de methode wel veel tijd”.

In meerdere gesprekken met docenten kwam het tijdrovende aspect van de parel en puzzel methode naar voren. Dit kan implementatie problemen opleveren, vooral gezien de hoge werkdruk die leraren nu al ervaren. Het is essentieel om naar een manier te zoeken om de methode minder tijdrovend te maken zonder dat dit ten koste gaat van de voordelen van de parel en puzzelmethode. Wellicht kan er een oplossing worden gevonden in digitale tools of apps om de administratieve en feedbackprocessen te versnellen.

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

Het enthousiasme van de leraren, hoofdzakelijk voor de formatieve beoordelingscriteria, zou een experimenteel effect kunnen introduceren waarbij de prestaties van de leerlingen mogelijk zijn beïnvloed door de extra aandacht en feedback die zij indirect hebben ontvangen.

Het is lastig te bepalen of het positieve effect dat de formatieve beoordelingscriteria met zich meebrengt zal aanhouden. Wellicht raken kinderen minder gemotiveerd wanneer zij gewend zijn aan de complimenten die zij op dagelijkse basis ontvangen. Een suggestie zal zijn om deze manier van beoordelen alleen toe te passen enkele weken voordat er een belangrijke toetsperiode aankomt, zoals de citotoetsen. Zo kunnen kinderen met een boost van hun zelfvertrouwen de citotoetsen ingaan.

Dit onderzoek levert interessante bevindingen die als basis kunnen dienen voor verder onderzoek. Een potentiële onderzoeksrichting is bijvoorbeeld de effectiviteit van formatieve feedback wanneer er sprake is van verschillende prestatieniveaus van kinderen die ver van het gemiddelde afzitten. Het is interessant om te onderzoeken of formatieve feedback ook effectief is voor kinderen waarbij een leerstoornis is geconstateerd, zoals dyscalculie. Enkele docenten, die aan dit onderzoek deelnamen, hebben geconstateerd dat leerlingen die normaal laag scoren, nu aanzienlijk hoger scoorden dan normaal. Dit roept opnieuw de vraag op of het effect van formatieve feedback op lange termijn stand zal houden. Vervolgonderzoek is dus noodzakelijk om de langetermijneffecten van formatieve feedback te onderzoeken.

Toekomstig onderzoek zou zich kunnen richten op de langetermijneffecten van dit experiment. Hiermee kan worden bepaald of de voordelen van de formatieve beoordelingscriteria blijvend zijn. Mocht het effect op lange termijn standhouden, zijn eventuele beleidsimplicaties mogelijk. Daarnaast zal onderzoek zich kunnen uitbreiden naar andere leeftijdsgroepen en culturele contexten.

Conclusie

In dit onderzoek wordt de invloed van beoordelingscriteria (summatief versus formatief) op de academische prestaties van basisschoolkinderen van zes tot en met acht jaar in Nederland onderzocht. Hiervoor wordt een kwantitatief onderzoek uitgevoerd waarbij er data van de prestaties van leerlingen op de basisschool in Nederland worden verzameld door middel van een experiment.

Uit de resultaten blijkt dat de groep die beoordeeld wordt aan de hand van formatieve feedback, in de vorm van het parel en puzzel formulier, beter presteert dan de groep die beoordeeld wordt aan de hand van cijfers. De groep die beoordeeld wordt aan de hand van formatieve feedback scoort gemiddeld 0.87 punt hoger op de eindtoets dan de groep die beoordeeld wordt aan de hand van cijfers.

Daarnaast blijkt uit het experiment dat de leerlingen die formatieve feedback krijgen meer vertrouwen hebben in het behalen van een goed resultaat voor de toets dan de leerlingen die worden beoordeeld aan de hand van cijfers.

De hypothese: De beoordelingscriteria (summatief versus formatief) hebben een significant effect op de academische prestaties van basisschoolleerlingen in Nederland, in de leeftijdscategorie zes tot en met acht jaar kan worden bevestigd aan de hand van de resultaten. De groep die beoordeeld wordt aan de hand van formatieve feedback scoort 0.87 punt hoger op de eindtoets dan de groep die beoordeeld is aan de hand van cijfers. Procentueel gezien halen meer leerlingen een voldoende wanneer zij voorafgaand aan de toets formatieve feedback hebben ontvangen (82%) in vergelijking met de leerlingen die cijfers hebben ontvangen (75%).

In het onderzoek wordt onderscheid gemaakt tussen mannen en vrouwen. Echter, uit de resultaten blijkt dat het geslacht van de leerling geen significant effect heeft op de toetsscore. Ook is er geen interactie-effect gevonden tussen de groep en het geslacht.

Referenties

- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>.
- Butler, R. A., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal Of Educational Psychology*, 78(3), 210–216. <https://doi.org/10.1037/0022-0663.78.3.210>
- De Groot, A. D. (1972). *Vijven en zessen: cijfers en beslissingen : het selectieproces in ons onderwijs*. <https://lib.ugent.be/nl/catalog/rug01:000023856>
- Dumont, H., Istance, D. & Benavides, F. (2010). The nature of learning. (2010). In *Educational research and innovation*. <https://doi.org/10.1787/9789264086487-en>
- Emons, W., Van Der Scheer, E., Hemker, B., Veldhuizen, J., & Stichting Cito, Onderzoek, Kennis & Innovatie. (2021). Een verkenning naar toetsing in het voortgezet onderwijs. In *Stichting Cito | Onderzoek, Kennis & Innovatie*. https://cito.nl/media/f15atcw4/2021_cito Onderzoeksrapport cijfers-cijfers-cijfers.pdf
- Gipps, C. (2002). Beyond testing. In *Routledge eBooks*. <https://doi.org/10.4324/9780203486009>
- Gordon, M. E., & Fay, C. H. (2010). The Effects of Grading and Teaching Practices on Students' Perceptions of Grading Fairness. *College Teaching*, 58(3), 93–98. <https://doi.org/10.1080/87567550903418586>
- Haimovitz, K., & Dweck, C. S. (2017). The Origins of Children's Growth and Fixed Mindsets: New Research and a New Proposal. *Child Development*, 88(6), 1849–1859. <https://doi.org/10.1111/cdev.12955>
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research*, 66(2), 99-136.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112
- Moneva, J. (2019). Students' level of self-confidence and performance tasks. *Asia Pacific Journal Of Academic Research in Social Sciences*, 5(1), 42-48
- Peeters, W. (2023, 31 oktober). *Wat zegt een cijfer?* Geraadpleegd op 30 april 2023, van <https://www.vernieuwewonderwijs.nl/wat-zegt-een-cijfer/?print=pdf>
- Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153- 189. <https://doi.org/10.3102/0034654307313795>
- Steffens, W. Hoe leren kinderen?. *kinderopvang* 32, 12–17 (2022). <https://doi.org/10.1007/s41189-021-1172-7>

HET EFFECT VAN FORMATIEVE FEEDBACK OP BASISCHOOLPRESTATIES

- Vansteenkiste, M., Sierens, E., Soenens, B., Lens, W., & Katholieke Universiteit Leuven. (2007). Willen, moeten en structuur in de klas : over het stimuleren van een optimaal leerproces. *ONDERWIJSKUNDIG & MOTIVATIE* (Vol. 1, Nummer 16, pp. 37–38).
- Visscher, A., Peters, M., & Staman, L. (2010). Het FOCUS-project: Opbrengstgericht werken op basis van prestatie-feedback. *Panama-post*, 29(4), 55-60.
- Visscher, A., Ehren, M., & Vakgroep Onderwijsorganisatie en -management, Universiteit Twente. (2011). De eenvoud en complexiteit van Opbrengstgericht Werken. In Kenniskamer van het Ministerie van Onderwijs, Cultuur en Wetenschap, *Kenniskamer van het Ministerie van Onderwijs, Cultuur en Wetenschap*.
- Von Bergen, C. W., Jr., Bressler, M. S., Campbell, K., & Southeastern Oklahoma State University. (2014). The sandwich feedback method: Not very tasty. In *Journal of Behavioral Studies in Business* (Vol. 7, pp. 1–3). <https://aabri.com/manuscripts/141831.pdf>

Bijlage A Parel en Puzzel Formulier



PARELS



PARELS



PUZZELS



Bijlage B Contact Scholen

Bijlage B1 Verzoek deelnamen experiment scholen

Onderwerp: Verzoek om experiment uit te voeren in groep 4 en 5

Geachte *[Naam van de directeur of contactpersoon]*,

Ik zou mijzelf even voorstellen; mijn naam is Louise Stoop, ik ben 21 jaar en ik studeer aan de Erasmus Universiteit Rotterdam. Momenteel schrijf ik mijn scriptie en doe ik onderzoek naar het effect van verschillende beoordelingscriteria op de academische prestaties van basisschoolkinderen in Nederland in groep 4 en groep 5. Met deze mail vraag ik u toestemming om een educatief experiment uit te voeren in de groepen 4 en 5 op uw basisschool.

Doel en opzet experiment

Het experiment richt zich op de impact van de formatieve beoordelingscriteria op de academische prestaties van leerlingen. Uit internationaal onderzoek is gebleken dat formatieve feedback een significant positieve invloed heeft op leerresultaten. Hoewel deze studies waardevol zijn, is er nog behoefte aan vergelijkbaar onderzoek binnen de Nederlandse context en specifiek voor basisschoolleerlingen in groep 4 en 5.

Volgens de Wet Primair Onderwijs (WPO) heeft het basisonderwijs als doel om kinderen een solide basis te bieden in hun leer- en ontwikkelingsproces. Uit onderzoek blijkt echter dat summatieve beoordelingen in de vorm van cijfers een negatief effect hebben op dit proces, omdat leren vaak stopt na het ontvangen van een cijfer. Bovendien kan het geven van cijfers het zelfvertrouwen van leerlingen schaden, wat leidt tot slechtere prestaties en een "fixed mindset" – het geloof dat capaciteiten vaststaan en extra inspanning zinloos is.

De Parel en Puzzel Methode

Om het leerproces te optimaliseren, is het belangrijk om naar alternatieve methoden te kijken, zoals de "parel en puzzel" methode. Deze methode maakt gebruik van formatieve feedback en richt zich op het verbeteren van de prestaties van leerlingen door hun sterke punten ("parels") te benadrukken en hen te motiveren. Dit helpt bij de overgang van een "fixed mindset" naar een "growth mindset", wat kan leiden tot betere prestaties en verbeterde vaardigheden zoals kritisch denken en samenwerken.

De "puzzels" in deze methode symboliseren de uitdagingen en verbeterpunten van de leerlingen. Leraren werken samen met de leerlingen om deze puzzels op te lossen en geven constructieve feedback, waarbij negatieve kritiek wordt vermeden. Onderzoek toont aan dat het uitspreken van vertrouwen in de verbetering van de leerling de grootste impact heeft op hun ontwikkeling.

Uitvoering en periode

Voor mijn experiment zou ik graag twee groepen 4 en 5 gedurende twee weken, van 21 mei tot en met 4 juni, volgen tijdens de rekenlessen. De twee groepen worden verdeeld en krijgen beide een verschillende behandeling.

- Groep 1 ontvangt na elke rekenles een cijfer
- Groep 2 ontvangt na elke rekenles formatieve feedback aan de hand van de parel en puzzel methode.

Na twee weken wordt de methodetoets afgenomen. De resultaten van de methodetoetsen zal ik verzamelen om het effect van de interventie te kunnen meten. Het experiment zal zorgvuldig worden ontworpen om minimale verstoring van het reguliere onderwijsproces te garanderen en de voortgang van de leerlingen nauwlettend te volgen.

Ik ben van plan om de resultaten te delen met uw school, wat mogelijk waardevolle inzichten kan bieden voor de verdere ontwikkeling van uw beoordelingsmethoden. Uiteraard ben ik bereid om alle details van het experiment te bespreken en eventuele zorgen die u heeft te adresseren.

Ik kijk uit naar de mogelijkheid om samen te werken en een positieve bijdrage te leveren aan de academische ontwikkeling van uw leerlingen. Mocht u nog vragen hebben neem dan gerust contact op. Bij voorbaat dank voor uw overweging en ik hoop spoedig van u te horen.

Met vriendelijke groet,

Louise Stoop



Louise Stoop



06-30134436



Louises2003@gmail.com



Erasmus Universiteit Rotterdam



Bijlage B2 Mail Docenten Controlegroep

Instructies experiment 21 mei tot en met 4 juni

Beste *naam onderwijzer*,

Nogmaals hartelijk dank dat ik dit experiment in uw klas mag uitvoeren. Na aanleiding van ons gesprek zal ik de onderzoeksopzet nogmaals uiteenzetten.

Startdatum en opzet

Op dinsdag 21 mei begint het experiment. De rekenlessen worden gegeven zoals u dit altijd doet, zonder aanpassingen in de lesinhoud of lesstijl.

Beoordelingsmethode

De gemaakte opdrachten worden door u nagekeken en beoordeeld met cijfers. De richtlijnen voor de beoordeling zijn als volgt:

- De foutmarge: 80% goed is een 6.
- De berekening van het cijfer : $(\text{Aantal goed} / \text{Totaal aantal opgaven}) \times 9 + 1$

*Wanneer een leerling 15 vragen van de 20 correct heeft beantwoord, dan is het cijfer $(15/20) \times 9 + 1 = 7.8$

Het cijfer schrijft u onderaan in het schrift waarin de opdrachten zijn gemaakt. Het is van belang voor het onderzoek dat de leerling dit cijfer te zien krijgt.

Verloop experiment

Dit proces herhaalt u na elke rekenles. Ofwel na elke rekenles ontvangt de leerling het behaalde cijfer voor de gemaakte opdrachten. Op 4 juni neemt u de methodetoets rekenen af. Graag ontvang ik van u de geanonimiseerde toets resultaten van de leerlingen. Ik u de volgende gegevens in te vullen in het formulier dat is meegestuurd in de bijlage. Als u vragen heeft of verdere verduidelijking nodig heeft, aarzel dan niet om contact met mij op te nemen.

Met vriendelijke groet,
Louise Stoop

Bijlage B3 Mail aan Docenten Behandelgroep

Instructies experiment 21 mei tot en met 4 juni

Beste *naam onderwijzer*,

Nogmaals hartelijk dank dat ik dit experiment in uw klas mag uitvoeren. Na aanleiding van ons gesprek zal ik de onderzoeksopzet nogmaals uiteenzetten.

Startdatum en opzet

Op dinsdag 21 mei begint het experiment. De rekenlessen worden gegeven zoals u dit altijd doet, zonder aanpassingen in de lesinhoud of lesstijl.

Beoordelingsmethode

De gemaakte opdrachten worden door u nagekeken en beoordeeld na aanleiding van de parel en puzzel methode. De “parel en puzzel methode” is een vorm van formatieve feedback. De methode is gericht op het geven van constructieve feedback. Hierbij worden de sterke punten (de parels) van de leerling benadrukt en wordt er hulp geboden bij het overwinnen van uitdagingen (puzzels). In de bijlage vind u een parel en puzzel formulier. Dit formulier dient u in te vullen op basis van de gemaakte opdrachten. Bij het invullen van het formulier dient u rekening te houden met de volgende beoordelingseisen:

Parels:

- Identificeer de sterke punten van de leerling
- Wees specifiek en gedetailleerd, geef concrete voorbeelden
- Wees motiverend

Puzzels:

- Identificeer de verbeterpunten van de leerling
- Concrete suggesties, geef specifiek en haalbare suggesties van verbetering
- Focus op groei, uit vertrouwen in de leerling dat de leerling capaciteit heeft om zijn prestaties te verbeteren

Voorbeeld:

Parel:

"Geweldig werk op de optelsommen! Je hebt 9 van de 10 optelsommen correct beantwoord. Ik zie dat je heel goed begrijpt hoe je getallen moet optellen, vooral bij sommen zoals $34 + 28$. Het is duidelijk dat je hard hebt gewerkt en je begrijpt deze vaardigheid uitstekend. Blijf zo doorgaan!"

Puzzel:

"Bij de aftreksommen heb je 6 van de 10 opgaven correct beantwoord. Ik merk dat je soms moeite hebt met het aftrekken van grotere getallen, vooral bij sommen zoals $72 - 45$. Een tip om je te helpen: probeer eerst de tientallen van elkaar af te trekken ($70 - 40 = 30$) en dan de eenheden ($2 - 5 = -3$), en combineer dan de resultaten ($30 - 3 = 27$). Dit kan het aftrekken eenvoudiger maken. Laten we samen naar enkele van deze opgaven kijken om te zien hoe je dit kunt verbeteren. Ik weet zeker dat als jij hard blijft oefenen volgende keer nog meer sommen goed zal maken"

Verloop experiment

Dit proces herhaalt u na elke rekenles. Op 4 juni neemt u de methodetoets rekenen af.

Graag ontvang ik van u de geanonimiseerde toets resultaten van de leerlingen. Ik vraag u de volgende gegevens in te vullen in het formulier dat is meegestuurd in de bijlage.

Als u vragen heeft of verdere verduidelijking nodig heeft, aarzel dan niet om contact met mij op te nemen.

Met vriendelijke groet,

Louise Stoop

Bijlage C: Rekenlessen Experiment

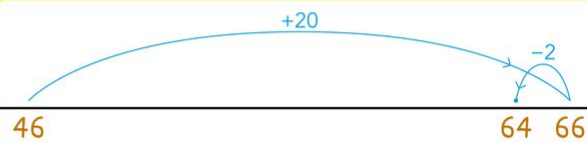
BLOK 9

LES 1

DOEL 1

• Je leert optellen tot en met 100 met de strategie: rijgen met te veel.

HULP



$$46 + 18 = 64$$
$$46 + 20 - 2 = 64$$

1

Reken met rijgen met te veel.

Reken uit op de getallenlijn.



 $54 + 19 = \dots\dots$

 $65 + 28 = \dots\dots$

 $47 + 38 = \dots\dots$

 $33 + 29 = \dots\dots$

2

Bij welke sommen rijg je met te veel?

Reken alleen die uit, op de getallenlijn.



 $37 + 19 = \dots\dots$

 $38 + 26 = \dots\dots$

 $63 + 29 = \dots\dots$

 $55 + 44 = \dots\dots$



3

Welke som hoort erbij? Reken uit op de getallenlijn.



Anne heeft 58 euro.
Zij krijgt 39 euro op haar verjaardag.
Hoeveel euro heeft zij nu?

Milan heeft 45 kilometer gereden.
Hij moet er nog 19.
Hoeveel heeft hij dan gereden?

som:

som:

antwoord: euro

antwoord: kilometer



4

Reken uit.

Kruis de sommen aan die je uitrekent met rijgen met te veel.



$28 + 28 = \dots\dots$

$67 + 29 = \dots\dots$

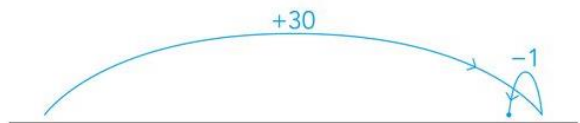
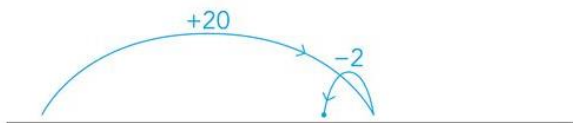
$48 + 47 = \dots\dots$

$81 + 14 = \dots\dots$



5

Maak sommen bij de getallenlijnen.



..... + = en + =

..... + = en + =



KIJK TERUG



Bedenk 2 plussommen waarbij rijgen met te veel een handige strategie is.

..... + =

..... + =



3

Reken uit op de getallenlijn.



$35 + 18 = \dots\dots$

$56 + 29 = \dots\dots$

$48 + 19 = \dots\dots$

$53 + 28 = \dots\dots$

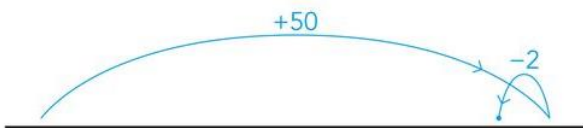
$62 + 29 = \dots\dots$

$72 + 19 = \dots\dots$

4

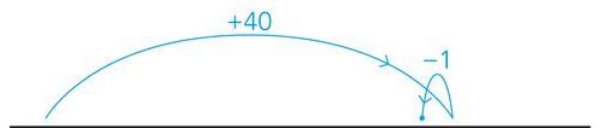
Welke som?

Schrijf ook het antwoord op.



35

som:



56

som:

5

Reken met rijgen met te veel.



$34 + 18 + 18 = \dots\dots\dots$

$45 + 29 + 38 = \dots\dots\dots$

$56 + 18 + 19 = \dots\dots\dots$

$24 + 39 + 29 = \dots\dots\dots$

KIJK TERUG



Bij $45 + 38$ maak je een sprong van 40 en dan 2 eraf.

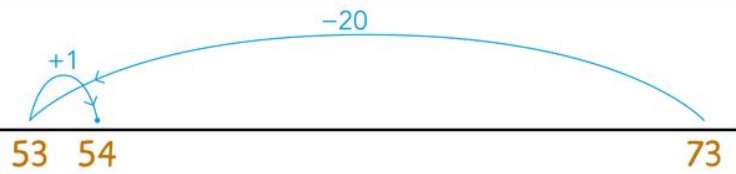
waar | niet waar



DOEL 2

• Je leert aftrekken tot en met 100 met de strategie: rijgen met te veel.

HULP



$$73 - 19 = 54$$

$$73 - 20 + 1 = 54$$

1

Reken met rijgen met te veel. Reken uit op de getallenlijn.



$$93 - 19 = \dots\dots\dots$$

$$52 - 28 = \dots\dots\dots$$

$$64 - 38 = \dots\dots\dots$$

$$86 - 29 = \dots\dots\dots$$

2

Bij welke sommen reken je met rijgen met te veel?

Reken alleen die sommen uit, op de getallenlijn.



$$77 - 19 = \dots\dots\dots$$

$$68 - 29 = \dots\dots\dots$$

$$43 - 16 = \dots\dots\dots$$

$$65 - 39 = \dots\dots\dots$$



3

Welke som hoort erbij? Reken uit op de getallenlijn.



Boris heeft 42 euro.
Hij koopt een fietsbel van 19 euro.
Hoeveel euro heeft hij nog over?

Juf Ilse heeft 64 schriften.
Alle 28 kinderen krijgen een schrift.
Hoeveel schriften houdt juf Ilse over?

som: _____

som: _____

antwoord: _____ euro

antwoord: _____ schriften



4

Reken uit.

Kruis de sommen aan die je uitrekent met rijgen met te veel.



$93 - 28 = \dots\dots$

$54 - 29 = \dots\dots$

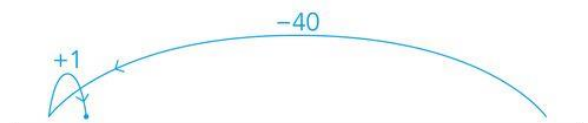
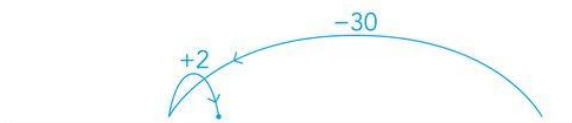
$76 - 38 = \dots\dots$

$88 - 47 = \dots\dots$



5

Maak sommen bij de getallenlijnen.



_____ - _____ = _____ en _____ - _____ = _____

_____ - _____ = _____ en _____ - _____ = _____



KIJK TERUG

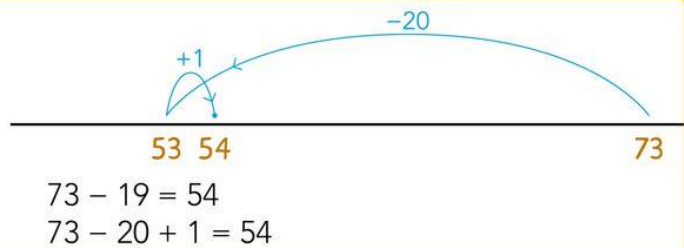


Bij $83 - 38$ maak je een sprong van 40 terug en dan 2 erbij. waar | niet waar

DOEL 2

- Je leert aftrekken tot en met 100 met de strategie: rijgen met te veel.
- Je begrijpt de strategie: rijgen met te veel.

HULP



1

Welke som hoort erbij?

Reken uit.

som:

antwoord: euro



2

Welke som hoort erbij? Reken uit op de getallenlijn.



Ik heb 57 boeken.

Kevin heeft 57 boeken.
Zijn broertje heeft er 29 minder.
Hoeveel boeken heeft zijn broertje?

som:
antwoord: boeken



Mijn buurvrouw is 54 jaar.
19 jaar geleden kwam ze naast ons wonen.
Hoe oud was ze toen?

som:
antwoord: jaar



De pauze duurt 45 minuten.
Er zijn al 18 minuten voorbij.
Hoelang kunnen we nog buiten spelen?

som:
antwoord: minuten



3

Reken uit op de getallenlijn.



$67 - 29 = \dots\dots$

$73 - 38 = \dots\dots$

$82 - 48 = \dots\dots$

$46 - 19 = \dots\dots$

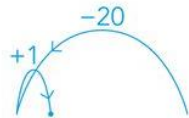


4

Welke som?

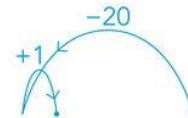
Schrijf ook het antwoord op.





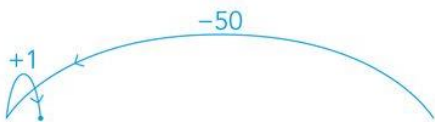
67

som:



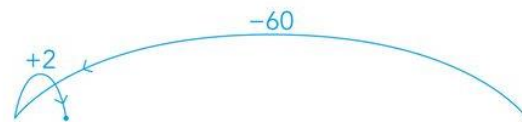
38

som:



84

som:



93

som:



5

Wat zit er onder de vlek?



$78 - 39 = \text{vlek}$

$98 - \text{vlek} = 49$

$\text{vlek} - 69 = 16$

$66 - \text{vlek} = 38$

$87 - 48 = \text{vlek}$

$\text{vlek} - 38 = 36$

$\text{vlek} - 38 = 8$

$\text{vlek} - 58 = 35$

$53 - \text{vlek} = 15$



KIJK TERUG



Bedenk 2 minsommen waarbij je moet rekenen met rijgen met te veel.

.....

.....

DOEL 1

1

Reken uit op de getallenlijn.



$38 + 19 = \dots\dots$

$65 + 28 = \dots\dots$

$56 + 28 = \dots\dots$

$53 + 39 = \dots\dots$

2

Welke som hoort erbij? Reken uit op de getallenlijn.



Kim koopt 65 gram drop. Eva koopt 18 gram meer. Hoeveel gram drop koopt Eva?

som:

antwoord: gram

3

Welke som hoort erbij? Hoeveel moeten ze betalen?



PRIJSLIJST

per week	
gezin	55 euro
gezin + huisdier	65 euro
tent	28 euro
caravan	39 euro



som:

antwoord: euro

TUSSENSTAND





DOEL 2

1

Reken uit op de getallenlijn.



48 - 19 =

83 - 39 =

72 - 48 =

92 - 69 =

2

Welke som hoort erbij? Reken uit op de getallenlijn.



Bij Kopi kost een printer 83 euro.
Bij Jopi kost hij 29 euro minder.
Hoeveel kost de printer bij Jopi?

som:

antwoord: euro

3

Welke som hoort erbij? Hoeveel geld blijft er over?



PRIJSLIJST

oorbellen	19 euro
armband	18 euro
ketting	28 euro
ring	29 euro



Ik heb 50 euro.
Ik koop oorbellen.



Ik heb 42 euro.
Ik koop een ketting
voor mijn moeder.

som:

antwoord: euro

som:

antwoord: euro

TUSSENSTAND



DOEL 3

• Je leert sommen als $42 : 3$ uitrekenen met de basisstrategie: splitsen.

HULP



$$\begin{array}{r} 42 : 3 = ? \\ \swarrow \quad \searrow \\ 30 \quad 12 \end{array}$$

stap 1: Meer dan $10 \times$? Ja \rightarrow splitsen
 stap 2: Welke splitsing? Eerst 10×3 eraf halen, nog 12 over om te delen.
 stap 3: hulpsommen: $30 : 3 = 10$ en $12 : 3 = 4 \rightarrow 42 : 3 = 14$

1



Welke som hoort erbij?

Reken uit met splitsen.



Er gaan 3 tennisballen in een blik.
 Hoeveel blikken kun je vullen?

$$\begin{array}{r} 48 : 3 = 16 \\ \swarrow \quad \searrow \end{array}$$

hulpsommen:

antwoord:

2



Reken uit.

Schrijf op zoals bij de Hulp.

$$\begin{array}{r} 72 : 4 = \dots\dots\dots \\ \swarrow \quad \searrow \end{array}$$

hulpsommen:

$$\begin{array}{r} 48 : 4 = \dots\dots\dots \\ \swarrow \quad \searrow \end{array}$$

hulpsommen:

$$\begin{array}{r} 65 : 5 = \dots\dots\dots \\ \swarrow \quad \searrow \end{array}$$

hulpsommen:

$$\begin{array}{r} 96 : 8 = \dots\dots\dots \\ \swarrow \quad \searrow \end{array}$$

hulpsommen:



3

Welke som hoort erbij?



Vul de tabel in.

Er gaan 7 appels in een zak.
Hoeveel zakken kun je vullen?



aantal appels	deelsom	hulpsommen	aantal zakken
77 appels	77 : 7	70 : 7 = 10 en 7 : 7 = 1	
84 appels			
91 appels			
105 appels			
119 appels			



4

Reken uit met splitsen.



72 : 6 =	56 : 4 =	80 : 5 =
.....
108 : 9 =	85 : 5 =	39 : 3 =
.....
84 : 6 =	104 : 8 =	108 : 6 =
.....
68 : 4 =	136 : 8 =	135 : 9 =
.....



5

Wat staat er onder de vlek?



 : 3 = 12	 : 4 = 15	 : 5 = 19	 : 7 = 12
 : 6 = 17	 : 8 = 15	 : 7 = 18	 : 9 = 16



KIJK TERUG



56 : 4 =

Hoe reken je dat uit?

.....

DOEL 3

• Je leert sommen als $42 : 3$ uitrekenen met de basisstrategie: splitsen.

HULP



$$\begin{array}{r} 42 : 3 = ? \\ \underline{30} \\ 12 \end{array}$$

stap 1: Meer dan $10 \times$? Ja → splitsen
 stap 2: Welke splitsing? Eerst 10×3 eraf halen, nog 12 over om te delen.
 stap 3: hulpsommen: $30 : 3 = 10$ en $12 : 3 = 4$ → $42 : 3 = 14$

1

Kruis de sommen aan die je uitrekent met splitsen.



Je hoeft de sommen niet uit te rekenen.

- | | | | |
|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <input type="checkbox"/> $55 : 5 =$ | <input type="checkbox"/> $56 : 8 =$ | <input type="checkbox"/> $91 : 7 =$ | <input type="checkbox"/> $32 : 8 =$ |
| <input type="checkbox"/> $90 : 5 =$ | <input type="checkbox"/> $96 : 8 =$ | <input type="checkbox"/> $24 : 3 =$ | <input type="checkbox"/> $72 : 4 =$ |
| <input type="checkbox"/> $36 : 6 =$ | <input type="checkbox"/> $48 : 3 =$ | <input type="checkbox"/> $44 : 4 =$ | <input type="checkbox"/> $28 : 7 =$ |

2

Vul de splitsing in en reken uit.



- | | | | |
|--|---|--|--|
| $102 : 6 =$
$\underline{}$
..... | $76 : 4 =$
$\underline{}$
..... | $120 : 8 =$
$\underline{}$
..... | $117 : 9 =$
$\underline{}$
..... |
|--|---|--|--|

3

Welke som hoort erbij?



Schrijf zelf het splitsdakje en de splitsing erbij en reken uit.



Op elk blad passen 6 foto's. Hoeveel bladen kun je vol plakken?

72 foto's
Hoeveel bladen vol?

som:

.....

antwoord:

90 foto's
Hoeveel bladen vol?

som:

.....

antwoord:

78 foto's
Hoeveel bladen vol?

som:

.....

antwoord:

108 foto's
Hoeveel bladen vol?

som:

.....

antwoord:





4

Splits en reken uit.



Schrijf zelf het splitsdakje en de splitsing erbij.

$96 : 6 =$

$70 : 5 =$

$52 : 4 =$

.....

.....

.....

$112 : 7 =$

$112 : 8 =$

$120 : 8 =$

.....

.....

.....

$54 : 3 =$

$76 : 4 =$

$90 : 5 =$

.....

.....

.....

$84 : 6 =$

$153 : 9 =$

$133 : 7 =$

.....

.....

.....



5

Welke som hoort erbij?



Bedenk de som en reken uit.

Schrijf zelf het splitsdakje en de splitsing erbij.

5 kinderen verdelen 75 stickers.
Hoeveel stickers krijgt ieder?

8 kinderen verdelen 128 stickers.
Hoeveel stickers krijgt ieder?

som:

som:

.....

.....

antwoord:

antwoord:

6 kinderen verdelen 114 stickers.
Hoeveel stickers krijgt ieder?

7 kinderen verdelen 126 stickers.
Hoeveel stickers krijgt ieder?

som:

som:

.....

.....

antwoord:

antwoord:



KIJK TERUG



$48 : 3 =$

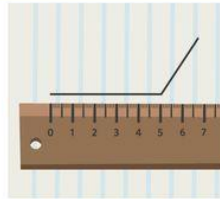
Hoe kun je zien dat je deze som kunt uitrekenen met splitsen?

.....

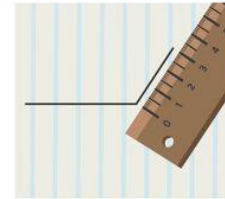
DOEL 4

• Je leert meten met een liniaal.

HULP



Meet het eerste stukje van de lijn.
5 centimeter (5 cm)



Meet het tweede stukje van de lijn.
3 centimeter (3 cm)

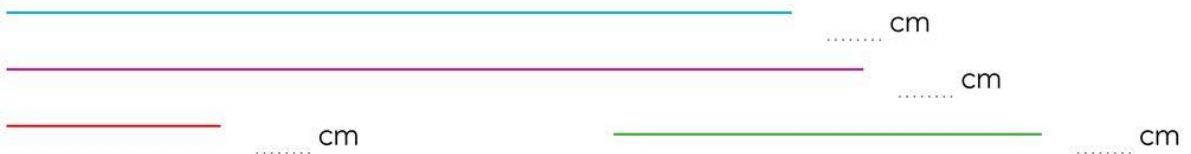
$$5 \text{ cm} + 3 \text{ cm} = 8 \text{ cm}$$



1

Hoe lang?

Meet met je liniaal.



2

Hoe lang?

Meet met je liniaal. Welke lijnen zijn even lang? Zet er een x bij.



3

Hoe lang?

Meet de potloden met je liniaal.





4

Hoe lang is het ongeveer?



Schrijf het voorwerp in het juiste vak.

Kies uit: schaar, schoen, gum, banaan, pink, hoofd.

5 cm

15 cm

25 cm

.....

.....

5

Hoe lang?



Meet met je liniaal. Welke lijnen zijn even lang? Zet er een x bij.



KIJK TERUG



Wat is ongeveer zo lang?



Zoek een voorwerp in de klas met de juiste lengte.

Schrijf het voorwerp in het juiste vak.

5 cm

15 cm

25 cm

.....

.....

DOEL 4

• Je leert meten met een meetlint van 1 meter.

HULP

De lijn begint bij 0.

De lijn stopt bij 95.
De lijn is 95 cm.



100 cm is 1 meter.

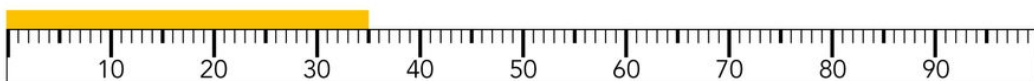


1

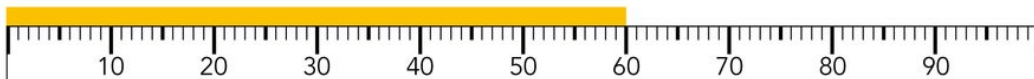
Hoe lang?



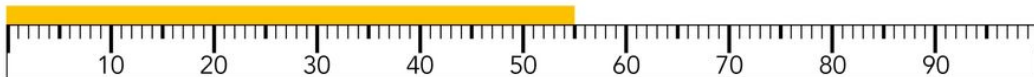
cm



cm



cm



cm

2

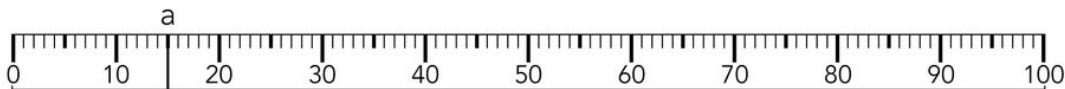
Zet een streep op de liniaal.



Schrijf de letter erbij.



- a bij 15 cm
- b bij 30 cm
- c bij 45 cm
- d bij 62 cm
- e bij 71 cm
- f bij 84 cm
- g bij 92 cm

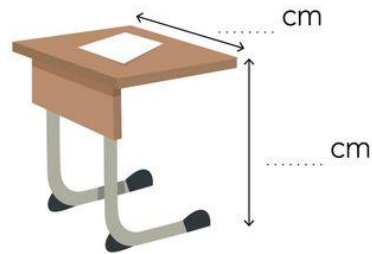
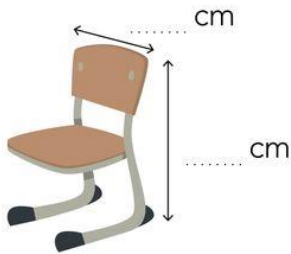




3

Meet je stoel en tafel.

Gebruik een meetlint.



4

Hoe lang?



30 cm

2 meter

15 cm

1 cm

1 meter

5

Teken een lijn.

De hele lijn is 61 cm lang.
Verdeel de lijn in 5 stukken.

KIJK TERUG

Bedenk 4 verschillende meetinstrumenten. Wat kun je er goed mee meten?

Met een kun je meten.

Met een kun je meten.

Met een kun je meten.

Met een kun je meten.

DOEL 3

1

Keer om en reken uit.



$3 \times 8 = 24$

$8 \times 3 = \dots\dots$

$7 \times 4 = \dots\dots$

$7 \times 6 = \dots\dots$

$7 \times 9 = \dots\dots$

$8 \times 9 = \dots\dots$

$8 \times 5 = \dots\dots$

2

Reken uit.



x	5	6	7
5			
8			
9			

x	3	4	9
6			
7			
8			

3

Reken uit.



dozen	1	5	6	7	8
eieren	6				



pakjes	1	4	8	9	10
stiften	8				



bossen	1	2	4	5	10
euro	3				

TUSSENSTAND



DOEL 4

1

Zet een streep op de liniaal.

Schrijf de letter erbij.

a bij 15 cm

b bij 25 cm

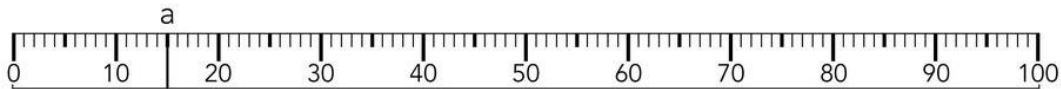
c bij 40 cm

d bij 54 cm

e bij 63 cm

f bij 71 cm

g bij 96 cm

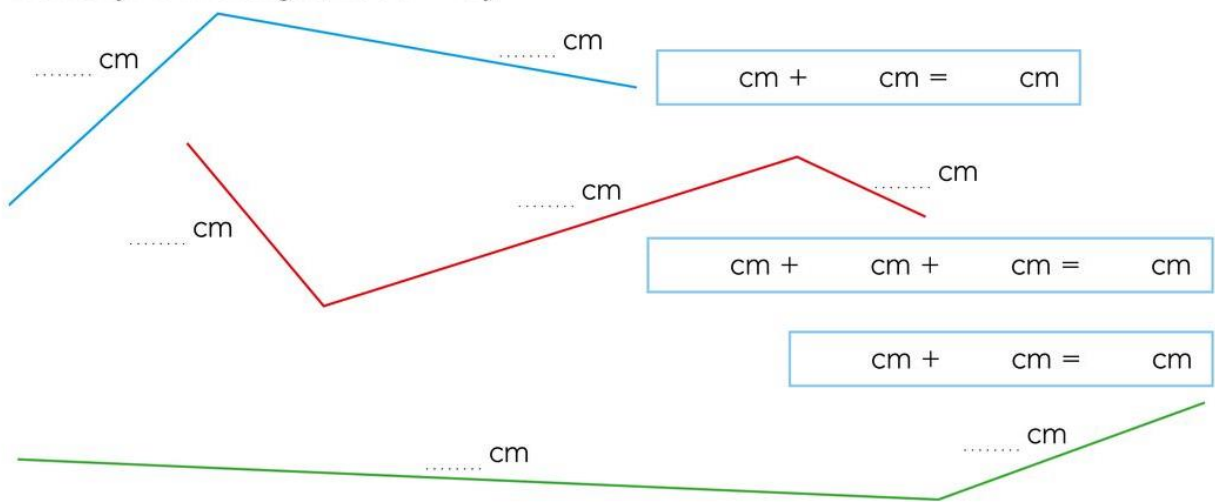


2

Hoe lang?

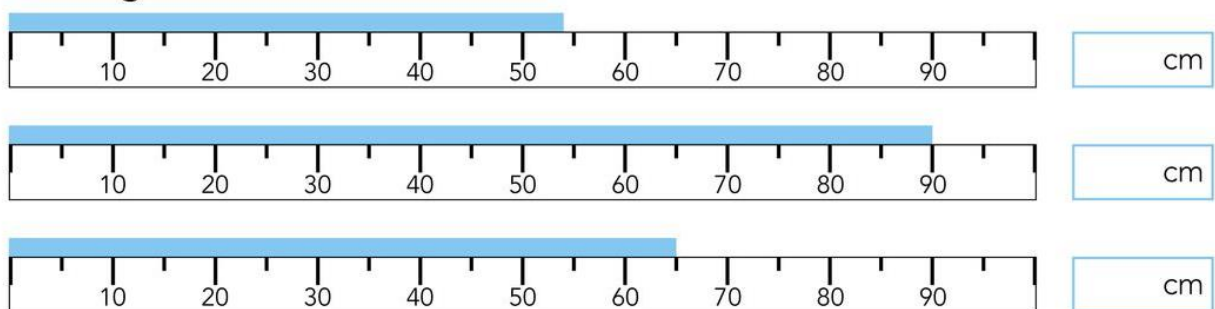
Meet met je liniaal.

Welke lijn is het langst? Zet er x bij.



3

Hoe lang?



TUSSENSTAND



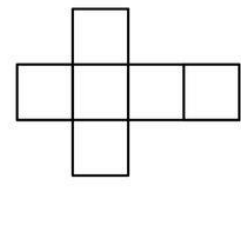
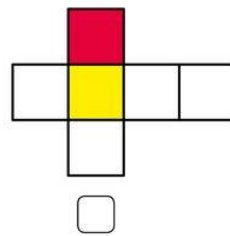
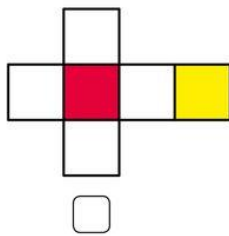
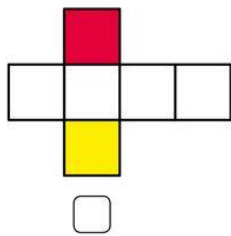
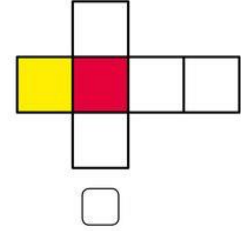
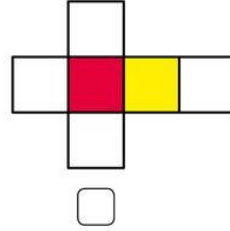
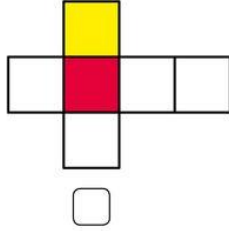
DOEL

- Je leert voorspellen hoe de uitslag van een figuur eruit zal zien.

1

Hoe kan de uitslag eruitzien?

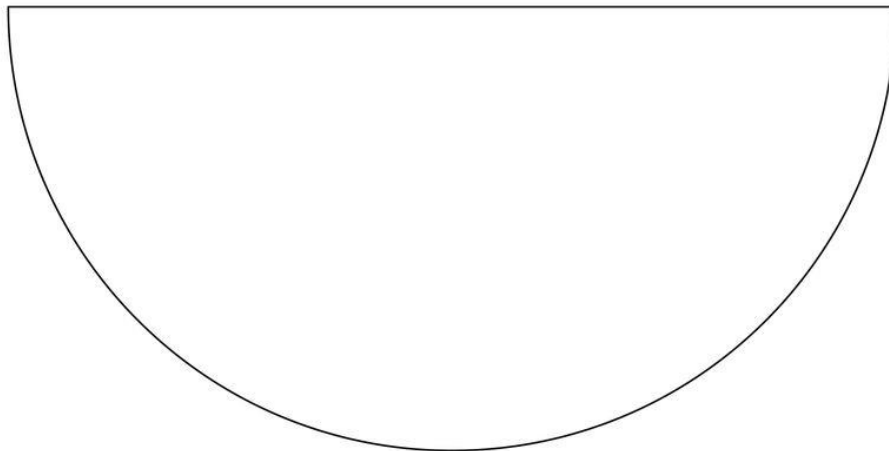
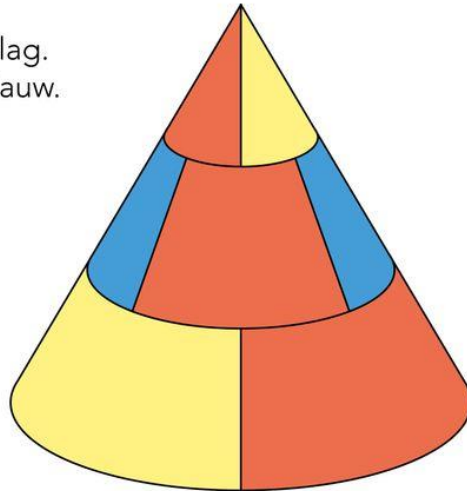
Bedenk zelf kleuren bij de laatste uitslag.



2

Maak de uitslag af.

Teken de lijnen op de uitslag.
Kleur met rood, geel of blauw.



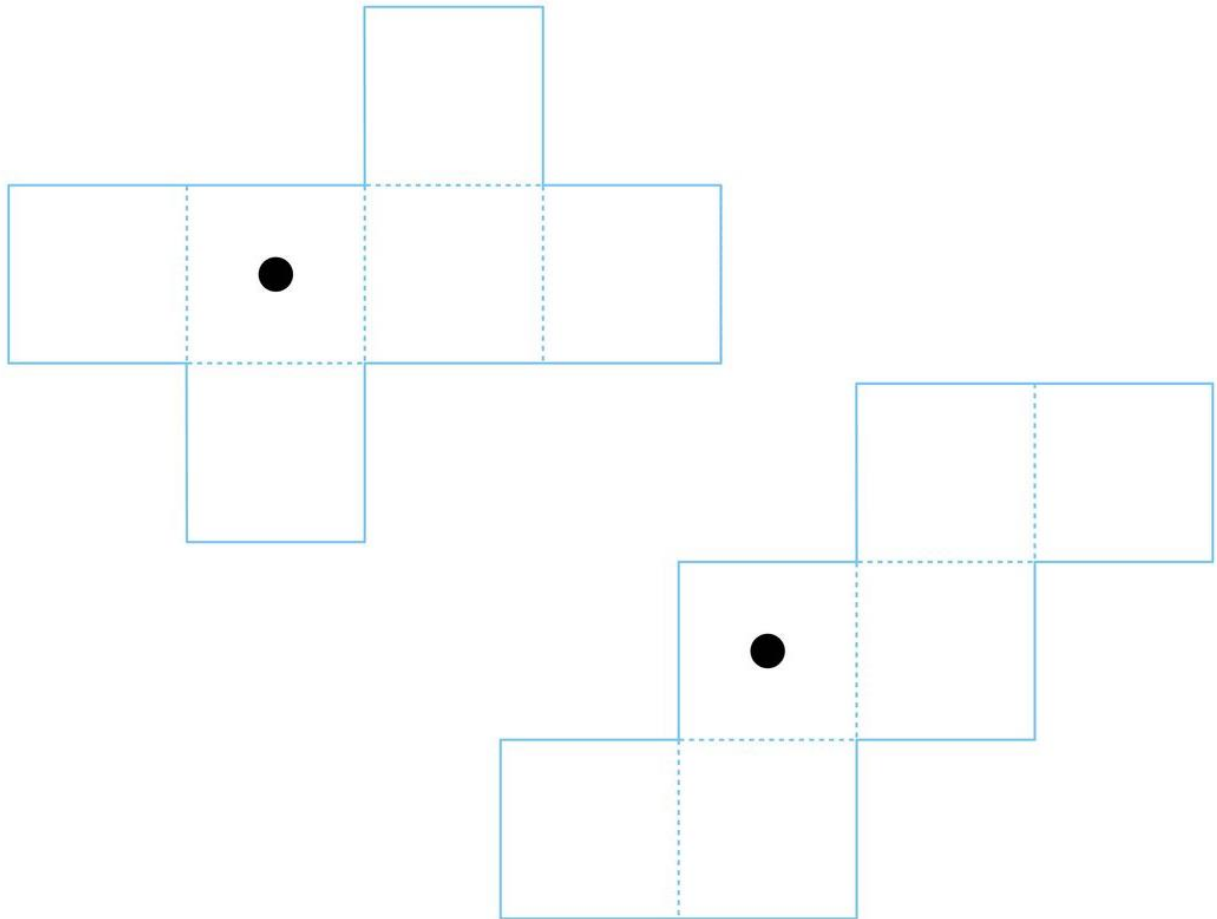


3

Maak een dobbelsteen die klopt.



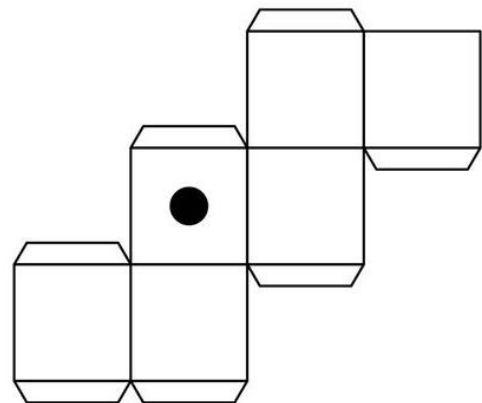
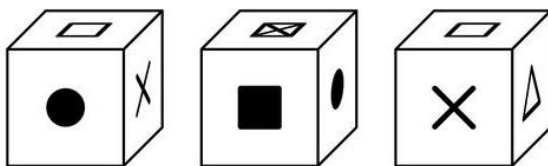
Teken de andere stippen op de goede plaats.
Controleer het met je printblad.
Knip uit en vouw de uitslag in elkaar.



KIJK TERUG



Teken de figuren op het goede vlak.



Bijlage D Eindtoets Rekenen

BLOK 9

TOETS

1 2 3

T

Reken uit.



$47 - 24 = \dots\dots$

$47 - 24 = \dots\dots$



$71 - 56 = \dots\dots$

$82 - 26 = \dots\dots$

$53 - 48 = \dots\dots$

$66 - 38 = \dots\dots$

$65 - 36 = \dots\dots$

$81 - 37 = \dots\dots$

$74 - 37 = \dots\dots$

$95 - 59 = \dots\dots$

1

Kruis de sommen aan waarbij aanvullen handig is.

Reken alleen die sommen uit op de getallenlijn.



$203 - 198 = \dots\dots$

$411 - 389 = \dots\dots$

$734 - 582 = \dots\dots$

$500 - 488 = \dots\dots$

$682 - 677 = \dots\dots$

$673 - 669 = \dots\dots$

$500 - 172 = \dots\dots$



1 2 3

2a

Reken uit met te veel.



$$7 \times 49 = 7 \times 50 - 7 \times 1 = 350 - 7 = 343$$

$$8 \times 19 = \dots\dots\dots$$

$$3 \times 59 = \dots\dots\dots$$

$$4 \times 48 = \dots\dots\dots$$

$$5 \times 98 = \dots\dots\dots$$

$$7 \times 39 = \dots\dots\dots$$

2b

Kruis de sommen aan waarbij je rekt met te veel.

Reken alleen die sommen uit. Laat zien hoe je rekt.



$6 \times 28 = \dots\dots\dots$

$5 \times 79 = \dots\dots\dots$

$6 \times 35 = \dots\dots\dots$

$3 \times 72 = \dots\dots\dots$

$5 \times 49 = \dots\dots\dots$

$3 \times 48 = \dots\dots\dots$

$3 \times 85 = \dots\dots\dots$

$4 \times 79 = \dots\dots\dots$

$5 \times 82 = \dots\dots\dots$

$8 \times 43 = \dots\dots\dots$

GA VERDER →



1 2 3

3b

Reken uit.



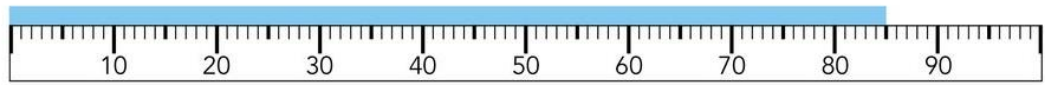
- $2 \times 9 = \dots\dots\dots$
- $3 \times 7 = \dots\dots\dots$
- $7 \times 4 = \dots\dots\dots$
- $8 \times 6 = \dots\dots\dots$
- $6 \times 5 = \dots\dots\dots$

- $5 \times 9 = \dots\dots\dots$
- $7 \times 8 = \dots\dots\dots$
- $7 \times 7 = \dots\dots\dots$
- $4 \times 8 = \dots\dots\dots$
- $9 \times 7 = \dots\dots\dots$

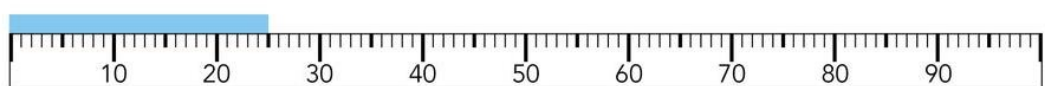
4

Hoe lang?

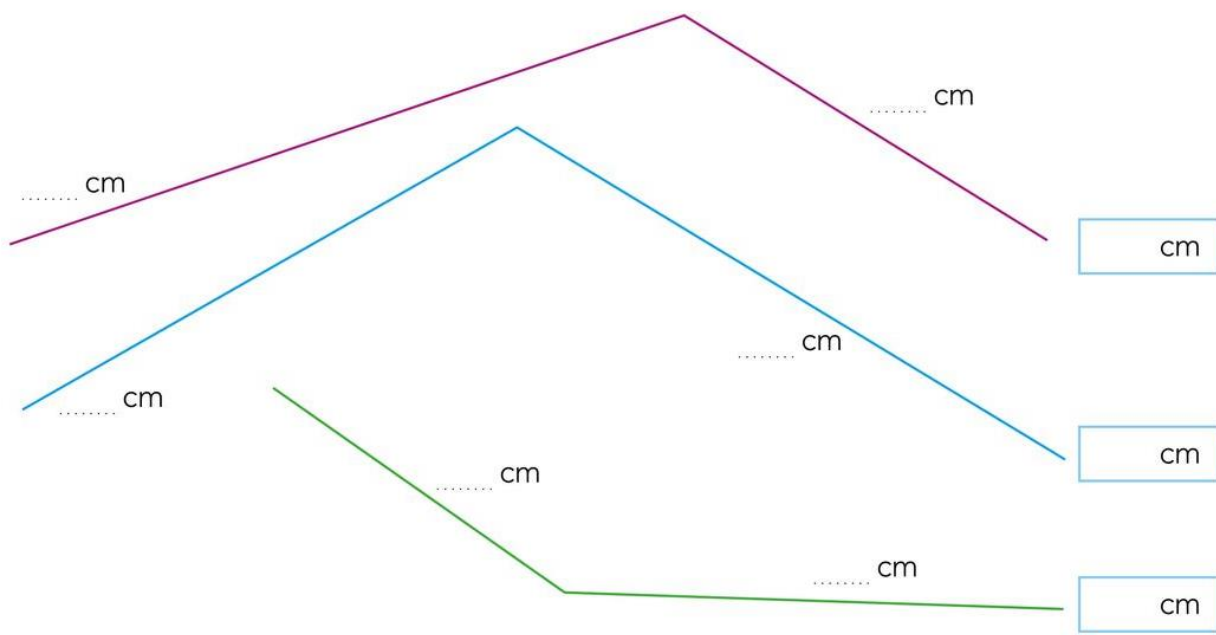
Meet met je liniaal.



cm



cm



KLAAR!

Appendix A: Poweranalyse Berekening

Parameters Instellen

Effectgrootte (d): 0.5

Significantieniveau (α): 0.05

Power ($1 - \beta$): 0.80

Z-scores bepalen

- Voor een α van 0.05 (tweezijdig), is de z-score $Z_{\alpha/2}$ ongeveer 1.96.
- Voor een β van 0.20 (power van 0.80), is de z-score Z_{β} ongeveer 0.84.

Formule voor steekproefgrootte

$$n = \left(\frac{(Z_{\alpha/2} + Z_{\beta})^2 \cdot (2\sigma^2)}{d^2} \right)$$

Berekening

$$n = \left(\frac{(1.96 + 0.84)^2 \cdot 2}{0.5^2} \right) = \left(\frac{(2.80)^2 \cdot 2}{0.25} \right) = \left(\frac{7.84 \cdot 2}{0.25} \right) = 62.72 \approx 63$$