

Kinderen ondersteunen tijdens onderzoekend leren: Instrueren of structureren?

Bachelorthese Psychologie – Instructie, Leren & Ontwikkeling

Opleiding Psychologie

Faculteit der Gedragswetenschappen

Universiteit Twente

Enschede, 26 juni 2013

1^e Begeleider: A. Lazonder

2^e Begeleider: Y. Mulder

Naam: Sjanou Drost

Studentnummer: s1121537

Samenvatting

Dit onderzoek richt zich op manieren om kinderen te ondersteunen bij het opstellen van hypothesen, experimenten en trekken van conclusies. Goede experimenten kunnen opgesteld worden door het toepassen van de control-of-variable strategie (CVS). Deze strategie leert dat alleen de waarde van de onderzochte variabele veranderd moet worden. Twee manieren om CVS gebruik te stimuleren zijn het aanbieden van instructie en het structureren van een taak. Verscheidene onderzoeken hebben de effectiviteit van deze strategieën vergeleken met onderzoekend leren zonder ondersteuning. Klahr en Li (2005) hebben de effectiviteit van CVS instructie op het ontwerpen van goede experimenten aangetoond. Lazonder en Kamp (2012) hebben dit aangetoond voor taakstructurering. Lazonder en Egberink (in press) hebben beide strategieën vergeleken, waaruit naar voren kwam dat beide strategieën effectiever waren met oog op gebruik van de CVS dan onderzoekend leren. Het bewijs voor de effectiviteit van beide strategieën is echter alleen aan de hand van logfiles onderzocht. Dit onderzoek was daarom gericht op het verkrijgen van direct bewijs; onder andere door kinderen te vragen naar hun hypothesen. De onderzoeksvraag die hierbij centraal stond luidde: *"Wat is de relatieve effectiviteit van directe instructie en taakstructurering bij het ondersteunen van kinderen tijdens onderzoekend leren?"*

Om deze onderzoeksvraag te beantwoorden hebben 55 leerlingen met een gemiddelde leeftijd van 10.89 van dezelfde basisschool met een computersimulatie van een gong gewerkt. Hiermee konden ze het effect van vier variabelen op nagalmtijd onderzoeken. De kinderen werden random toegewezen aan één van de drie condities, namelijk directe instructie ($n=18$), taakstructurering ($n=18$) en controle conditie ($n=19$). Kinderen uit de directe instructie conditie en controle conditie werkten met één algemene vraag, welke bestond uit vier variabelen. De kinderen uit de directe instructie conditie kregen CVS -instructie aangeboden. Dit werd gedaan door ze uit te leggen hoe ze goede experimenten op kunnen stellen en ze hier zelf mee te laten oefenen. Kinderen uit de taakstructureringsconditie kregen vier smalle vragen die ze achtereenvolgens moesten beantwoorden. Deze vier vragen bevatte elk steeds één variabele om te onderzoeken. Kinderen uit de controle conditie namen deel aan onderzoekend leren zonder ondersteuning.

Uit dit huidige onderzoek kwam naar voren dat kinderen uit de directe instructie conditie de CVS significant beter toepasten dan kinderen uit de uit de controle conditie. Kinderen uit beide experimentele condities trokken significant geldiger conclusies dan kinderen uit de controle conditie. Ook bleken kinderen uit beide experimentele condities significant meer variabelen te onderzoeken dan kinderen uit de controle conditie. Verder is er een substantiële correlatie tussen het gebruik van de CVS en de scores op de CVS natest gevonden. Kennis en gebruik van de CVS gaan dus hand in hand. De resultaten tonen aan dat het aanbieden van directe instructie de meest effectieve manier lijkt te zijn om kinderen op de basisschool te ondersteunen bij het ontwikkelen van wetenschappelijke vaardigheden. Een aanbeveling voor vervolgonderzoek is de hardop-denken methode om mogelijke cueing effecten bij het vragen naar hypothesen te voorkomen.

Abstract

Young children need support to develop scientific skills like preparation of hypothesis, preparation and implementation of experiments and drawing conclusions. This research focuses on ways to support children during inquiry learning. Good experiments can be designed by applying the control-of-variable strategy (CVS). This strategy learns to change only the variable of focus during experimentation. Two ways to stimulate use of CVS are offering instruction in the CVS and structuration of the task. Several studies have compared the effectiveness of those two strategies with unguided inquiry learning. Klahr en Li (2005) for example, demonstrated the effectiveness of CVS instruction on designing good experiments. Lazonder and Kamp (2012) have demonstrated this for the structuration of the task. Lazonder and Egberink (in press) compared both strategies, which showed that both experimental strategies proved to be better than unguided inquiry learning in stimulating the use of CVS. Evidence for the effectiveness of these strategies is only based on the examination of log files. The present research wants to find direct prove for the results from earlier research; including by asking children about their hypothesis. The central question was: *"What is the relative effectiveness of direct instruction and task structuring in supporting children during inquiry learning?"*

To answer this question, 55 pupils with an average age of 10.89 from the same school, worked with a computer simulation of a gong. With the simulation they could investigate the effect of four variables on reverberation time. The children were randomly allocated to one of three conditions, namely direct instruction ($n=18$), task structure ($n=18$) and control group ($n=19$). Children from the direct instruction condition and control condition worked with a general question, which consisted of four variables. The children in the direct instruction condition were offered CVS instruction. This instruction explained how good experiments could be designed and offered children practice in designing good experiments. Children in the task structuring condition worked with four narrow questions, which they needed to answer successively. Each of these four questions contained one variable to investigate. Children in the control group participated in unguided inquiry learning.

This present study showed that children from the direct instruction condition applied CVS better than children which participated in the control condition. Children from both experimental conditions drew more valid conclusions than children in the control condition. Also, children from both experimental conditions investigated more variables with their experiments. Further, there was found a substantial correlation between the use of the CVS and the scores on the CVS post-test. This shows that knowledge and use of CVS go together. The results show that the most effective way to support primary school children in the development of scientific skills seems to be direct instruction. A recommendation for further research is a think-aloud method to prevent possible cueing effects when asking children about their hypotheses.

Inhoudsopgave

Samenvatting	2
Abstract	3
Inhoudsopgave.....	4
1. Inleiding.....	5
2. Methode.....	11
2.1 Participanten	11
2.2 Materialen	11
2.2.1 Leeromgeving.....	11
2.2.2 CVS instructie	12
2.2.3 Testen.....	14
2.3 Procedure	15
2.4 Metingen en scoringsprocedures	16
3. Resultaten	18
4. Discussie.....	22
5. Referentielijst.....	27
6. Bijlages.....	30
6.1 Voorkennis test	30
6.2 Coderingsinstructies voorkennistest	32
6.3 CVS test.....	33
6.4 Coderingsinstructies CVS-test	43

1. Inleiding

Op de basisschool worden kinderen voorbereid op de toekomst. Dit betekent onder andere dat in toenemende mate gebruik gemaakt wordt van ICT om kinderen te ondersteunen in het leerproces. (van Galen & Gravemeijer, 2008). Leren met behulp van ICT stelt kinderen vanaf jonge leeftijd in staat om bepaalde ideeën te ervaren, ermee te werken en op de ideeën verder te bouwen, wat eerder niet mogelijk was door de complexiteit van deze ideeën (Kilderry et al., 2003).

Een onderdeel van leren van kinderen op de basisschool is onderzoek doen. Dit begint in de onderbouw waar kinderen vanaf acht jaar les krijgen over natuurkundige objecten en materialen. Ook doen kinderen zelf eenvoudige proefjes om te oefenen met het ontwerpen en uitvoeren van onderzoek (National Research Council, 1996). Deze lessen kunnen door ICT ondersteund worden, bijvoorbeeld door het gebruik van eenvoudige computersimulaties. De omgeving waarin ICT gebruikt wordt moet kinderen echter wel ondersteunen tijdens het leerproces om het leren van kinderen toe te laten nemen (Dwyer, 2007).

Kinderen leren dus al op de basisschool onderzoek te doen en de leeractiviteiten die ze hierbij uitvoeren vertonen parallellen met de wetenschap. Wetenschappelijk denken is het onderzoeken van een fenomeen om het te leren begrijpen. Een aantal vaardigheden die behoren tot het wetenschappelijk denken zijn het opstellen van hypothesen, ontwerpen van experimenten, bewijs evalueren en conclusies trekken voor het vormen of herzien van theorieën (Zimmerman, 2007). Voor het doen van wetenschappelijk onderzoek is wetenschappelijk denken nodig, wat bestaat uit cognitieve en metacognitieve vaardigheden. De ontwikkeling en het gebruik van deze vaardigheden vereisen veel oefening (Zimmerman, 2007). Daarom is het belangrijk dat kinderen hiermee op de basisschool al mee in aanraking komen en ook de mogelijkheid krijgen om te oefenen.

Hoewel wetenschappelijk denken een complexe vaardigheid is, blijken kinderen in staat om hypothetisch te denken en deductief te redeneren. Volgens Piaget is deze manier van denken en redeneren een manier van probleemoplossen die wordt gekarakteriseerd door het vermogen om hypothesen te testen en conclusies te trekken op basis van de resultaten (de Wit & van der Veer, 1989). Veel ontwikkelaars van leerplannen en lesmateriaal gingen uit van de theorie van Piaget, die benadrukte dat kinderen vooral concreet denken en dat het concreet-operationele kind niet de mogelijkheden heeft om wetenschappelijk te denken. Dit had ook betrekking op de vaardigheid om bij het uitvoeren van een experiment steeds één variabele te variëren en de andere variabelen constant te houden. Piaget en Inhelder (1958) suggereerden daarom dat kinderen op de basisschool niet in staat zouden zijn om hypothesen te formuleren, systematisch te experimenteren en conclusies te trekken op basis van resultaten. Later werd echter duidelijk dat kinderen naast een aantal kernactiviteiten (zoals classificatie en meten)–waar ze volgens de theorie van Piaget in de concreet-operationele fase toe in staat zijn–ook hypothetisch-deductief konden redeneren (Metz, 1995). Dit betekent dat kinderen vanaf

jonge leeftijd zich kunnen ontwikkelen in wetenschappelijk denken. Kortom, kinderen zijn competentere dan verwacht werd (Zimmerman, 2007).

Wetenschappelijk denken speelt een belangrijke rol bij inquiry learning (onderzoekend leren). De National Science Foundation definieert inquiry learning als “een benadering tot leren die zich richt op het proces van het ontdekken van de natuurlijke of materiele wereld en leidt tot het stellen van vragen, doen van ontdekkingen en tot het testen van deze ontdekkingen in de zoektocht naar nieuw begrip” (2000, p.2.). Inquiry learning is dus de wereld leren begrijpen door zelf onderzoek te doen (Lazonder, in press). De suggestie wordt gedaan dat, hoewel kinderen competentere zijn dan verwacht, ze op alle leeftijden aanvullende ondersteuning nodig hebben door de complexiteit van opstellen van hypothesen, ontwerpen van experimenten en het evalueren van bewijs. In het vervolg van deze paragraaf wordt beschreven hoe competent kinderen nu werkelijk zijn in deze vaardigheden.

Het opstellen en testen van hypothesen zijn belangrijke vaardigheden voor succesvol inquiry learning (Lazonder, in press). Jonge kinderen vormen vaak geen hypothesen. Wanneer kinderen ouder worden en de leeftijd van 12-14 jaar bereiken, vormen ze vaak aannemelijke hypothesen vanuit de kennis die ze al hebben. Deze kinderen hebben al meer specifieke kennis en hebben vaak ook sterke overtuigingen over een onderwerp (Penner & Klahr, 1996; Tomkins & Tunnicliffe, 2001). Daarom kiezen zij vaak één hypothese om te onderzoeken (Klahr, Fay & Dunbar, 1993). Ook volwassenen beginnen hun onderzoek met het focussen op hypothesen die in overeenstemming zijn met eerdere overtuigingen, maar verschillen van kinderen in pogingen om vervolgens andere hypothesen op te gaan stellen. Kinderen neigen hypothesen waarmee ze bekend zijn op te blijven stellen en vinden het vaak moeilijk hypothesen op te stellen die onwaarschijnlijk zijn (Lazonder, in press). Kinderen ervaren dus confirmation bias (Tschirgi, 1980). Volwassenen zijn echter beter in staat om hypothesen op te stellen die niet in overeenstemming met hun kennis zijn. Verder vinden zowel kinderen als volwassenen het moeilijk om nieuwe of alternatieve hypothesen op te stellen aan de hand van de data (Lazonder, in press)

De tweede centrale vaardigheid is experimenteren. Om goede experimenten te bedenken, moet de control-of-variable strategie (CVS) beheerst worden. Deze strategie leert ze om alleen de variabele waarvan het effect onderzocht wordt te manipuleren; de waarden van de overige variabelen moeten constant gehouden worden (Lazonder, in press). De CVS wordt vooral spontaan gebruikt om een aangeboden negatieve hypothese te falsifiëren en uit onderzoek blijkt dat zelfs jonge kinderen hiertoe in staat zijn (Tschirgi, 1980). Bij een hypothese die een wenselijke situatie beschrijft, gebruiken basisschoolleerlingen de CVS echter beduidend minder vaak. Kinderen vanaf 6 jaar lijken in staat te zijn om te kunnen onderscheiden tussen uitsluitende en niet uitsluitende experimenten (Zimmerman, 2007). Verder lijken kinderen vanaf 12 jaar goed te begrijpen hoe een goed experiment eruit hoort te zien (Schauble et al., 1995). De kennis over de CVS lijkt dus op jonge leeftijd al aanwezig te zijn. Toch lijkt het gebruik van de CVS, resulterend in goede experimenten, zich langzaam te ontwikkelen zonder instructie of oefening (Lazonder, in press).

De derde vaardigheid, het evalueren van bewijs, is het afleiden van algemeenheden uit data om een hypothese te genereren, testen of te verfijnen. Deze vaardigheid blijkt zich langzaam te ontwikkelen (Kuhn, Amsel & O'Loughlin, 1988). Hierbij zijn een tweetal aspecten belangrijk: observatie en interpretatie. Allereerst is het belangrijk vaardig te observeren om data te verkrijgen. Dit vereist coördinatie van domein specifieke kennis en algemene sensorische perceptuele vaardigheden. Bij kinderen zijn beide vaardigheden nog niet goed ontwikkeld (Eberbach & Crowley, 2009). Specifieke kennis van het onderwerp kan helpen om goed te observeren, vooral wanneer sensorische stimuli dubbelzinnig zijn (Penner & Klahr, 1996). Daarnaast richten kinderen zich bij het observeren vaak te veel op oppervlakkige kenmerken en negeren ze de onderliggende kenmerken van het fenomeen (Chi, Feltovich & Glaser, 1981). Na het verkrijgen van data is het belangrijk deze data te interpreteren. Uit onderzoek van Amsel en Brock (1996) blijkt dat wanneer kinderen de domein specifieke kennis nog niet hebben, ze de neiging hebben om eigen theorieën te vormen. Wanneer kinderen over een onderwerp al wel domein specifieke kennis hebben, blijkt dat bewijs dat consistent is met de overtuiging de overtuiging versterkt, terwijl inconsistent bewijs de overtuiging niet verzwakt. (Koslowski, Marasia, Chelenza & Dublin, 2008). Kortom, op elke leeftijd is er aanvullende ondersteuning nodig voor de observatie en interpretatie van data (Lazonder, in press).

In dit onderzoek wordt ingegaan op het ondersteunen van kinderen tijdens onderzoekend leren, met name tijdens het opstellen en uitvoeren van experimenten. Een manier van ondersteuning is het aanbieden van CVS instructie. Chen en Klahr (1999) hebben onderzoek gedaan naar het aanbieden van directe instructie voorafgaand aan het experimenteren. Kinderen bleken door deze instructie in staat om de CVS te begrijpen en vervolgens toe te passen op het ontwerpen van eenvoudige experimenten. Ook Klahr en Nigam (2004) hebben onderzoek gedaan naar het effect van CVS-instructie op het experimenteergedrag van basisschoolleerlingen. De leraar heeft directe instructie aangeboden door uit te leggen waarom CVS werkt en heeft kinderen goede en slechte voorbeelden gegeven. Dit onderzoek toonde aan dat het aanbieden van directe instructie voorafgaand aan het experimenteren ertoe leidt dat kinderen systematisch experimenteren. Daarnaast blijkt uit onderzoek van Klahr & Li (2005) dat directe instructie leidt tot vaardiger gebruik van de CVS dan onderzoekend leren zonder instructie.

De fundering voor een alternatieve vorm van ondersteuning, taakstructurering, werd aangeboden door Kuhn en Dean (2005). Kuhn en Dean (2005) gaven kinderen de hint zich tijdens de les te richten op één van de variabelen. Vervolgens moesten ze zelf de waarde van die variabele vaststellen. Uiteindelijk konden kinderen conclusies trekken op basis van de resultaten van de experimenten. Uit dit onderzoek bleek dat de kinderen die een hint ontvingen doelgerichter waren in het ontwerpen van experimenten en betere conclusies trokken. Deze conclusies gaven aan dat kinderen die de hint ontvingen vaardiger waren in het gebruik van de CVS dan de controlegroep die experimenteerde zonder hints.

Lazonder en Kamp (2012) bouwden verder op het onderzoek van Kuhn en Dean (2005). Dit deden zij door een brede onderzoeksvraag op te delen in vier specifieke vragen, die achtereenvolgens

onderzocht en beantwoord moesten worden, zodat kinderen steeds één variabele tegelijk onderzochten en het effect van deze variabele vast konden stellen. Kinderen die deze taakstructurering ontvingen, pasten de CVS vaker toe en trokken betere conclusies dan kinderen uit de controle conditie die zonder ondersteuning experimenteerden.

Lazonder en Egberink (in press) hebben de effectiviteit van de beide hierboven genoemde strategieën vergeleken. Hun onderzoek bestond uit drie condities: directe instructie in de CVS, taakstructurering, en een controle conditie waarin geen ondersteuning voor het systematisch experimenteren werd gegeven (unguided inquiry). Uit de vergelijking tussen deze condities bleek dat kinderen uit de taakstructureringsconditie net zo systematisch experimenteerden als kinderen uit directe instructie conditie en beter dan kinderen uit de controle conditie. Taakstructurering was bovendien effectiever voor het trekken van geldige conclusies op basis van de experimentele resultaten vergeleken met beide andere condities.

Het bewijs voor de effectiviteit van taakstructurering, en in mindere mate directe instructie, komt echter uit indirecte metingen (log file analyse). Hierdoor bestaan er weliswaar sterke aanwijzingen voor de effectiviteit van beide interventies, maar deze zijn nog niet gevalideerd door directe metingen van het onderzoekend leerproces. In dit huidige onderzoek wordt daarom geprobeerd direct bewijs te vinden voor de gevonden resultaten uit eerder onderzoek.

Dit wordt gedaan aan de hand van het Scientific Discovery as Dual Search Model (SDDS). Het model is door Klahr en Dunbar (1988) opgesteld om wetenschappelijk redeneren te beschrijven. Het model bestaat uit drie componenten, namelijk het genereren van hypothesen, het bedenken en uitvoeren van experimenten en het trekken van conclusies op basis van resultaten van het experimenteren. Het model beschrijft het onderzoekend leerproces als een zoektocht door een ruimte van mogelijke hypothesen en een zoektocht door een ruimte van relevante experimenten om deze hypothesen te testen. Het doorzoeken van de hypothese ruimte heeft betrekking op het opstellen van hypothesen uit eerdere kennis of op basis van experimenten (Lazonder, Wilhelm & Hagemans, 2008). Het doorzoeken van de experimentruimte heeft betrekking op de drie stappen die al eerder genoemd zijn, namelijk het selecteren van één variabele als focus voor het experiment, de waarde van deze focale variabele kiezen en de waarden van de andere variabelen constant houden (Klahr & Dunbar, 1988). Verder hebben Klahr en Dunbar (1988) ontdekt dat wanneer kinderen een taak onderzoeken die bestaat uit meerdere variabelen, zoals in het geval van de directe instructie en controle conditie, ze de experimentruimte met een ander doel onderzoeken. Wanneer kinderen bij een taak één variabele tegelijkertijd onderzoeken (single variabele taak), zou de experimentruimte effectief beperkt worden. Door deze beperking zijn kinderen telkens gericht op één variabele en mogelijk hebben ze over deze variabele mogelijk specifiekere hypothesen dan wanneer ze hun aandacht hadden moeten verspreiden over meerdere variabelen. De derde component van het SDDS model is het trekken van conclusies op basis van resultaten. Resultaten van de experimenten kunnen vragen om aanpassing van de eigen theorieën.

Directe instructie en taakstructurering kunnen deze drie componenten van het SDDS model beïnvloeden. Zo bleek uit onderzoek van Kuhn en Dean (2005) dat het aanbieden van de hint om één variabele tegelijk te onderzoeken, de eerste stap uit het SDDS model, samenging met doelgerichte experimenten en geldige conclusies. Dit is ook gebleken uit onderzoek van Lazonder en Egberink (in press). Uit het laatste onderzoek is ook gebleken dat het instrueren van kinderen in het SDDS model resulteert in systematischer experimenteren en het trekken van betere conclusies dan het geval was bij onderzoekend leren zonder ondersteuning.

In dit huidige onderzoek wordt de effectiviteit van taakstructurering en directe instructie met oog op de drie componenten van het SDDS model onderzocht. De onderzoeksvraag die hierbij centraal staat luidt: *"Wat is de relatieve effectiviteit van directe instructie en taakstructurering bij het ondersteunen van kinderen tijdens onderzoekend leren?"*

De experimentele condities ontvangen in dit huidige onderzoek steun bij het opstellen van experimenten, beide op de eigen karakteristieke manier. Directe instructie zou kinderen helpen bij het opstellen en uitvoeren van experimenten door voorafgaand aan het experimenteren kinderen de CVS uit te leggen, waardoor ze leren systematisch te experimenteren. Taakstructurering zou het gebruik van de CVS op een andere manier stimuleren. Het begint met de veronderstelling dat kinderen geen goede experimenten opstellen, omdat de gestelde onderzoeksvraag, waarin meerdere variabelen genoemd worden, niet effectief zou zijn. Taakstructurering deelt de onderzoeksvraag daarom op in deelvragen waarin steeds één variabele centraal staat. Hierdoor zijn kinderen aan het begin van hun onderzoek telkens gericht op één variabele. Dit vervangt de eerste stap van het SDDS model, namelijk het kiezen van een variabele als focus voor het onderzoek. Hieruit volgen het kiezen van waarden voor de variabele van focus en de waarden van overige variabelen dan vanzelf. Taakstructurering beperkt dus het zoeken in de experimentruimte door kinderen te focussen op één variabele voorafgaand aan het experimenteren. Hierdoor zou de CVS dus afgeleid kunnen worden van een goed geformuleerde onderzoeksvraag (Lazonder & Egberink, in press). Kinderen uit de unguided inquiry conditie experimenteren zonder ondersteuning. Uit onderzoek van Lazonder (in press) blijkt dat instructie of oefening nodig is in de CVS om deze goed toe te kunnen passen en dus systematische experimenten op te stellen.

De eerste *hypothese* luidt dan ook dat directe instructie en taakstructurering beiden effectiever zijn dan geen strategie in het ondersteunen van kinderen bij het opstellen en uitvoeren van experimenten. Bovendien wordt er verwacht dat directe instructie en taakstructurering beiden even effectief zijn.

De tweede *hypothese* luidt dat kinderen uit de taakstructurering specifiekere hypothesen opstellen dan kinderen uit de directe instructie conditie en de controle conditie, omdat bij taakstructurering de eerste stap (het kiezen van een variabele als focus) expliciet wordt ondersteund bij de taakstructureringsconditie.

De derde *hypothese* luidt dat zowel directe instructie als taakstructurering leiden tot betere conclusies dan onderzoekend leren, omdat kinderen in beide experimentele condities volgens de verwachting systematischer zullen experimenteren, wat vervolgens zal leiden tot geldiger conclusies. Verder wordt geen verschil verwacht tussen kinderen uit de directe instructie conditie en kinderen uit de taakstructurering, omdat beide condities gestimuleerd worden de CVS toe te passen.

De vierde *hypothese* luidt dat kinderen uit de taakstructureringsconditie meer variabelen onderzoeken bij het experimenteren dan kinderen uit de directe instructie conditie en controle conditie, omdat de aandacht van kinderen uit de taakstructureringsconditie specifiek gericht wordt op de variabele die ze moeten onderzoeken en dit niet het geval is bij de beide andere condities.

2. Methode

2.1 Participanten

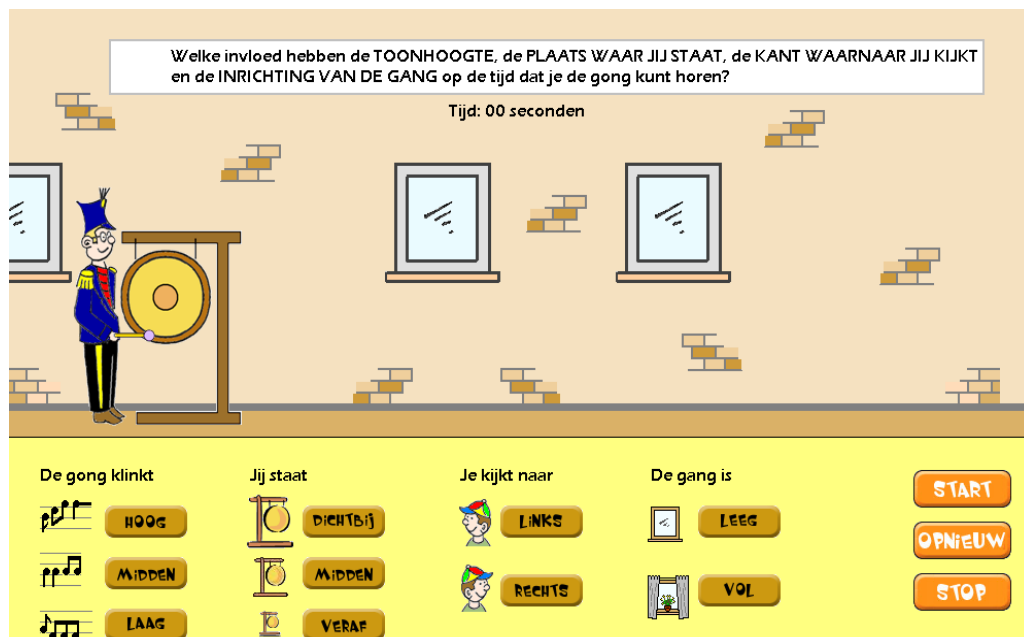
Het onderzoek is afgenomen op twee basisscholen in Nederland. De steekproef bestond uit 55 leerlingen, waarvan de gemiddelde leeftijd 10.89 ($SD = 0.86$) was. Verder bestond deze uit 26 jongens en 29 meisjes. Alle kinderen werden random toegewezen aan één van drie condities, namelijk directe instructie ($n=18$), taakstructurering ($n=18$) of aan de controle conditie ($n=19$).

2.2 Materialen

2.2.1 Leeromgeving

In dit onderzoek werd een computersimulatie van een gong gebruikt om het experimenteelgedrag bij kinderen te onderzoeken. Kinderen interacteerden met de simulatie en onderzochten hoe elk van de variabelen nagalmtijd beïnvloedde. De variabelen waren toonhoogte (laag, medium, of hoog), afstand vanaf de gong (dichtbij, halverwege of ver weg), de kijkrichting (richting de gong of van de gong af) en het interieur van de ruimte (leeg of gevuld met decoratie). Toonhoogte had een positief effect op nagalmtijd: hoe hoger de toon, hoe langer de nagalmtijd. Afstand had een differentieel effect: een persoon dichtbij de gong zou de echo langer horen dan iemand in het midden. De persoon in het midden zou de galm net zo lang horen als de persoon aan de overkant van de kamer. De derde variabele, kijkrichting, had geen effect op de nagalmtijd. Een lege kamer veroorzaakte een langere nagalmtijd dan een kamer gevuld met decoratie.

Kinderen konden de invloed van deze variabelen onderzoeken door proefjes te doen. Ze konden deze proefjes opstellen door onderin het scherm een bepaalde waarde voor elke variabele te selecteren. Na het opstellen van een proefje, konden kinderen de simulatie runnen door op de 'START' button te klikken, waardoor ze het effect van de waarde van de variabele op de nagalmtijd zagen. De nagalmtijd werd weergegeven op het scherm wanneer het geluid niet meer te horen was. De opstelling en resultaten van de proefjes konden kinderen eventueel op papier noteren. De conclusies die kinderen trokken op basis van de resultaten moesten de kinderen noteren in de simulatie, maar ze hadden ook de mogelijkheid deze te noteren op papier. Alle acties van de kinderen werden opgeslagen in een logfile. In Figuur 1 is de interface van de gong simulatie te zien.



Figuur 1: Interface van de gong simulatie

Er zijn twee verschillende versies ontworpen voor de drie condities. Deze verschilden alleen met oog op de onderzoekstaak die kinderen kregen. Kinderen in de directe instructie en controle conditie, kregen een algemene vraag die alle variabelen in de simulatie bevatte (“Wat zijn de invloeden van toonhoogte, afstand, kijkrichting en interieur van de ruimte op de tijd dat je de echo hoort?”). Kinderen kregen de vraag te zien zodra ze inlogden in de leeromgeving. Tijdens het onderzoeken van de variabelen konden ze de vraag bovenaan het scherm zien. Kinderen moesten de vraag lezen, onderzoeken en beantwoorden. Kinderen in de taak structureringsconditie moesten dit onderzoek voor elk van de vier variabelen doen, dus in totaal vier keer. Voor elke variabele, kregen ze een vraag te zien die hun aandacht richtte op die variabele, waardoor ze één variabele tegelijkertijd onderzochten (“Wat is de invloed van toonhoogte op de tijd dat je de echo hoort?”). Kinderen moesten deze vraag beantwoorden, voordat ze de volgende vraag gingen onderzoeken. Het terugkeren naar een vorige vraag was niet mogelijk.

2.2.2 CVS instructie

De CVS instructie werd individueel aan kinderen aangeboden. Er werd bij deze CVS instructie een andere simulatie gebruikt dan de simulatie die gebruikt werd bij het experimenteren, namelijk een simulatie over hellingen. De onderzoeker begon deze instructie met uit te leggen dat veel variabelen effect kunnen hebben op hoe ver de ballen van de helling afrollen. Vervolgens toonde de onderzoekster aan de kinderen een tweetal voorbeeldexperimenten. Deze waren gericht op de hoogte van de helling. Voor deze variabele was het eerste voorbeeld een niet systematisch experiment; het tweede experiment wel. Dit betekent dat de onderzoeker eerst een verkeerd voorbeeld gaf, omdat de twee hellingen niet alleen in hoogte verschilden, maar ook op de andere variabelen. Het kind werd gevraagd of dit experiment een goede test was om uit te vinden of hoogte een verschil maakte in de lengte die de bal aflegde op de helling. Vervolgens werd het kind gevraagd, wanneer één van de ballen verder zou

rollen, je met zekerheid kon zeggen dat dit kwam door de hoogte van de helling. De onderzoeker bevestigde of corrigeerde het antwoord van het kind door uit te leggen waarom dit experiment geen goede test was (Zie Tabel 1; verklaring slechte test).

Vervolgens gaf de onderzoeker een voorbeeld van een systematisch experiment en vroeg het kind of dit een goede test was en of het kind met zekerheid kon oordelen over de hoogte van de helling. De onderzoeker bevestigde of corrigeerde het kind, waarbij vervolgens werd uitgelegd waarom deze systematische vergelijking een goede test was (Zie Tabel 1; verklaring goede test).

Tot slot gaf de onderzoeker een samenvatting van de CVS (Zie Tabel 1; samenvatting CVS).

Tabel 1. *Uitleg van de proefleider tijdens de CVS instructie*

Soort uitleg	Uitwerking uitleg
Verklaring slechte test	“Eigenlijk kon je niet met zekerheid zeggen dat het de hoogte was, wat de ene bal verder liet rollen dan de andere, omdat deze twee hellingen niet alleen verschilden in hoogte. Het kan zijn dat de ene helling hoger is dan de andere waardoor de ene bal verder rolt dan de andere, maar zie je dat de hellingen ook verschillende ballen en lengte hebben? Het gewicht van de ballen en de lengte van de helling verschillen ook, waardoor de ene bal verder zou kunnen rollen dan de andere bal. Zoals je ziet, kun je met deze vergelijking niet zeggen of het de hoogte of het gewicht van de bal was die de ene bal verder de helling af liet rollen dan de andere.”
Verklaring goede test	“Eigenlijk kon je met zekerheid zeggen door deze vergelijking of hoogte zorgde voor een verschil in de lengte die de ballen op de hellingen afleggen, omdat de hellingen alleen verschillen in hoogte. Het gewicht van de ballen is gelijk, zie je dat? En zie je dat de hellingen even lang zijn? Dus wanneer één van de ballen verder van de helling af eindigt, weet je dat alleen de hoogte hiervoor verantwoordelijk kon zijn, omdat de hellingen verder gelijk zijn.”
Samenvatting CVS	“Oké, nu weet je hoe je goede proefjes op kun stellen met de hellingen. De voorbeelden waarmee we begonnen, waren niet goed, maar we hebben ze zo aangepast dat het nu goede proefjes zijn toch? En nu weet je dat wanneer je wilt zien of iets van de hellingen een verschil maakt in hoe ver de bal vanaf de helling rolt, je ervoor moet zorgen dat de hellingen maar op één manier verschillen. Je wilt dat de hellingen verschillen op datgene waarvan jij wilt onderzoeken of het verschil maakt. Net zoals in de voorbeelden van net, maakten we hoogte van de helling anders (hoog of laag), maar we zorgden ervoor dat de andere dingen, zoals het gewicht van de ballen en de lengte van de hellingen gelijk waren, zodat we konden vertellen of een hoge of lage baan verschil maakte. En we gebruikten dezelfde ballen. Alleen wanneer je deze vorm van vergelijkingen maakt, kun je vertellen of iets een verschil maakt.”. Wanneer je een goed proefje wilt opstellen, moet je dus eerst één onderwerp kiezen wat je wilt onderzoeken. Dit onderwerp kun je op bijvoorbeeld hoogte (hoog of laag) laten verschillen, maar je zorgt ervoor dat de dingen die je niet wilt onderzoeken niet verschillen. Vervolgens werd gevraagd: ‘Begrijp je dit?’

Na het verduidelijken van deze stappen konden kinderen zelf aan de slag gaan met het simulatieprogramma. Eerst kregen kinderen twee voorbeelden van experimenten te zien die ze af moesten maken zodat het goede experimenten zouden zijn. Het eerste voorbeeld experiment begon met de uitleg: “Nu wil je onderzoeken hoe de plaats van de bal op de helling te maken heeft met hoe ver de bal rolt.” De proefleider koos de waarden voor het eerste deel van het experiment en klikte op

de 'START' button. Het kind moest daarna de waarden van het tweede deel invullen, zodat het een goede proef was. Het tweede voorbeeld was gelijk aan het eerste, met als enige verschil dat het betrekking op de variabele gewicht had. Na deze twee proefjes kregen kinderen de mogelijkheid om zelf een experiment op te stellen. Op deze experimentjes ontvingen ze feedback van de onderzoekster met de uitleg waarom het onderzoek goed of fout was. Dit verduidelijken ging door tot de onderzoeker er van overtuigd was dat de kinderen wisten hoe en waarom ze de CVS toe moesten passen.

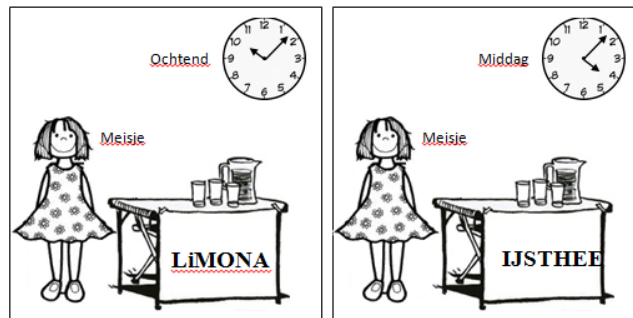
2.2.3 Testen

Een test, bestaande uit vier items, werd afgenomen om vooraf de overtuigingen van kinderen over de nagalmtijd te onderzoeken. Onderdeel van deze test was het uitleggen van het concept nagalmtijd: "De echo is de tijd dat je het geluid hoort nadat het poppetje op de gong geslagen heeft". Elke vraag bevatte de invloed van één variabele op de nagalmtijd. Het waren gesloten vragen en kinderen moesten beantwoorden of ze dachten dat een bepaalde variabele invloed had op de nagalmtijd en welke richting het effect van de variabele was. Een voorbeeld was: "Maakt het uit of de gong een hoog of laag geluid maakt?". Als hier "ja" op geantwoord werd, moest de richting van effect aangegeven worden. De richting moesten kinderen aangeven door te omcirkelen of een hoog geluid langer of korter gehoord zou worden dan een laag geluid.

De kennis van kinderen over de CVS is gemeten met de CVS test die zowel als voortest en als natest werd afgenomen. Deze test bestond uit drie domeinen, namelijk het verkopen van drankjes, groei van planten en modellen van vliegtuigen. Elk domein bestond uit drie items die waren ontleend aan de Story Problem test (Chen & Klahr, 1999). Voor elk domein bevatte één item een systematisch experiment en de andere twee items bevatten niet systematische experimenten. Bij alle items moesten kinderen een korte tekst lezen die het probleem introduceerde en de focale variabele introduceerde, de twee plaatjes bekijken die de experimentele setup visualiseerde en besluiten of de vergelijking systematisch was ja of nee. Kinderen konden antwoorden met: "ja, het is een goede test" of "Nee, het is een slechte test". Wanneer ze kozen voor de optie dat het een slechte test was, moesten ze de slechte test corrigeren door het plaatje te veranderen. In Figuur 2 staat vraag 1 van de CVS-test weergegeven.

Meneer Mulder verkoopt drankjes.
Maar hij verkoopt niet altijd even veel.

Meneer Mulder wil weten of dat komt door het moment van de dag (ochtend of middag)
In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet meneer Mulder straks zeker of het moment van de dag iets uitmaakt voor hoeveel drankjes er worden verkocht? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

Figuur 2: Vraag 1 van de CVS-test

2.3 Procedure

Het onderzoek vond plaats tijdens reguliere lessen. De eerste les begon met de afname van de test om de overtuigingen over nagalmtijd te onderzoeken. De onderzoeker illustreerde eerst het concept nagalmtijd en las vervolgens één voor één de vragen voor. Kinderen gaven hun antwoorden aan op papier na elke voorgelezen vraag. Hierna werd de CVS voortest afgenomen om de kennis van kinderen over de CVS te onderzoeken. De onderzoekster doorliep samen met de klas de voorbeeldvraag. De twee plaatjes werden doorlopen en kinderen werd verteld dat ze één antwoord moesten kiezen: het was een goede of een slechte test. Daarnaast werd kinderen verteld dat als ze kozen voor het antwoord “Het is een slechte test”, ze de plaatjes zo aan moesten passen dat het een goede test zou worden. Ook was er ruimte voor vragen. Hierna konden kinderen de negen vragen beantwoorden.

Vervolgens werden de kinderen één voor één uit de klas gehaald. De kinderen uit de directe instructie conditie ontvingen eerst de CVS instructie. Na de directe instructie ontvingen ze een algemene introductie in de simulatie. Kinderen uit de taakstructurering en de controle conditie begonnen met deze algemene introductie in de leeromgeving. De onderzoeker toonde de simulatie op een laptop, legde de interface elementen uit en liet de kinderen zien hoe ze experimenten op konden stellen en uit konden voeren om uit te vinden wat het effect van een variabele op de nagalmtijd was. Ook demonstreerde de onderzoekster hoe kinderen de resultaten af konden lezen. De kinderen werd expliciet verteld dat ze de aanwijzing op het scherm moesten volgen en de antwoorden op de vraag in een apart tekstveld in de computersimulatie moesten noteren. Vervolgens werd ingelogd in de simulatie en de onderzoekster stimuleerde kinderen tot het beginnen van het onderzoek door de vraag

in de simulatie zorgvuldig te lezen. Voor elk proefje dat kinderen opstelden, werd er ze gevraagd even te stoppen en werd ze een drietal vragen voorgelegd om erachter te komen met welk doel kinderen de hypothesen opstelden. Vragen die de kinderen gesteld werden voor het uitvoeren van de experimentjes waren:

- Vraag 1: “Wat ga je onderzoeken met dit proefje?”
- Vraag 2: “Wat denk je dat uit dit proefje naar voren zal komen?”
- Vraag 3: “Waarom denk je dat?”

Hierbij werd ook duidelijk geobserveerd wat de acties van kinderen waren. Wanneer het erop leek dat kinderen al voorafgaand aan het stellen van de vragen op de ‘START’ button klikten, werd ze gevraagd even te stoppen om toch de vragen te beantwoorden. Gebeurde dit wel dan was de proefleider erop voorbereid en draaide het computerscherm zo snel mogelijk weg.

Na het beantwoorden van deze drie vragen, kon het kind op de ‘START’ button klikken om het effect op nagalmtijd te onderzoeken. Dit herhaalde zich tot het kind het antwoord op de vraag die gesteld werd in de computersimulatie kon beantwoorden. De onderzoekster controleerde dit met de vraag: “Denk jij dat je het antwoord op de vraag kunt geven?”. In het geval van een bevestigend antwoord werd er op de ‘VERDER’ button geklikt en kon het kind het antwoord op de vraag invullen in het tekstveld. Vervolgens sloot de onderzoekster de computersimulatie af. Het kind kon weer terug naar de klas en de volgende leerling werd uit de klas gehaald.

2.4 Metingen en scoringsprocedures

Door middel van de voorkennistest werden de overtuigingen van kinderen over de nagalmtijd van de gong onderzocht. Per vraag konden twee punten verdiend worden, namelijk voor het correct aangeven of een variabele wel of geen invloed had en een extra punt wanneer ze de richting van de variabele ook goed hadden. Op deze test konden kinderen dus een maximum score van 7 punten behalen.

Daarnaast werd de kennis van kinderen over de CVS onderzocht door middel van de CVS-test. Voor het goed beantwoorden van de vraag kregen kinderen 1 punt. Wanneer een item geen correcte test was, konden kinderen een extra punt verdienen door het correct verbeteren van deze test. Het maximum aantal punten wat kinderen konden verdienen was op deze test 15 punten. De interne consistentie van de items, zowel van de CVS voortest als de CVS natest, zijn berekend door middel van Cronbach's α . De Cronbach's α van de CVS als voortest 0.85; van de CVS natest 0.88. De Cohen's κ inter-beoordelaars betrouwbaarheid voor het berekenen van de scores op de CVS-natest bleek bij 10 willekeurige testen 1.00 te zijn.

Aan de hand van de hypotheseformulieren en de logfiles werden de interacties van kinderen met de simulatie onderzocht. Hierbij is gekeken naar het aantal proefjes dat de kinderen opgesteld hadden en hoe lang ze met de simulatie gewerkt hadden. Als proefje werd elke keer wanneer het kind

op 'START' klikte meegerekend in de logfiles. Dit werd samengevoegd tot één experiment wanneer de doelen van de hypothesen van de proefjes gelijk waren. Het *gebruik van de CVS* werd berekend door het aantal goede experimenten te delen door het totaal aantal experimenten. Er werd dus een percentage berekend wat duidde op een CVS-score. Hoe hoger het percentage, hoe systematischer kinderen experimenteerden. De Cohen's κ inter-beoordelaars betrouwbaarheid voor deze methode is berekend en bleek 0.86 te zijn.

Ook de *specificiteit van de hypothesen* werd onderzocht. Deze werd berekend door het toekennen van punten, namelijk één punt voor het benoemen van het effect van een variabele, een extra punt voor het aangeven van de juiste richting van het effect en nog een extra punt voor het benoemen van het exacte aantal seconden dat de echo van de gong te horen was.

Ook de *geldigheid van conclusies*, die kinderen genoteerd hadden in het tekstveld of de tekstvelden (in het geval van de taakstructurering) van de simulatie, is onderzocht. Deze werd ook berekend door het toekennen van punten, namelijk één punt voor het wel of niet van invloed zijn van een variabele op nagalmtijd en een extra punt voor het aangeven van de juiste richting van het effect van deze variabele. Deze manier om antwoorden te scoren is ook in eerdere onderzoeken gebruikt en de Cohen's κ inter-beoordelaars betrouwbaarheid was 0.88 (Lazonder & Kamp, 2012).

De *mate van volledigheid van experimenten* is op twee manieren onderzocht. Allereerst is er gekeken naar het aantal genoemde variabelen tijdens het experimenteren. Vervolgens is er gekeken hoe veel variabelen kinderen genoemd hadden in hun conclusies.

3. Resultaten

Tabel 2 toont de gemiddelde scores van de kinderen uit de drie condities. Deze zijn geanalyseerd door middel van ANOVA voor de condities directe instructie, taakstructurering en controle. Als de ANOVA een significant verschil aantoonde, is er een post hoc analyse volgens de methode van Bonferroni uitgevoerd. De grootte van de condities verschilde bij de verschillende onderdelen van het onderzoek vanwege leerlingen die door ziekte afwezig waren.

Tabel 2. *Gemiddelden en standaardafwijking voor de performance van kinderen in experimenteren*

	Directe instructie (<i>n</i> =18)	Taakstructurering (<i>n</i> =18)	Controle (<i>n</i> =19)
<i>Voormetingen</i>			
Voorkennis test (max. 7) ¹	4.06 (1.70)	4.12 (1.27)	4.63 (1.61)
CVS-test (max. 15) ²	10.22 (4.18)	8.82 (4.85)	9.65 (4.50)
<i>Interactie met simulatie</i>			
Tijdsduur (min.)	19.11 (3.34)	21.16 (5.65)	17.54 (5.58)
Aantal proefjes	9.00 (2.47)	11.00 (2.50)	7.74 (1.54)
Aantal experimenten	4.61 (1.42)	5.33 (1.85)	3.58 (1.58)
Gebruik CVS (%)	75.04 (32.03)	51.09 (29.86)	40.50 (36.30)
Specificiteit hypothesen (max. 3)	1.70 (0.57)	1.93 (0.58)	1.55 (0.90)
Geldigheid conclusies	3.33 (1.82)	3.39 (1.72)	1.26 (1.52)
Volledigheid experimenten(max.4)	3.56 (0.62)	3.44 (0.62)	2.75 (1.15)
Volledigheid conclusies (max.4)	2.78 (1.48)	2.61 (1.34)	1.05 (1.35)
<i>Nametingen</i>			
CVS-test (max. 15) ³	13.11 (2.83)	11.12 (4.74)	11.06 (4.54)

Note.

¹ Bij de voorkennis test was de grootte van de drie condities respectievelijk *n*= 19, *n*=18, *n*=17.

² Bij de CVS-test als voormeting was de grootte van de drie condities respectievelijk *n*= 17, *n*=18, *n*=17.

⁵ Bij de CVS-test als nameting was de grootte van de drie condities respectievelijk *n*= 16, *n*=18, *n*=17.

De scores van kinderen op de voorkennistest tonen aan dat kinderen enige kennis hadden van de invloed van de variabelen op de gong (zie Tabel 2). Een ANOVA toonde aan dat de scores tussen de drie condities op de voorkennis test niet significant verschilden, $F(2,51) = 0.78$, $p = .464$. De scores op de CVS-voortest tonen aan dat kinderen redelijk deze benadering tot experimenteren toe wisten te

passen, maar ook hier geen volledige kennis van hadden (zie Tabel 2). Wederom werd geen significant verschil gevonden tussen de drie condities, $F(2,49) = 0.42, p = .657$.

Tijdens het onderzoek werkten de kinderen gemiddeld 19.24 minuten ($SD = 5.12$) met de computersimulatie. De gemiddelde tijdsduur verschilde niet tussen de drie condities, $F(2,52) = 2.43, p = .098$.

De kinderen hebben gemiddeld 9.13 proefjes ($SD = 2.61$) en 4.49 experimenten ($SD = 1.75$) uitgevoerd, voordat ze kwamen met hun antwoord op de gestelde vraag. De proefjes en experimenten zijn verder geanalyseerd om te onderzoeken in welke mate kinderen in de drie condities systematisch experimenteerden. Allereerst is er gekeken naar het aantal proefjes die kinderen in de drie condities opstelden. Het aantal proefjes bleek significant te verschillen tussen de drie condities, $F(2,52) = 11.87, p < .001$. Uit een post hoc analyse bleek dat kinderen uit de taakstructureringsconditie meer proefjes opstelden dan kinderen uit de controle conditie, $p < .001$ en meer proefjes dan kinderen uit de directe instructie conditie, $p = .026$. Het aantal proefjes bleek niet significant te verschillen tussen de controle conditie en de directe instructie, $p = .120$.

Vervolgens is er gekeken naar het aantal opgestelde experimenten. Uit een ANOVA kwam naar voren dat er een significant verschil bestond tussen de drie condities wat betreft het aantal experimenten, $F(2,52) = 5.48, p = .007$. Een post hoc analyse toonde aan dat kinderen uit de taakstructureringsconditie significant meer experimenten opgesteld hadden dan kinderen uit de controle conditie, $p = .005$. Er werd geen significant verschil gevonden tussen de taakstructureringsconditie en de directe instructie, $p = .563$. Ook tussen de directe instructie en de controle conditie werd wederom geen significant verschil gevonden, $p = .176$. De covariaat is voor alle ANOVA's berekend om te corrigeren voor het significante verschil in aantal experimenten, maar dit bleek geen significante verschillen met zich mee te brengen. Omwille van het leesgemak is besloten om deze dan ook niet te vermelden.

Het gebruik van de CVS werd aangeduid door het percentage correcte experimenten in de drie condities. Een ANOVA toonde aan dat de drie condities significant verschilden in het gebruik van de CVS, dus het systematisch experimenteren, $F(2,52) = 5.31, p = .008$. Uit een post hoc analyse bleek dat de kinderen uit de directe instructie systematischer experimenteerden dan kinderen uit de controle conditie, $p = .007$. Het systematisch experimenteren bleek echter niet te verschillen tussen de directe instructie conditie en de taakstructureringsconditie, $p = .101$. Ook werd er geen verschil in systematisch experimenteren gevonden tussen de controle conditie en de taakstructureringsconditie, $p = .999$.

Ook is er gekeken naar de specificiteit van hypothesen. De gemiddelde specificiteit bleek 1.72 (0.71). Uit een ANOVA kwam naar voren dat de drie condities niet significant verschilden met oog op de specificiteit van hun hypothesen, $F(2,52) = 1.38, p = .261$.

Aan het eind van de sessie werd kinderen uit de directe instructie en controle conditie gevraagd om de vraag te beantwoorden. Kinderen uit de taakstructureringsconditie moesten na elke

onderzochte variabele hun antwoord op de vraag noteren. Aan de hand van het antwoord van de kinderen is de geldigheid van conclusies bepaald. De drie condities bleken significant te verschillen wat betreft de geldigheid van conclusies, $F(2,52) = 9.62$, $p < .001$. Vervolgens is een post hoc analyse waaruit bleek dat kinderen uit de directe instructie significant betere conclusies trokken dan kinderen uit de controle conditie, $p = .001$. Verder bleek dat kinderen uit de taakstructureringsconditie ook significant betere conclusies trokken dan kinderen uit de controle conditie, $p = .002$. Er bleek geen significant verschil te zijn tussen de directe instructie conditie en taakstructurering, $p = 1.000$.

Vervolgens is er gekeken naar de volledigheid van de experimenten die de kinderen opstelden. Dit bleek significant te verschillen tussen de drie condities, $F(2,52) = 5.23$, $p = .009$. Uit een post hoc analyse bleek dat de volledigheid van de experimenten significant hoger was in de taakstructureringsconditie dan in de controle conditie, $p = .004$. De volledigheid bleek wederom significant hoger in de directe instructie conditie dan in de controle conditie, $p = 0.014$. Tussen de taakstructureringsconditie en de directe instructie conditie werd echter geen significant verschil gevonden wat betreft volledigheid van experimenten, $p = 1.000$.

Er is ook gekeken naar de volledigheid van conclusies die kinderen trokken. Uit een ANOVA die uitgevoerd is bleek dat deze significant verschilden tussen de drie condities, $F(2,52) = 8.75$, $p < .001$. Uit een post hoc analyse, kwam naar voren dat er een significant verschil bestond tussen de directe instructie conditie en controle conditie, $p = .001$. Verder bleek er ook een significant verschil te bestaan tussen de taakstructureringsconditie en de controle conditie, $p = .004$. Er werd echter geen significant effect gevonden tussen de directe instructie conditie en de taakstructureringsconditie, $p = 1.000$.

De kinderen hebben vervolgens de CVS als natest ingevuld. De scores zijn geanalyseerd aan de hand van een mixed model ANOVA, waarbij voor elke conditie de CVS voortest vergeleken werd met de CVS natest. Hieruit kwam naar voren dat er een significant effect was voor de within-subject factor tijd, $F(1,45) = 25.12$, $p < .001$. Daarnaast bleek er geen significant effect te zijn voor de between-subject factor conditie, $F(2,45) = 0.95$, $p = .395$. Verder is er ook geen significant interactie effect gevonden, $F(2,45) = 0.275$, $p = .761$. Dit betekent dat de kinderen uit alle drie condities hoger scoorden op de CVS natest en dat deze toename in alle drie de condities ongeveer even groot was en er dus in alle condities ongeveer evenveel geleerd is.

Vervolgens is er nog gekeken naar de relaties tussen afhankelijke variabelen uit dit onderzoek. De eenzijdige Pearsons correlatie is weergegeven in Tabel 3.

Tabel 3. *Pearsons Correlatie tussen de afhankelijke variabelen*

	1.	2.	3.	4.	5.	6.	7.
1. Voorkennistest	—						
2. CVS gebruik	0.12	—					
3. Specificiteit hypotheses	0.12	0.20	—				
4. Geldigheid conclusies	0.09	0.54***	0.25	—			
5. Volledigheid experimenten	0.01	0.40**	0.42**	0.39**	—		
6. Volledigheid conclusies	0.12	0.60***	0.27**	0.97***	0.41**	—	
7. CVS natest	0.27	0.54***	0.18	0.40**	0.24	0.41**	—

Note. *** = $p < 0,001$ (tweezijdig)

** = $p < 0,01$ (tweezijdig)

* = $p < 0,05$ (tweezijdig)

Hieruit kwam naar voren dat er geen significante correlaties waren tussen voorkennistest en één van de andere variabelen. Verder bleek dat er significante correlaties waren tussen het CVS gebruik en de andere afhankelijke variabelen, behalve voor de correlatie tussen het CVS gebruik en de specificiteit van hypotheses. De geldigheid van hypotheses correleerde significant met alle andere afhankelijke variabelen. Verder was er in het geval van specificiteit van hypotheses geen significante correlatie met het CVS gebruik en de CVS natest. In beide gevallen van volledigheid van experimenten waren alle correlaties significant. De correlaties voor de CVS natest waren allen significant, behalve de correlatie met de specificiteit van hypotheses.

4. Discussie

Het doel van dit onderzoek was om meer inzicht te krijgen in de relatieve effectiviteit van twee verschillende strategieën voor het verbeteren van de vaardigheden van kinderen in het experimenteren. Deze twee strategieën waren het geven van directe instructie en het structureren van de taak. Ten tweede is gekeken naar de specificiteit van de hypothesen die kinderen formuleerden. Tot slot is gekeken hoe volledig de kinderen waren in het experimenteren.

De eerste hypothese was dat directe instructie en de taakstructurering beiden effectiever zouden zijn dan geen strategie, de controle conditie, in het ondersteunen van kinderen in het experimenteren. Deze verwachting werd niet geheel bevestigd. Weliswaar experimenteerden de kinderen uit de directe instructie conditie systematischer dan kinderen uit de controle conditie, maar het verschil tussen de taakstructurering en controle conditie was niet significant. Daarnaast bleek dat beide experimentele condities tot even systematisch experimenteelgedrag leidde.

De tweede hypothese was dat kinderen uit de taakstructureringsconditie specifiekere hypothesen zouden opstellen dan de kinderen uit de beide andere condities. De resultaten bevestigden deze hypothese echter niet: kinderen uit alle drie de condities hadden even specifieke hypothesen.

De derde hypothese was dat zowel directe instructie als taakstructurering leidt tot betere conclusies dan onderzoekend leren. Verder werd er geen verschil verwacht tussen beide experimentele condities. De resultaten bevestigen deze hypothese.

Tot slot is als vierde hypothese de volledigheid van de experimenten onderzocht. De verwachting was dat kinderen in de taakstructurering conditie meer variabelen zouden noemen in hun experimenten en conclusies, dan kinderen uit de directe instructie en controle conditie. De resultaten toonden aan dat kinderen uit beide experimentele condities vollediger experimenten opstelden dan kinderen uit de controle conditie. De resultaten blijken dus wederom niet geheel in overeenstemming met de verwachting.

Het instrueren van kinderen in de CVS strategie is dus effectiever gebleken dan onderzoekend leren zonder ondersteuning. De goede prestatie van kinderen uit de directe instructie conditie in het systematisch experimenteren, bleek ook uit onderzoek van Lazonder en Egberink (in press). Blijkbaar is het effectief om kinderen de CVS-strategie te onderwijzen en resulteert dit ook in de toepassing van de strategie door kinderen. Dit wordt ook ondersteund door de substantiële correlatie tussen het gebruik van de CVS en de scores op de CVS natest die gevonden is in dit onderzoek.

Er is echter wel verschil gevonden in het percentage systematische experimenten wat Lazonder en Egberink (in press) gevonden hebben en het percentage wat in dit huidige onderzoek gevonden is. Het percentage bedroeg in het onderzoek van Lazonder en Egberink (in press) 69 en in dit huidige onderzoek 75. De standaarddeviatie uit beide onderzoeken is ongeveer gelijk, namelijk respectievelijk 32% en 36%. Dit zou kunnen betekenen dat de directe instructie in dit huidige onderzoek effectiever is geweest. Een verklaring voor dit verschil zou verband kunnen houden met de

manier van aanbieden van de directe instructie. In dit onderzoek is het voorbeeld van Chen en Klahr (1999) gevolgd in het aanbieden van de directe instructie. Zij boden kinderen individueel instructie aan, zodat ze konden controleren of de kinderen de CVS instructie ook werkelijk begrepen. Mogelijk heeft dit huidige onderzoek hier ook voordeel bij gehad. De onderzoekster kon namelijk één op één controleren of kinderen de directe instructie ook werkelijk begrepen en juist toe konden passen. In het onderzoek van Lazonder en Egberink (in press) werd de directe instructie klassikaal aangeboden. Hieruit kan geconcludeerd worden dat het individueel aanbieden van de directe instructie wel degelijk winst heeft boven klassikaal aanbieden van de directe instructie.

Er is in dit onderzoek een nieuwe manier gebruikt om het CVS-gebruik te meten. In dit onderzoek zijn proefjes samengevoegd tot experimenten wanneer ze hetzelfde doel hadden. Vervolgens is aan de hand van deze samengestelde proefjes een CVS-score berekend. Het was de vraag of dit een valide maat was om het CVS-gebruik te meten. De scores op de CVS natest blijken te correleren met de geldigheid van het trekken van conclusies, wat een indicatie is voor de validiteit van het CVS-gebruik (zie Kuhn & Dean, 2005). Belangrijker is echter de substantiële correlatie tussen het gebruik van de CVS en de scores op de CVS natest. Hieruit blijkt dat kinderen die relatief veel systematische experimenten opstelden, ook hoger scoorden op de CVS natest. Kennis en gebruik van de CVS gaan dus hand in hand.

Tegen de verwachting in bleek de taakstructurering in dit huidige onderzoek niet effectiever te zijn in het systematisch experimenteren dan de controle conditie. Een verklaring zou kunnen zijn dat kinderen uit de taakstructureringsconditie de vraag over de eerste variabele niet goed begrepen. Kinderen gaven vaak aan moeite te hebben met de vraag over toonhoogte. In de toekomst zou ook onderzocht kunnen worden of kinderen de variabele toonhoogte uit dit onderzoek te ingewikkeld vinden en dit resulteert in minder systematisch experimenteren. Een tweede verklaring zou gezocht kunnen worden in de manier waarop dit onderzoek opgesteld is. Tijdens het experimenteren moesten kinderen drie vragen van de proefleider beantwoorden. Dit zou ertoe geleid kunnen hebben dat de eerste stap uit het SDDS model, namelijk het focussen op één variabele, gestimuleerd wordt. Lazonder en Egberink (in press) suggereren dat uit deze eerste stap de andere stappen van het SDDS model volgen. Een verklaring zou dus kunnen zijn dat, door de vragen die de proefleider stelt, de kinderen in de controle conditie ook gestimuleerd worden tot het focussen op één variabele, waaruit vervolgens de andere stappen volgen. Hoewel onbedoeld, zou het zo kunnen zijn dat de kinderen uit de controle conditie dus eigenlijk dezelfde benadering gehad hebben als de kinderen uit de taakstructurering.

Wat betreft de geldigheid van het trekken van conclusies is er wel overstemming gevonden in resultaten tussen dit huidige onderzoek en het onderzoek van Lazonder en Egberink (in press). Uit beide onderzoeken is naar voren gekomen dat kinderen uit de twee experimentele condities beter waren in het trekken van geldige conclusies dan kinderen uit de controle conditie. Dit was ook in overeenstemming met het onderzoek van Kuhn en Dean (2005). In dit onderzoek kwam naar voren dat de hint aan kinderen om zich op één variabele te richten effectief was met oog op het trekken van

geldige conclusies. Dit is een vroege versie van taakstructurering, wat uitgebreid is door Lazonder en Kamp (2012). Uit hun onderzoek is wederom gebleken dat taakstructurering een goede strategie kan zijn om de vaardigheden van kinderen te verbeteren in het trekken van geldige conclusies. Zoals blijkt uit het onderzoek van Lazonder en Egberink (in press) en het huidige onderzoek, is ook directe instructie effectief voor het trekken van geldige conclusies. Verder is, net als in het onderzoek van Lazonder en Egberink (in press), ook een significante correlatie gevonden tussen het systematisch experimenteren en het CVS-gebruik. Dit betekent dat kinderen die systematischer experimenteren ook betere conclusies trekken.

Het is natuurlijk de vraag hoe verklaard kan worden dat er voor de taakstructurering geen significant effect gevonden is in het systematisch experimenteren, maar wel voor de geldigheid van conclusies. Een verklaring hiervoor zou mogelijk gevonden kunnen worden in de manier waarop de geldigheid van de conclusies berekend is. Kinderen uit de taakstructureringsconditie kregen bij elk van de vier vragen de mogelijkheid om een antwoord in te vullen, terwijl kinderen uit de directe instructie conditie en de controle conditie hier slechts één kans voor kregen, namelijk helemaal aan het einde van de sessie. Kinderen uit de taakstructureringsconditie hadden dus meer kans om variabelen te noemen en hebben hier mogelijk dus meer punten voor gekregen dan wanneer ze slechts één kans hadden gekregen om de variabelen te noemen. Dit zou een mogelijke verklaring kunnen zijn voor het gevonden resultaat dat taakstructurering wel effectief bleek in het trekken van geldige conclusies, maar niet in het systematisch experimenteren. Een andere verklaring voor dit resultaat zou kunnen zijn dat de geldigheid van conclusies een logisch gevolg is van de volledigheid van experimenten.

Verder heeft dit onderzoek aangetoond dat er geen verschil is in specificiteit van hypothesen tussen de drie condities. Zoals uit de inleiding al blijkt, vormen jonge kinderen vaak geen hypothesen. Vaak zijn ze in staat om hypothesen te vormen wanneer ze de leeftijd van 12-14 jaar bereikt hebben (Penner & Klahr, 1996; Tomkins & Tunnicliffe, 2001). Dit onderzoek is op de basisschool uitgevoerd en uit de resultaten van dit onderzoek is onder andere gebleken dat de kinderen gemiddeld 10.89 jaar waren. De kinderen zijn dus jonger dan de leeftijd waarop gedacht wordt dat kinderen hypothesen gaan vormen. Maar uit dit huidige onderzoek blijkt echter dat kinderen juist heel specifieke hypothesen opstellen. De gemiddelde specificiteit van hypothesen was namelijk 1.72, wat op een schaal van 0 tot 3 redelijk specifiek te noemen is. Dit resultaat doet vermoeden dat kinderen wel in staat zijn om (specifieke) hypothesen op te stellen wanneer ze hiertoe gestimuleerd worden. In dit onderzoek werden kinderen steeds drie vragen gesteld om de specificiteit van hun hypothesen te achterhalen. Een mogelijke verklaring zou dus kunnen zijn dat kinderen specifieke hypothesen opstellen omdat ze door de gestelde vragen gestimuleerd worden dit te doen, maar dit vanuit zichzelf dit niet zouden doen. Voor vervolgonderzoek zou het een aanbeveling zijn om door middel van hardop denken te onderzoeken of kinderen uit zichzelf geen hypothesen opstellen in de leeftijd van 12-14, maar hier wel toe in staat zijn wanneer ze ook gestimuleerd worden om dit te doen.

De onbedoelde *cueing* effecten zouden bovendien een verklaring kunnen zijn waarom de tweede hypothese niet kon worden bevestigd. Verwacht werd dat kinderen uit de taakstructureringsconditie meer gericht zouden zijn op het onderzoeken van specifieke hypothesen, omdat de taakstructurering zou helpen bij het richten van de aandacht op één variabele, wat de eerste stap is uit het SDDS model. Uit dit onderzoek blijkt echter dat kinderen uit de andere condities ook vrij specifieke hypothesen opstellen. Voor kinderen uit de directe instructie conditie valt dit te verklaren doordat zij geleerd hebben om zelf de focus van hun onderzoek te bepalen. Maar voor kinderen uit de controle conditie geldt dit niet; wellicht dat hun specifieke hypothesen veroorzaakt zijn door de vragen die door de proefleidster werden gesteld. Deze vragen--en vooral de derde vraag: 'Waarom denk je dat?'-- stimuleert kinderen om met een verklaring te komen en hun aandacht te richten op één variabele. Om dit effect van de gestelde vraag en de verklaringen die gegeven worden waarom er tussen directe instructie en taakstructurering geen verschil gevonden wordt, te onderscheiden, zou er een minder sturende onderzoeksopzet gebruikt moeten worden om de specificiteit van hypothesen te onderzoeken. De hardop-denken methode biedt hiervoor goede mogelijkheden; het gebruik ervan is een aanbeveling voor toekomstig onderzoek.

Tot slot heeft dit onderzoek gekeken naar de volledigheid van de experimenten. Hieruit kan geconcludeerd worden dat het aanbieden van directe instructie en het structureren van de taak kinderen ondersteunen in het opstellen van volledige experimenten. Een mogelijke verklaring voor het gevonden effect bij de derde hypothese is dat taakstructurering de aandacht richt op de eerste stap van het SDDS model, namelijk het kiezen van een variabele als focus, waardoor kinderen aandacht hebben voor alle variabelen, deze onderzoeken en dus meer variabelen onderzoeken met het experimenteren. De directe instructie ondersteunt deze eerste stap niet expliciet, maar legt kinderen wel de werking van de CVS strategie uit. In de directe instructie conditie worden de drie stappen uit het SDDS model benadrukt en uitgelegd. Deze stappen zijn het kiezen van een variabele als focus, het kiezen van waarden bij deze variabele en het kiezen van waarden bij de overige variabelen. Mogelijk onderzoeken kinderen in deze conditie alle mogelijke variabelen, doordat ze de stappen van het SDDS model volgen. Dit volgen van het SDDS model, door de CVS instructie dus gestimuleerd, zou een verklaring kunnen zijn voor de volledigheid van experimenten bij kinderen uit de directe instructie conditie. Kinderen uit de controle conditie volgen deze stappen, expliciet ondersteund bij de taakstructurering en uitgelegd bij de directe instructie conditie, niet. Dit zou een verklaring kunnen zijn voor het gevonden resultaat dat er meer variabelen genoemd worden door beide experimentele condities. Het zou natuurlijk zo kunnen zijn dat het alleen de eerste stap van het SDDS model is wat hiervoor verantwoordelijk is in plaats van het volgen van alle drie de stappen. Toekomstig onderzoek zou moeten onderzoeken het gevonden effect te verklaren is door het ondersteunen van de eerste stap uit het SDDS model of dat de daaropvolgende stappen hier ook een belangrijke bijdrage aan leveren.

Tot slot is het zeker noemenswaardig om aan te geven dat tijdens het experimenteren naar voren is gekomen dat veel kinderen moeite hadden met de formulering van de vraag, vooral de

zinsnede 'de invloed van'. Daarnaast dachten sommige kinderen dat het de bedoeling was om te onderzoeken wanneer de echo van de gong 10 seconden duurde of wanneer deze het langste duurde. Ze gaven dan als antwoord op de vraag: *'Wat is de invloed van toonhoogte, afstand tot de gong, kijkrichting en volheid van de ruimte op de nagalmtijd van een gong?'* de samenstelling variabelen die 10 seconden duurde of de samenstellingen die het langst duurde. Een aanbeveling van Lazonder en Egberink (in press) is dat bij vervolgonderzoek kinderen er duidelijk op gewezen moeten worden dat ze antwoord moeten geven op de vraag en niet het experiment moeten vermelden dat 10 seconden duurde of het langste experiment moeten vermelden. Bij deze aanbeveling sluit dit huidige onderzoek zich aan.

Ondanks de hierboven besproken beperkingen heeft dit onderzoek ook belangrijke informatie opgeleverd in het debat over de verschillende strategieën in het ondersteunen van kinderen in het experimenteren. Op school worden kinderen al vroeg betrokken bij het onderzoek doen. Vanaf 8 jaar worden kinderen al voorbereid op wetenschappelijk onderzoek door ze dit uit te leggen. Verder doen ze ook proefjes om te oefenen met het ontwerpen en uitvoeren van onderzoek (National Research Council, 1996). ICT kan deze instructies en het doen van proefjes ondersteunen. Uit dit huidige onderzoek is wederom gebleken dat onderzoekend leren zonder ondersteuning niet zo effectief is. Bij de keuze voor een vorm van ondersteuning kunnen leraren ervoor kiezen om een strategie uit te leggen of een taak in kleine stukjes op te delen. Beide manieren zijn effectiever dan geen ondersteuning, en in dit onderzoek bleek directe instructie de meest effectieve manier. Dit heeft implicaties voor de keuze van leraren om te kiezen voor een van beide strategieën. Voor leraren is het instrueren van kinderen in een bepaalde strategie, namelijk vrij eenvoudig toe te passen. Het is echter lastiger voor leraren om een complexe taak op te delen in kleinere stukjes eenvoudiger taken. Dus leraren hoeven op basis van dit onderzoek niet per se een voor hen ingewikkelde taak uit te voeren om kinderen te steunen, maar kunnen kinderen ook eenvoudigweg instrueren.

5. Referentielijst

- Amsel, E., & Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, 11, 523–550. doi: 10.1016/S0885-2014(96)90016-7.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Children's acquisition of the control of variables strategy. *Child Development*, 70(5), 1098 – 1120.
- Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152. doi: 10.1207/s15516709cog0502_2.
- De Wit, J. & van der Veer, G. (1989). Psychologie van de adolescentie. Nijkerk, Intro.
- Dwyer, J. (2007). Computer-based Learning in a Primary School: Differences between the early and later years of primary schooling. *Asia-Pacific Journal of Teacher Education*, 35, 89–103. doi: 10.1080/13598660601111307.
- Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research*, 79, 39-68. doi:10.3102/0034654308325899.
- Galen, F. van, & Gravemeijer, K. (2008). Experimenteren met Grafieken. *Jeugd School en Wereld*, 19, p.p. 16-21.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Kilderry, A., Yelland, N., Lazaridis, V., & Dragicevic, S. (2003). ICT and numeracy in the knowledge era: Creating contexts for new. *Childhood Education*, 79, 293–296.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-48. doi:10.1207/s15516709cog1201_1.
- Klahr, D., Fay, A., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, 25, 111–146. doi: 10.1006/cogp.1993.1003.

Klahr, D., & Li, J. (2005). Cognitive research and elementary science instruction: From the laboratory, to the classroom, and back. *Journal of Science Education and Technology*, 14, 217-238.
doi:10.1007/s10956-005-4423-5.

Koslowski, B., Marasia, J., Chelenza, M., & Dublin, R. (2008). Information becomes evidence when an explanation can incorporate it into a causal framework. *Cognitive Development*, 23, 472-487.
doi:10.1016/j.cogdev.2008.09.007.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. Orlando, FL: Academic Press.

Kuhn, D., & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science*, 16, 866-870. doi:10.1111/j.1467- 9280.2005.01628.x.

Lazonder, A.W. (in press). Chapter 41: Inquiry Learning.

Lazonder, A. W., & Egberink, A.G.G.J. (in press). Children's acquisition and use of the control-of-variable strategy: Effects of explicit and implicit instructional guidance.

Lazonder, A. W., & Kamp, E. (2012). Bit by bit or all at once? Splitting up the inquiry task to promote children's scientific reasoning. *Learning and Instruction*, 22, 458-464.
doi:10.1016/j.learninstruc.2012.05.005.

Metz, K.E. (1995). Re-assessment of developmental assumptions in children's science instruction. *Review of Educational Research*, 65(2), 93-127.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Science Foundation (2000). An introduction to inquiry. *Inquiry: Thoughts, views and strategies for the K-5 classroom* (Vol. 2, pp. 1-5). Washington, DC: National Science Foundation

Penner, D. E., & Klahr, D. (1996). The interaction of domain-specific knowledge and domain-general discovery strategies: A study with sinking objects. *Child Development*, 67, 2709-2727.
doi:10.1111/1467-8624.ep9706244829.

Schauble, L., Glaser, R., Duschl, R.A., Schulze, A., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences*, 4, 131-166. doi: 10.1207/s15327809jls0402_1

Tomkins, S. P., & Tunnicliffe, S. D. (2001). Looking for ideas: Observation, interpretation and hypothesis-making by 12-year-old pupils undertaking science investigations. *International Journal of Science Education*, 23, 791-813. doi: 10.1080/09500690119322.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, 51, 1-10.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27, 172-223. doi: 10.1016/j.dr.2006.12.001.

6. Bijlages

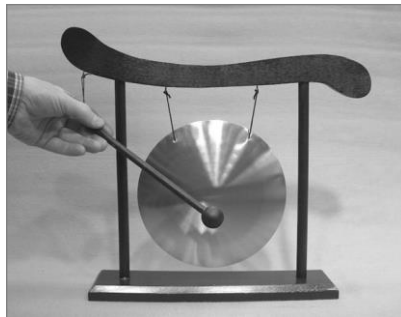
6.1 Voorkennis test

Naam:

Leeftijd:

Jongen / Meisje:

Je staat in een lange gang. In die gang staat een grote gong. Als ik 1 keer op de gong sla, maakt hij geluid. Je hoort dan “boing”. Je kunt dit geluid nog een paar seconden horen. Dat heet de echo. Soms duurt de echo heel lang, soms heel kort. Hoe zou dat komen? Daarover gaat deze test.



Kruis aan welke dingen iets uitmaken voor hoe lang je de echo kunt horen.

Let op. Er kunnen meer goede antwoorden zijn. Dus als jij denkt dat 2 of 3 of misschien wel 4 dingen iets uitmaken, dan kan dat!

Vraag 1

Maakt het uit of de gong een hoog of een laag geluid maakt?

☐ **Nee**, dat maakt niet uit voor hoe lang je de echo kunt horen
→ Ga door met vraag 2

☐ **Ja**, dat maakt wel uit voor hoe lang je de echo kunt horen
→ Omcirkel het goede woord:

Bij een hoge toon hoor je de echo **langer / korter** dan bij een lage toon

Vraag 2

Maakt het uit of de gong voor of achter jou staat?

- ☐ **Nee**, dat maakt niet uit voor hoe lang je de echo kunt horen

→ *Ga door met vraag 3*

- ☐ **Ja**, dat maakt wel uit voor hoe lang je de echo kunt horen

→ *Omcirkel het goede woord:*

Als de gong voor je staat, hoor je de echo **langer / korter** dan als de gong achter je staat

Vraag 3

Maakt het uit of jij dichtbij staat of veraf?

- ☐ **Nee**, dat maakt niet uit voor hoe lang je de echo kunt horen

→ *Ga door met vraag 4*

- ☐ **Ja**, dat maakt wel uit voor hoe lang je de echo kunt horen

→ *Omcirkel het goede woord:*

Als de gong ver weg staat, hoor je de echo **langer / korter** dan als de gong dichtbij staat

Vraag 4

Maakt het uit of de gang leeg is of vol met spulletjes staat?

- ☐ **Nee**, dat maakt niet uit voor hoe lang je de echo kunt horen

→ *Je bent klaar met de test*

- ☐ **Ja**, dat maakt wel uit voor hoe lang je de echo kunt horen

→ *Omcirkel het goede woord:*

Als de gang leeg is, hoor je de echo **langer / korter** dan als de gang vol is

6.2 Coderingsinstructies voorkennistest

Doel en opzet van de test

De voorkennistest is bedoeld om de kennis over het onderwerp van de simulatie te meten. De test bestaat uit 4 vragen, 1 voor elke factor uit de simulatie. Bij elke vraag moet worden aangegeven of de betreffende factor invloed heeft op de nagalmtijd van de gong. Bij een positief antwoord moet bovendien worden aangegeven wat de richting van het effect is. Hierbij wordt steeds een dichotome vergelijking gepresenteerd, ook voor factoren die in de simulatie 3 levels hebben. Tabel 1 geeft een overzicht van de inhoud van test.

Tabel 1

Overzicht van de items uit de CVS-test

Item	Factor	Levels	Effect?	Richting van het effect?
1	Toonhoogte	Hoog, Laag	Ja	Hoge toon = Langere echo
2	Plaats van de gong	Voor, Achter	Nee	N.v.t.
3	Afstand tot de gong	Dichtbij, Veraf	Ja	Veraf = Kortere echo
4	Inrichting van de gang	Leeg, Vol	Ja	Leeg = Langere echo

De test bevat een korte uitleg van het begrip nagalmtijd / echo en een plaatje van een gong. Zie verder in de paragraaf “Instructies”.

Scoring

De antwoorden kunnen op 2 manieren worden gescoord. Bij een grofmazige scoring moet een vraag helemaal goed zijn beantwoord, dus zowel het bestaan als de richting van het effect moeten correct zijn. In dat geval wordt 1 punt gegeven. Dit leidt tot een maximale score van 4 punten. Bij een fijnmazige scoring wordt 1 punt gegeven voor het bestaan van een effect (kolom 4 uit Tabel 1). Indien van toepassing wordt voor de de richting van het effect ook 1 punt gegeven (kolom 5 uit Tabel 1). Het totaal aantal punten komt hiermee op 4+3=7 punten. Beide deelscores kunnen ook apart worden geanalyseerd.

Instructies

-Voordat we beginnen; schrijf allemaal je naam onderaan de eerste bladzijde. En je leeftijd en of je een jongen of een meisje bent.

-We gaan nu deze test maken. En om jullie een beetje te helpen, lees ik de vragen voor. Lees maar mee.

-(Inleiding voorlezen)

-(Plaatje bespreken): Op de foto zie je een gong. Bij deze test moet je aankruisen welke dingen iets uitmaken voor hoe lang je de echo van de gong kunt horen. (rest voorlezen).

-(Vragen 1 voor 1 voorlezen.)

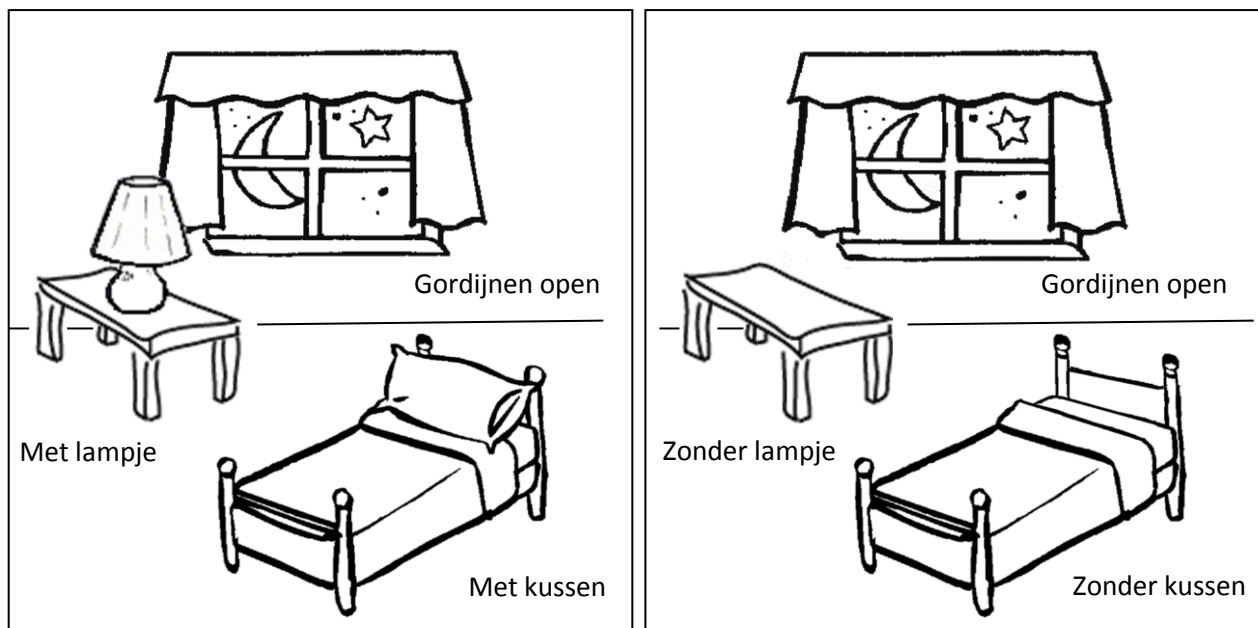
6.3 CVS test

Voorbeeld

Henk kan soms niet goed slapen.

Hij wil weten of dat komt door zijn kussen.

In de plaatjes zie je hoe hij dat heeft onderzocht.



Is dit een goede test? Weet Henk straks zeker of zijn kussen iets uitmaakt voor hoe goed hij slaapt? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

Naam:

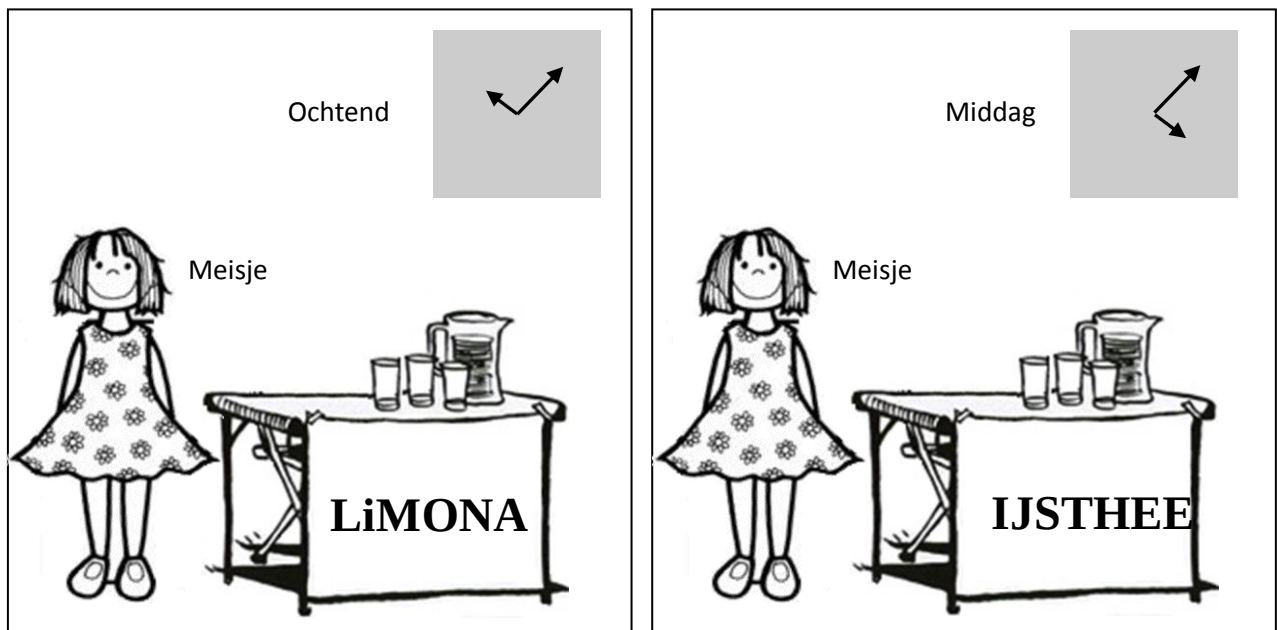
Leeftijd:

Jongen / Meisje:

Vraag 1

Meneer Mulder verkoopt drankjes.
Maar hij verkoopt niet altijd even veel.

Meneer Mulder wil weten of dat komt door het moment van de dag (ochtend of middag)
In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet meneer Mulder straks zeker of het moment van de dag iets uitmaakt voor hoeveel drankjes er worden verkocht? Omcirkel het goede antwoord:

JA, dit is een goede test

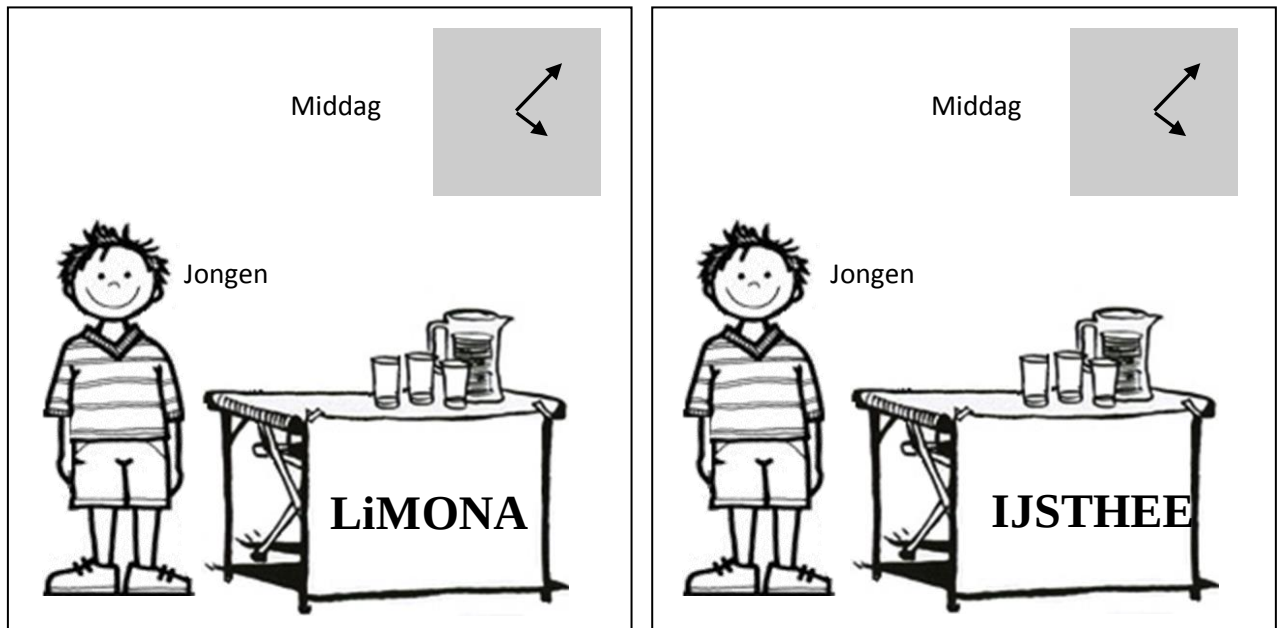
NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

Vraag 2

Meneer Mulder verkoopt drankjes.
Maar hij verkoopt niet altijd even veel.

Meneer Mulder wil weten of dat komt door het soort drankje (limonade of ijsthee).
In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet meneer Mulder straks zeker of het soort drankje iets uitmaakt voor hoeveel drankjes er worden verkocht? Omcirkel het goede antwoord:

JA, dit is een goede test

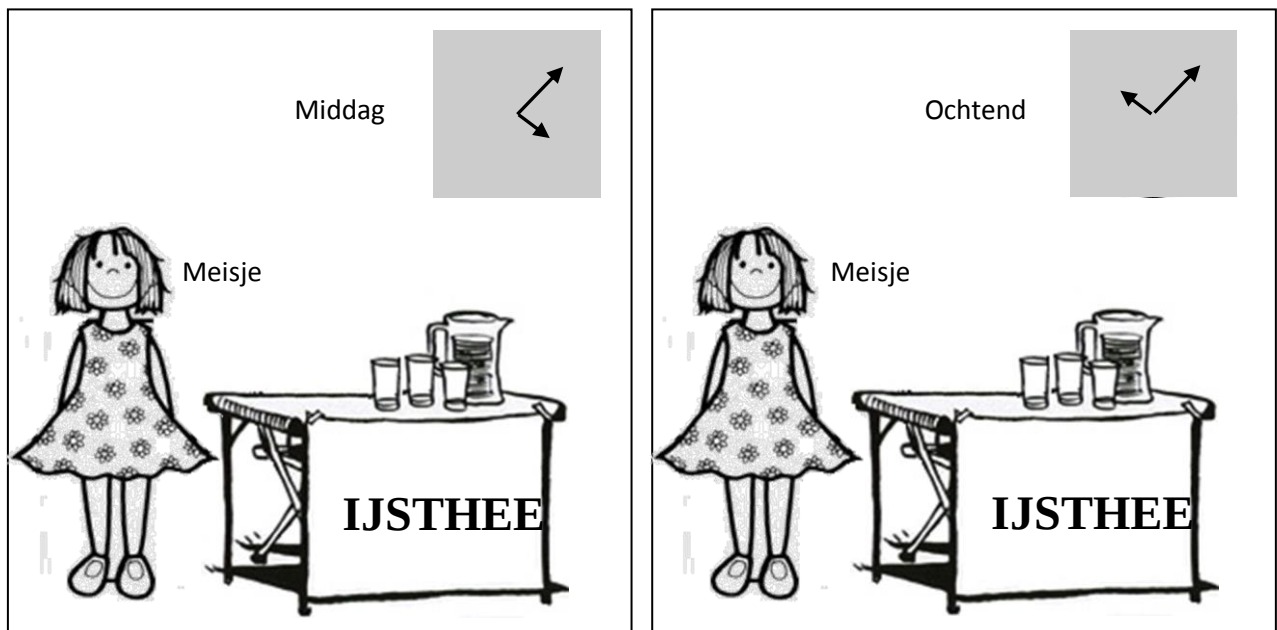
NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

Vraag 3

Meneer Mulder verkoopt drankjes.
Maar hij verkoopt niet altijd even veel.

Meneer Mulder wil weten of dat komt door **de verkoper** (jongen of meisje).
In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet meneer Mulder straks zeker of **de verkoper** iets uitmaakt voor hoeveel drankjes er worden verkocht? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

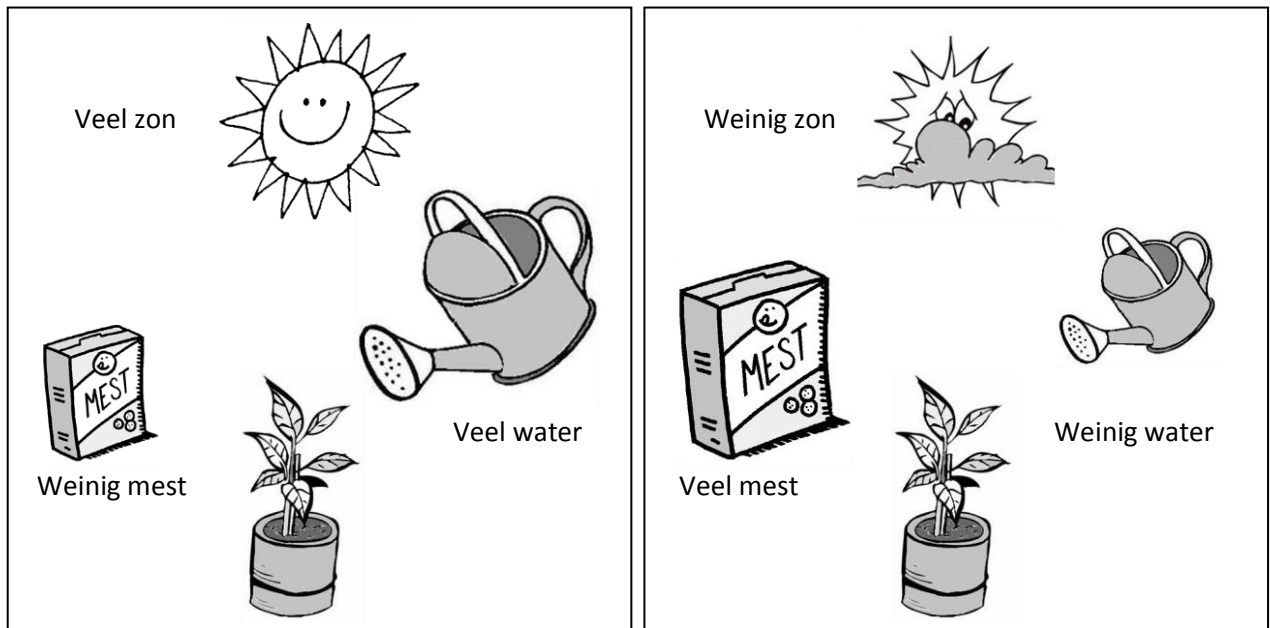
Vraag 4

Mieke heeft veel planten.

Maar die groeien niet altijd even goed.

Mieke wil weten of dat komt door **de hoeveelheid zonlicht**.

In de plaatjes zie je hoe zij dat heeft getest.



Is dit een goede test? Weet Mieke straks zeker of **de hoeveelheid zonlicht** iets uitmaakt voor de groei van het plantje? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

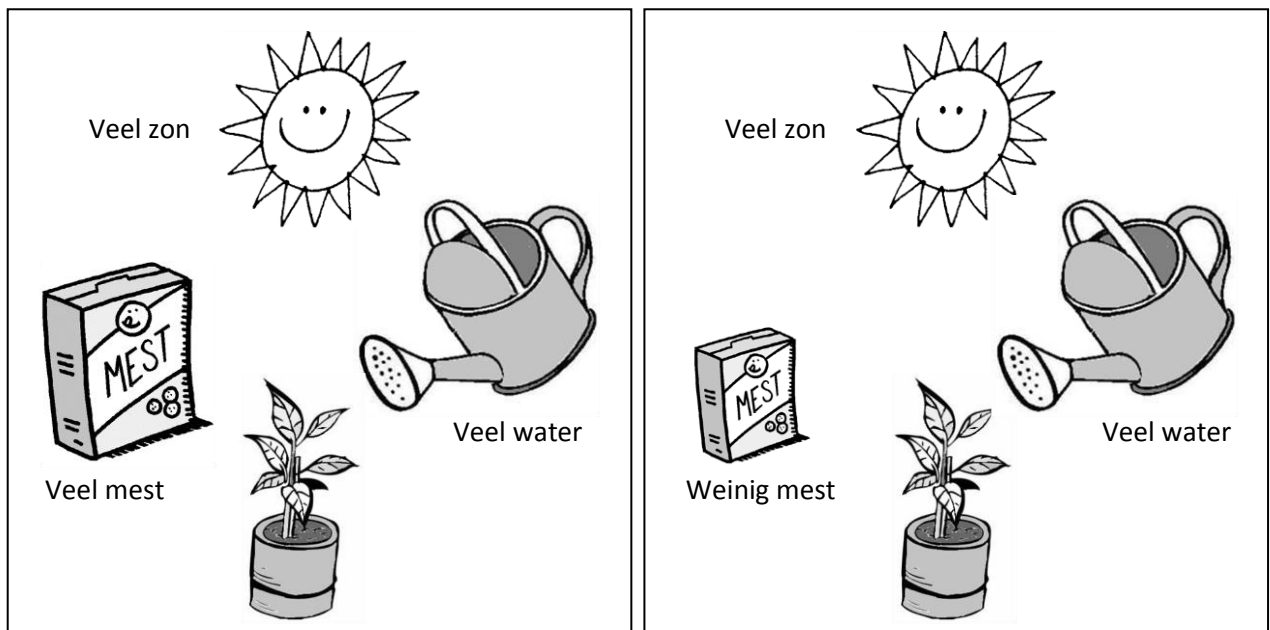
Vraag 5

Mieke heeft veel planten.

Maar die groeien niet altijd even goed.

Mieke wil weten of dat komt door **de hoeveelheid mest**.

In de plaatjes zie je hoe zij dat heeft getest.



Is dit een goede test? Weet Mieke straks zeker of **de hoeveelheid mest** iets uitmaakt voor de groei van het plantje? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

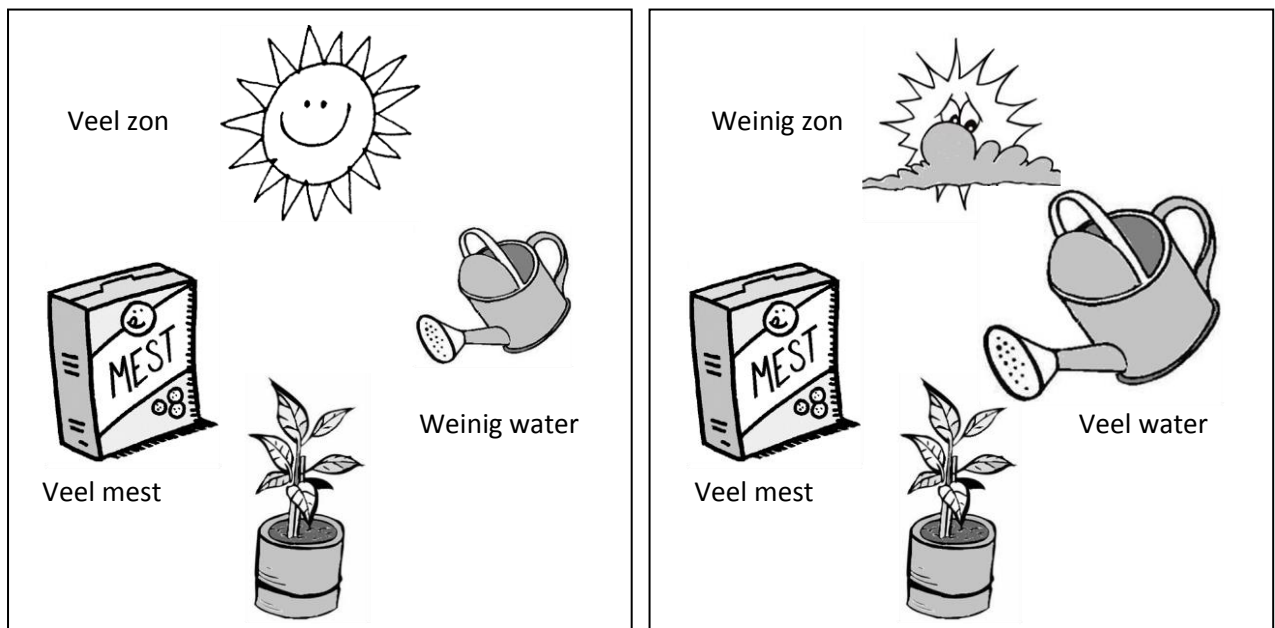
Vraag 6

Mieke heeft veel planten.

Maar die groeien niet altijd even goed.

Mieke wil weten of dat komt door **de hoeveelheid water**.

In de plaatjes zie je hoe zij dat heeft getest.



Is dit een goede test? Weet Mieke straks zeker of **de hoeveelheid water** iets uitmaakt voor de groei van het plantje? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

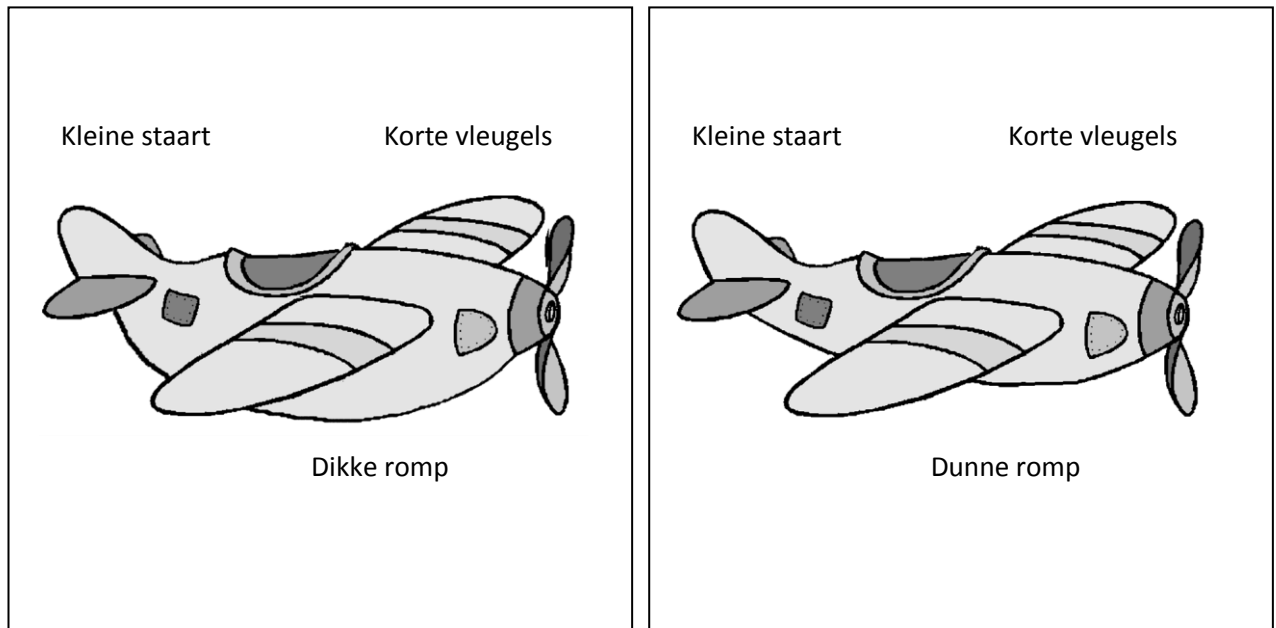
Vraag 7

Peter maakt vliegtuigjes.

Maar die vliegen niet altijd even snel.

Peter wil weten of dat komt door **de vorm van de romp**.

In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet Peter straks zeker of **de vorm van de romp** iets uitmaakt voor de snelheid van het vliegtuigje? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

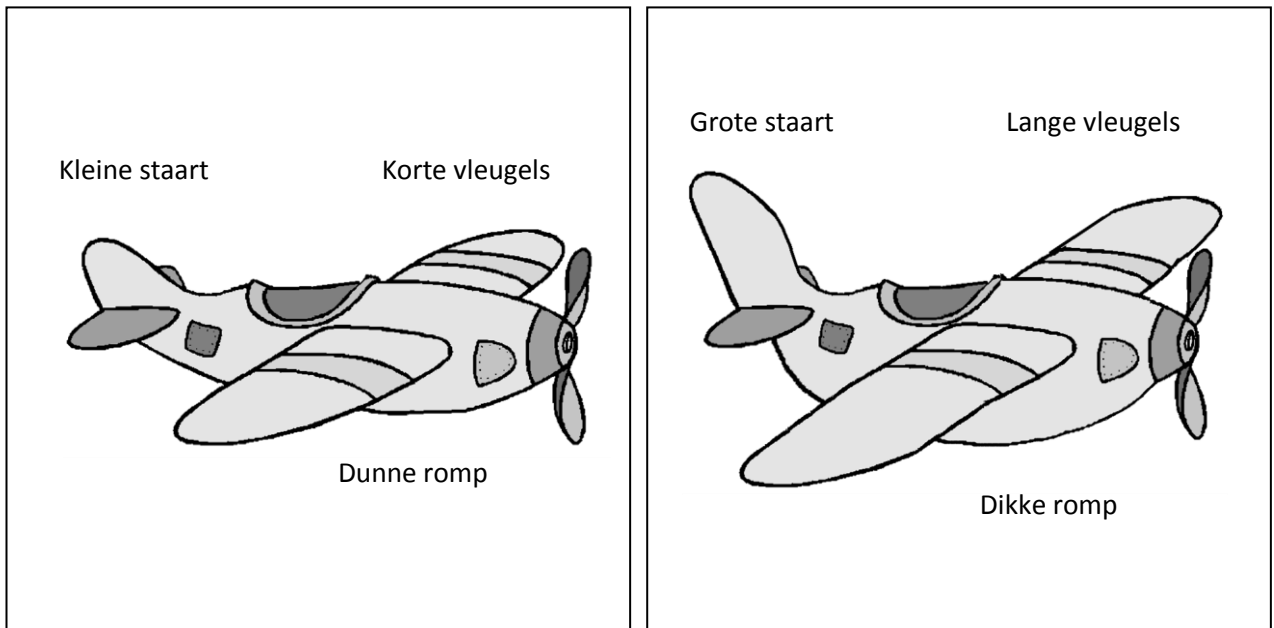
Vraag 8

Peter maakt vliegtuigjes.

Maar die vliegen niet altijd even snel.

Peter wil weten of dat komt door de lengte van de vleugels.

In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet Peter straks zeker of de lengte van de vleugels iets uitmaakt voor de snelheid van het vliegtuigje? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

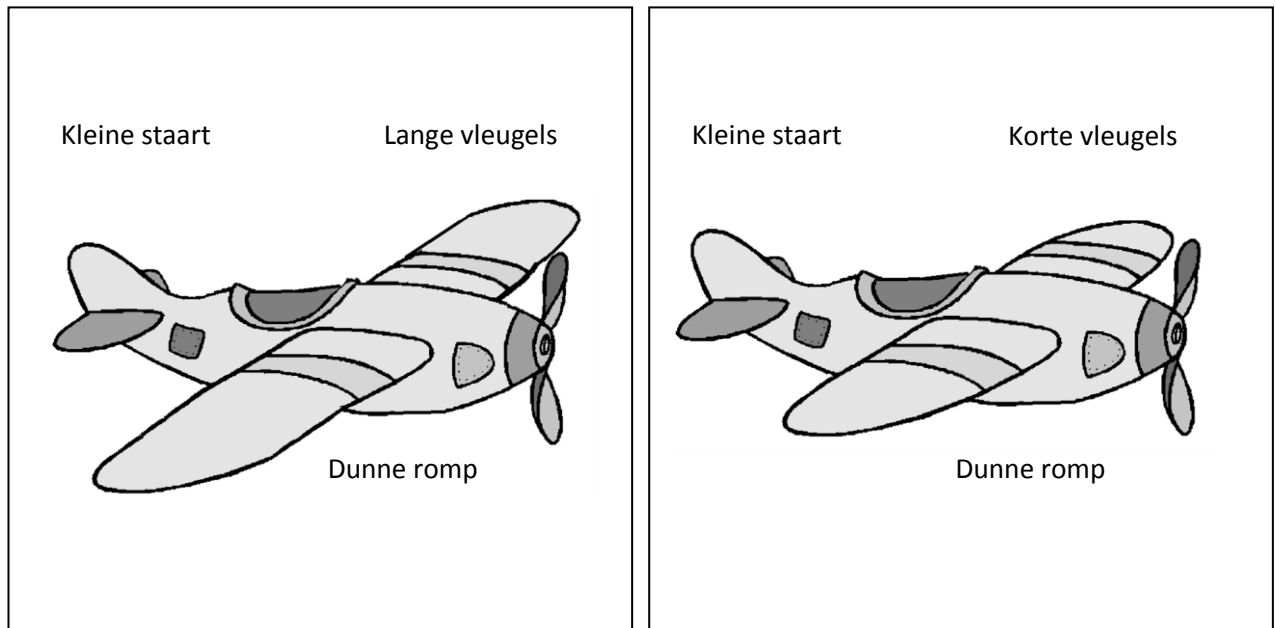
Vraag 9

Peter maakt vliegtuigjes.

Maar die vliegen niet altijd even snel.

Peter wil weten of dat komt door de lengte van de staart.

In de plaatjes zie je hoe hij dat heeft getest.



Is dit een goede test? Weet Peter straks zeker of de lengte van de staart iets uitmaakt voor de snelheid van het vliegtuigje? Omcirkel het goede antwoord:

JA, dit is een goede test

NEE, dit is een slechte test

Is jouw antwoord NEE? Verander dan de plaatjes om er een goede test van te maken.

6.4 Coderingsinstructies CVS-test

Doel en opzet van de test

De CVS test is bedoeld om de vaardigheid in het systematisch experimenteren te meten. De test is een vrijwel letterlijke vertaling van de CVS-test uit de onderzoeken van Klahr (zie Chen & Klahr, 1999; Klahr, Chen, & Toth, 2001; Matlen & Klahr, in press; Strand-Carey & Klahr, 2008; Siler et al., 2009). De CVS-test bestaat uit 9 items die betrekking hebben op 3 verschillende domeinen, namelijk limonade verkopen, plantengroei en vliegtuigjes. Per domein worden 3 vragen gesteld waarin steeds een vergelijking tussen 2 situaties moet worden beoordeeld. In 1 van die vragen is sprake van een systematisch experiment; in de andere 2 vragen is het experiment niet systematisch. Hierbij kan sprake zijn van 3 verschillende soorten fouten, die elk 2 keer voorkomen en random zijn verdeeld over de 3 domeinen. De fouten zijn: (1) er wordt behalve de doelvariabele nog 1 andere variabele veranderd, (2) er worden behalve de doelvariabele nog 2 andere variabelen veranderd, (3) er wordt 1 variabele veranderd, maar dat is niet de doelvariabele. Tabel 1 geeft een overzicht van de inhoud van de 9 items.

Tabel 1
Overzicht van de items uit de CVS-test

Item	Domein	Doelvariabele	Situatie 1	Situatie 2	Soort fout
1	Limonade	Tijdstip	Ochtend, Limonade, Meisje	Middag, IJsthee, meisje	(1)
2	Limonade	Soort drankje	Middag, Limonade, Jongen	Middag, IJsthee, boy	Geen
3	Limonade	Geslacht	Middag, IJsthee, Meisje	Ochtend, IJsthee, Meisje	(3)
4	Planten	Zonlicht	Veel zon, Weinig mest, Veel water	Weinig zon, Veel mest, Weinig water	(2)
5	Planten	Mest	Veel zon, Veel mest, Veel water	Veel zon, Weinig mest, Veel water	Geen
6	Planten	Water	Veel zon, Veel mest, Weinig water	Weinig zon, Veel mest, Veel water	(1)
7	Vliegtuigjes	Romp	Dikke romp, Korte vleugels, Kleine staart	Dunne romp, Korte vleugels, Kleine staart	Geen
8	Vliegtuigjes	Vleugels	Dunne romp, Korte vleugels, Kleine staart	Dikke romp, Lange vleugels, Grote staart	(2)
9	Vliegtuigjes	Staart	Dunne romp, Lange vleugels, Kleine staart	Dunne romp, Korte vleugels, Kleine staart	(3)

De goede antwoorden op de vragen worden getoond in Tabel 2

Tabel 2

Antwoordsleutel CVS-test

Item	Goede antwoord	Verbetering(en) om het een “goede test” te maken
1	Nee	Het soort drankje moet gelijk zijn
2	Ja	n.v.t.
3	Nee	Geslacht moet verschillend zijn én tijdstip van de dag moet gelijk zijn
4	Nee	Hoeveelheid mest moet gelijk zijn én hoeveelheid water moet gelijk zijn
5	Ja	n.v.t.
6	Nee	De hoeveelheid zonlicht moet gelijk zijn
7	Ja	n.v.t.
8	Nee	Vorm van de romp moet gelijk zijn én lengte van de staart moet gelijk zijn
9	Nee	Lengte van de vleugels moet gelijk zijn én lengte van de staart moet verschillend zijn

De test bevat een voorbeeldvraag die wordt gebruikt om de kinderen uit te leggen hoe het format van de test in elkaar zit en hoe ze de vragen moeten beantwoorden. Deze voorbeeldvraag komt uit een ander domein (slaapproblemen). Zie verder in de paragraaf “Instructies”.

Scoring

De antwoorden kunnen op 2 manieren worden gescoord. Bij een grofmazige scoring, die wordt gebruikt in de studies van David Klahr, moet een vraag helemaal goed zijn beantwoord, inclusief de eventuele verbetering(en). In dat geval wordt 1 punt gegeven. Dit leidt tot een maximale score van 9 punten. Bij een fijnmazige scoring wordt 1 punt gegeven per goed aangekruist antwoord (kolom 2 uit Tabel 2). Indien van toepassing worden voor de correcte verbetering(en) ook 1 punt gegeven (kolom 3 uit Tabel 2). Het totaal aantal punten komt hiermee op $9+6=15$ punten. Let op: als bij een vraag 2 verbeteringen nodig zijn, moeten beide verbeteringen correct zijn om 1 extra punt te verdienen; is slechts 1 correcte verbetering aangebracht dan wordt geen extra punt gegeven.

In het Engels: Story responses received one point if children correctly identified experiments as either consistent or inconsistent with CVS, and an additional point if they correctly modified incorrect experiments. All other responses were assigned a 0. Scores ranged from 0 to 15. (Bron: Matlen & Klahr, in press).

Betrouwbaarheid en interne consistentie

Berekend met de voortoetsscores van Angelique (N = 65).

Bij grofmazige codering: Cohen's $\kappa = .996$. Interne consistentie: Cronbach's $\alpha = .84$ (kan .85 worden als vraag 7 wordt verwijderd; heeft een lage item-rest correlatie van .208).

Bij fijnmazige codering: Cohen's $\kappa = .979$. Interne consistentie: Cronbach's $\alpha = .83$ (kan .84 worden als vraag 7 wordt verwijderd; heeft een lage item-rest correlatie van .179).

Bij berekening met natoets scores van Angelique (N = 67)

Grofmazige codering: Cohen's $\kappa = .997$. Interne consistentie: Cronbach's $\alpha = .86$. Kan niet verbeterd worden.

Fijnmazige codering: Cohen's $\kappa = .998$. Interne consistentie: Cronbach's $\alpha = .84$. Kan niet verbeterd worden.

Instructies

-(Indien mogelijk de voorbeeldvraag via het digibord laten zien)

-Voordat we beginnen; schrijf allemaal je naam onderaan de eerste bladzijde. En je leeftijd en of je een jongen of een meisje bent.

-We gaan nu de test maken. En om jullie een beetje te helpen, doen we de eerste vraag samen. Dan weet je een beetje hoe de vragen in elkaar zitten en hoe je antwoord moet geven. Kijk op de eerste bladzijde naar de vraag. En lees maar mee.

-(Eerste 3 zinnen voorlezen, kinderen laten meelezen)

-(Plaatjes bespreken): Henk gaat een paar dagen slapen met de gordijnen open, lampje aan en een kussen. Dat zie je in het linker plaatje. Elke ochtend schrijft hij op of hij goed heeft geslapen. Daarna gaat Henk een paar dagen slapen zoals in het rechter plaatje. Dus met de gordijnen open, zonder lampje en zonder kussen. Kijk maar even goed naar de plaatjes.

-Onder de plaatjes staat de vraag. (voorlezen). Als je denkt dat het een goede test is, dat Henk zo 100% zeker weet dat zijn kussen iets te maken heeft met of hij goed kan slapen, dan zet je een cirkel om het antwoord JA. Vind je het een slechte test, dan moet je het antwoord NEE omcirkelen. Maar let op, je mag bij elke vraag steeds maar 1 antwoord geven. Snappen jullie dat? Dan mag je nu jouw antwoord omcirkelen.

-Als je het antwoord JA hebt omcirkeld ben je klaar en kun je verder met de volgende vraag. Is jouw antwoord NEE, dan moet je ook aangeven hoe je de slechte test kunt verbeteren. Dat kun je doen door dingen in de plaatjes te veranderen. Als je bijvoorbeeld denkt dat de gordijnen dicht moeten, dan schrijf je dat in beide plaatjes erbij. Of als je vindt dat het lampje altijd aan moet, dan teken je in het rechter plaatje een lampje.

-Hebben jullie nog vragen? Nee? Dan kunnen jullie nu beginnen met de toets. Heel veel succes. En denk er aan: lees de vragen heel goed door en kijk goed naar de plaatjes.

Belangrijk: vertel niet wat het goede antwoord is. Vertel ook niet dat de test bedoeld is om de vaardigheid in systematisch experimenteren te meten, en evenmin hoe dit is behandeld tijdens de instructie en het doen van de proefjes met de simulatie.