



Effecten van hulp bij het toetsen van ontwerp- en onderzoeksvaardigheden

Ontwerp, gebruik en resultaten van dynamische toetsen in de context van een peilingsonderzoek Natuur en Techniek

Januari 2017

© Cito, Arnhem

Erik Roelofs

Inhoud

1	Achtergrond, theorie en onderzoeksdoel	1
1.1	Theoretische achtergrond	1
1.2	Context en onderzoeksvragen.....	1
2	Onderzoeksmethode.....	3
2.1	Dataverzameling.....	3
2.1.1	Keuze van praktische opdrachten	3
2.1.2	Afname, begeleiding en scoring van de praktische opdrachten	9
2.1.3	Kennistoets.....	10
2.1.4	Leerlingvragenlijsten: attitude, buitenschoolse activiteiten en zelfbeoordeling.....	11
2.1.5	Subjecten.....	13
3	Resultaten	14
3.1	De verleende hulp aan leerlingen	14
3.2	Leerlingenmerken en prestaties in de conditie zonder hulp	21
3.3	Effecten van hulp.....	22
3.3.1	Overall effecten van hulp	22
3.3.2	Differentiële effecten van hulp	27
3.3.3	Effecten van verschillende soorten hulp en hulpcombinaties	32
3.3.4	Verschillen tussen leerlingen bij wie wel en geen modeling plaatvond na inhoudelijke hulp	34
3.3.5	Leerlingkenmerken en effecten van hulp.....	37
4	Conclusies.....	39
4.1	Soorten hulp (onderzoeksvraag 1)	39
4.2	Soorten hulp naar leerlingkenmerken (onderzoeksvraag 2).....	39
4.3	Leerlingkenmerken en prestaties op de praktische opdrachten in de conditie zonder hulp (onderzoeksvraag 3. deelvraag i)	40
4.4	Overall effecten van hulp (onderzoeksvraag 3. deelvraag ii).....	40
4.5	Differentiële effecten van hulp (Onderzoeksvraag 3. deelvraag iii).....	41
4.6	Effectiviteit van specifieke typen hulp en combinaties van hulp (Onderzoeksvraag 3. deelvraag iv)	41
4.7	Samenhang tussen leerlingkenmerken en prestatieverbetering (onderzoeksvraag 4)	42
5	Discussie en aanbevelingen.....	45
5.1	Een terugblik op de resultaten	45
5.1.1	Aan leerlingen verstrekte hulp	45
5.1.2	Effecten van hulp op prestatieverbetering	45

5.2	Perspectieven op verdere ontwikkeling en theorievorming.....	46
5.3	Aanbevelingen: ontwerpen van ontwikkelingsgerichte prestatiematen.....	48
5.4	Methodologische beperkingen en aanbevelingen.....	49
6	Referenties	49
7	Bijlagen	51
Bijlage 7.1	Ontvangen hulpvormen in percentages leerlingen.....	51
7.1.1	Ontvangen hulpvormen bij vier praktische opdrachten	51
7.1.2	Ontvangen hulpvormen bij deelopdrachten van de Knikkerbaan	51
7.2	Prestaties voor en na hulp van p10, p50 en p90 leerlingen.....	52
7.2.1	Toelichting op de totstandkoming van p-waarden.	52
7.2.2	P-waarden van prestatie-indicatoren op de opdracht Brug voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore.....	54
7.2.3	P-waarden van prestatie-indicatoren op de opdracht Fietsbel voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore	54
7.2.4	P-waarden van prestatie-indicatoren op de opdracht Bibberspel voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore	55
7.2.5	P-waarden van prestatie-indicatoren op de opdracht Knikkers te koop voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore	55
7.2.6	P-waarden van prestatie-indicatoren op de opdracht Knikkerbaan voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore	56

1 Achtergrond, theorie en onderzoeksdoel

1.1 Theoretische achtergrond

Een van de problemen met gestandaardiseerde en uniforme toetsen is dat ze weinig informatie opleveren over leerlingen die (ver) achterblijven in hun ontwikkeling. Critici geven aan dat uniforme toetsen vooral laten zien waartoe leerlingen niet in staat zijn, maar dat onduidelijk is wat ze wel zouden kunnen met extra didactische ondersteuning. Dit laatste is wat vastgesteld wordt bij gebruik van dynamische toetsen (DT). Deze beogen het leerpotentieel van leerlingen te meten op een bepaald leergebied (Lidz & Elliott, 2000). Het leerpotentieel wordt gezien als het verschil tussen waartoe een leerling zelfstandig in staat is en wanneer hij/zij hulp krijgt van een deskundige (volwassene). Anders gezegd, bij dynamische toetsing wordt de zone van naaste ontwikkeling van leerlingen onderzocht. Hierbij voert een leerling toetstaken uit, begeleid door een getrainde toetsleider, die waar nodig hulp op maat biedt.

Voorstanders van DT benadrukken dat door deze vorm van toetsing verschillen tussen leerlingen zichtbaar worden die anders verborgen zouden zijn gebleven. Twee leerlingen kunnen gelijk presteren op een statische toets, die aangeeft wat ze tot nu toe geleerd hebben, maar heel verschillend op een dynamische toets, als het gaat om het potentieel om nieuwe dingen te leren. Recent onderzoek laat zien dat dynamische toetsen een duidelijke meerwaarde hebben ten opzichte van statische toetsen in het correct voorspellen van leerlingen die risico lopen achterop te raken in hun ontwikkeling (context van aanvankelijk lezen: Swanson, 2011; context van wiskundige woordproblemen: Fuchs et al., 2011).

Bij dynamische toetsing is sprake van actieve interactie tussen de toetsleider en de leerling, die erop gericht is de leerling vooruit te helpen bij het uitvoeren van toetstaken. Daarbij wordt precies die mate van hulp gegeven die de leerling nodig heeft om verder te komen met de taak. Ondertussen worden alle vormen van hulp en de reacties van de leerling, waaronder het type gemaakte fouten en benodigde toetstijd, geregistreerd. Behalve vaststellen van het leerpotentieel biedt deze toetsvorm handvatten voor het bepalen van het te geven vervolgonderwijs, aansluitend bij wat zich tijdens dynamische toetsing voordoet.

Bij dynamische toetsing kan volgens twee modellen hulp worden verstrekt (Sternberg & Grigorenko, 2002): het interventiemodel en het interactiemodel. In het interventiemodel wordt gekozen uit vooraf gestandaardiseerde vormen van minder of meer expliciete hulp, volgens een vooraf bepaalde hiërarchie van hints. Bij het interactionele model van DT wordt de traditionele rolverdeling tussen toetsleider en kandidaat min of meer losgelaten en vervangen door een docent-leerling relatie (Feuerstein, Rand & Hoffman, 1979). De mediator past niet alleen de stimuli of de taak aan, maar laat de interactie met de leerling meer spontaan verlopen (Feuerstein et al, 1979).

1.2 Context en onderzoeksvragen

In de context van de nieuwe opzet van periodieke onderwijspeilingen in Nederland, waarvoor de Nederlandse onderwijsinspectie eindverantwoordelijk is, heeft Cito een meer dynamische toetsvorm ontwikkeld, in aanvulling op bestaande uniforme (statische) vormen van toetsing. Het idee hierachter is dat fijnmazige informatie wordt verkregen over de scholen en leerlingen.

In dit rapport wordt verslag gedaan van de aard van de hulpinterventies die toetsleiders verstrekken bij de toepassing van de dynamische toetsvorm en de invloed ervan op de kwaliteit van het ontwerp- en onderzoeksgedrag van de leerlingen.

De volgende onderzoeksvragen en deelvragen staan centraal.

1. Welke vormen van hulp worden bij de begeleiding van praktische ontwerp- en onderzoeksopdrachten verstrekt om leerlingen een stap verder te brengen in de oplossing van een onderzoeks- of ontwerpprobleem?

2. In hoeverre hangen verschillen in verleende typen hulp samen met demografische kenmerken, te weten leeftijd, geslacht, sociaal-economische achtergrond en psychologische leerlingkenmerken, te weten attitude ten aanzien van Natuur en Techniek, en kennis van Natuur en Techniek?
3. In hoeverre gaat de gegeven hulp gepaard met prestatieverbetering.
 - (i) In hoeverre verschillen de effecten van hulp naar het niveau van presteren vóórdat hulp werd gegeven?
 - (ii) Wat is het overall effect van hulp op de hele opdracht en op de onderdelen van de opdracht?
 - (iii) In hoeverre profiteren leerlingen verschillend van hulp?
 - (iv) Verschillen de effecten naar het type verleende hulp of de combinatie van hulptypen?
4. Bestaat er een verband tussen achtergrondkenmerken en psychologische kenmerken van leerlingen en de geboekte prestatieverbetering bij hulp?

2 Onderzoeksmethode

2.1 Dataverzameling

2.1.1 Keuze van praktische opdrachten

Voor het doel van deze studie zijn de resultaten gebruikt van een aantal individueel afgenomen (deel)opdrachten, zoals die aan de orde kwamen in drie ontwerp opdrachten en twee onderzoeksoopdrachten van de peilingsstudie (Cito, 2016). In Tabel 1 is in de kolom 'gebruikt in deelstudie' aangegeven om welke deelopdrachten het gaat.

Tabel 1: Gekozen deelopdrachten voor onderhavige studie

Opdracht	Aantal prestatie-indicatoren	Gebruikt in deelstudie?
1. Brug: een dragende brug maken van papier	8	x
2. Fietsbel		
Deelopdracht 1: losse onderdelen van een fietsbel uit onderdelen in elkaar zetten tot een functionerende fietsbel	4	x
Deelopdracht 2 de fietsbel correct monteren op een buisje van passende dikte	2	
3. Bibberspel	7	x
Deelopdracht 1: in elkaar zetten Bibberspel uit de onderdelen, daarbij een gesloten stroomcircuit aanleggen, zodat spel functioneert	2	x
Deelopdracht 2: herstellen één mankement: het lampje is losgedraaid	1	
Deelopdracht 3: herstellen twee mankementen: - Er is een lege batterij aangesloten - de gekleurde draad is aangesloten op de isolator op de spiraal linksonder.	2	
Deelopdracht 4: verslag leggen van de verrichte reparaties	2	
4. Knikkers te koop	14	x
Deelopdracht 1: meten van afmetingen en gewicht knikkers	2	
Deelopdracht 2: meten van afmetingen en gewicht knikkers	4	
Deelopdracht 3: metingen deelopdracht 2 weergeven in een diagram	7	x
Deelopdracht 3 Voorspellen van het gewicht van een bakje knikkers	1	
5. Knikkerbaan	37	x
Deelopdracht 1: voorspellen van de uitkomst van een experiment	4	
Deelopdracht 2: uitvoeren van een experiment met ingestelde knikkerbaan	5	
Deelopdracht 3: vragen over wat een experiment is	1	
Deelopdracht 4: klaarzetten experimentele opstellingen volgens een getekend voorbeeld	4	
Deelopdracht 5a: klaarzetten experimentele opstelling volgens opgegeven vraagstelling	4	x
Deelopdracht 5b: voorspelling doen over uitkomst	1	x
Deelopdracht 5c: verrichten van metingen volgens de vraagstellingen	3	x
Deelopdracht 5d: conclusies trekken	1	x
Deelopdracht 6a: klaarzetten experimentele opstelling volgens een eigen vraagstelling	4	x
Deelopdracht 6b: voorspelling doen over uitkomst	1	x
Deelopdracht 6c: verrichten van metingen volgens de vraagstellingen	3	x
Deelopdracht 6d: conclusies trekken	1	x
Deelopdracht 7: schematische experimentele opstelling interpreteren	5	

De reden om te kiezen voor juist deze deelopdrachten is dat ze alle een vergelijkbare structuur bezitten. Het gaat daarbij steeds om één complete, afgeronde (deel)opdracht, waarbij sprake kan zijn van hulp door de toetsleider. Voorbereidende, kleinere deelopdrachten of losse 'trouble shooting'-opdrachten zijn niet in deze deelstudie betrokken. De zesde individuele praktische opdracht 'Kopje onder' die in de Peilingsstudie is afgenomen, heeft een afwijkende structuur vergeleken de eerste vijf en is om die reden niet opgenomen in de deelstudie. Deze opdracht bestaat uit een aantal korte

deelopdrachten die elk ondersteunende kennis opleveren (scaffolds) voor het correct voorspellen van het resultaat van de eindopdracht (Cito, 2016).

De gekozen opdrachten bestrijken verschillende deelvaardigheden van ontwerpen en onderzoeken, zoals weergegeven in Tabel 2 en 3. Bij de operationalisering van de onderliggende ontwerpvaardigheden is voortgeborduurd op de methodiek van de VROO (Vaardigheden Rubrics Onderzoeken en Ontwerpen; Van Keulen en Slot, 2013), welke bestaat uit een verzameling observatie-rubrics voor de deelvaardigheden van ontwerpen en onderzoeken. Deze rubrics zijn per opdracht uitgewerkt naar beoordelingsuitspraken voor de handelingen en tussenproducten van leerlingen (zie bijvoorbeeld Tabel 12a).

Tabel 2: Onderliggende ontwerpvaardigheden en domeinen bij de drie onderzochte ontwerp opdrachten

Ontwerpvaardigheden	Brug	Fietsbel	Bibberspel
D1. Een probleem herkennen en verkennen			
D1.1 Probleem onderkennen			
D1.2 Voorkennis gebruiken			
D1.3 Probleem verkennen en verwoorden	X		
D1.4 Probleem afbakenen			
D1.5 Programma van eisen			
D2 Ontwerpen van de oplossing			
D2.1 Oplossing bedenken	X		
D2.2 Oplossing kiezen	X		
D2.3 Werkplan maken			
D3 Het ontwerp uitvoeren			
D3.1 Materialen en gereedschap gebruiken	X	X	X
D3.2 Maken van het ontwerp	X	X	X
D4 Het ontwerp uittesten en verbeteren			
D4.1 Testen van het ontwerp	X	X	X
D4.2 Trouble shooting	X	X	X
D4.3 Herontwerp	X		X
D5 Presenteren			
D5.1 Oplossing presenteren			
D5.2 Verantwoorden			
D5.3 Delen			
Domeinen			
Natuurkunde			
Elektriciteit			x
Techniek			
Eigenschappen van materialen			
Constructies en verbindingen	x	x	
Overbrengingen		x	
Geautomatiseerde systemen			x

Tabel 3: Onderliggende Onderzoeksvaardigheden en domeinen bij de twee onderzochte onderzoeksoopdrachten

	Knikkerbaan	Knikkers te koop
Onderzoeksvaardigheden		
I1 Verwonderen en verkennen		
I1.1 Vragen stellen		
I1.2 Voorkennis gebruiken		
I1.3 Probleem verkennen		
I1.4 Probleem afbakenen		
I1.4a Onderscheiden onderzoeksvariabelen		
I1.4b Formuleren onderzoeksvraag	X	
I1.5R Verwachtingen formuleren: voorspellen/(hypothese formuleren en verklaren	X	
I2 Antwoorden op de onderzoeksvraag verzamelen		
I2.1 Onderzoeksplan maken	X	
I2.2 Experiment uitvoeren	X	
I2.2a Onderzoekscondities neerzetten in experimentele opstelling	X	
I2.2b Uitvoeren metingen	X	
I2.3 Waarnemen	X	
I3.1 Gegevens vastleggen		
I3.1 Gegevens vastleggen	X	
I3.2 Gegevens ordenen en bewerken	X	X
I3.3 De kern weergeven	X	X
I4 Conclusies trekken en bediscussiëren		
I4.1 Conclusies trekken	X	
I4.1a Conclusies trekken	X	
I4.1b Verklaren resultaten	X	
I4.2 Kritische discussie		
I5 Presenteren		
I5.1 Presentatie voorbereiden		
I5.2 Presenteren		
I5.3 Betekenis van de resultaten		
I5.4 Delen van de bevindingen		
Domeinen		
Domein natuurkunde		
Krachten	X	
Domein Techniek		
Eigenschappen van materialen		X

Hieronder wordt de opzet van de verschillende praktische opdrachten nader toegelicht.

Brug

Bij de opdracht Brug maakt de leerling een brug van papier. De brug moet stevig genoeg zijn om twee speelgoedautootjes te kunnen dragen. Ook moet het speelgoedbootje overal onder de brug door kunnen varen. Deze ontwerpopdracht valt onder het inhoudelijke domein *Constructies*. Daarbij komen de volgende vaardigheden aan de orde: probleem verkennen en verwoorden, oplossing bedenken, oplossing kiezen, materialen en gereedschap gebruiken, maken van het ontwerp, testen van het ontwerp, trouble shooting en herontwerp.

De opdracht Brug is een bewerking van een praktische opdracht in de PPON peiling van Natuur en Techniek van 2010 (Kneepkens, Van der Schoot, en Hemker, 2011). Voor de opdracht is maximaal 30 minuten beschikbaar. Het materiaal en gereedschap voor de opdracht bestaat uit: een blauw vel papier A4 (120 gram) met rivierprint, een bootje (7 cm hoog), een speelgoedautootje van 32 gram en 68 gram, 2 houten blokjes van 8x5x5 cm, een A4-vel groen 80 grams papier, een half vel in de lengte gedeeld wit 120 grams papier, een half vel in de lengte gedeeld wit 160 grams papier, een schaar, een potlood, en een liniaal. De toetsleider beschikt over reservepapier van alle kleuren.

Fietsbel

In de gekozen deelopdracht Fietsbel wordt de leerling gevraagd de losse onderdelen van een fietsbel in elkaar te zetten. De opdracht valt onder het inhoudelijke domein *Overbrengingen*.

Daarbij komen de volgende vaardigheden aan de orde: materialen en gereedschap gebruiken, maken van het ontwerp, testen van het ontwerp, en trouble shooting.

De opdracht *De fietsbel* is een bewerking van een praktische opdracht in de periodieke peiling van Natuur en Techniek van 2010 (Kneepkens, Van der Schoot, en Hemker, 2011). Voor deelopdracht 1 (in elkaar zetten van de fietsbel) is maximaal 10 minuten beschikbaar. Het materiaal en gereedschap voor de deelopdracht bestaat uit: negen onderdelen van een gedemonteerde fietsbel, waarbij ook het veertje eruit is gehaald; een schroevendraaier, en een tang. De toetsleider beschikt over een kleine fietsbel die kan worden ingezet als model.

Bibberspel

In de volledige opdracht Bibberspel maakt de leerling de elektrische aansluiting voor een bibberspel, en doet dit ook in situaties waarin het spel een defect vertoont, een zogeheten 'trouble shooting' taak (Jonassen, 2000). Bij deze laatste taken test hij en levert een testrapport. In deze studie wordt alleen de deelopdracht van het aansluiten van alle onderdelen betrokken. De opdracht valt onder het inhoudelijke domein Elektriciteit. Daarbij komen de volgende vaardigheden aan de orde: materialen en gereedschap gebruiken, maken van het ontwerp, testen van het ontwerp, trouble shooting, herontwerp. Voor de deelopdracht is maximaal 10 minuten beschikbaar. Het materiaal en gereedschap voor de opdracht bestaat uit: een plankje, een fasedraad, een fitting, een lampje, drie krokodillebeksnoertjes, een 4,5 volts batterij, een (speel)staafje gemaakt van fasedraad en een rietje. Verder beschikt de leerling over een tester voor een batterij, een reservelampje en een lege batterij (te herkennen aan een zwarte stip).

In de afgeronde toestand is sprake van een gesloten stroomcircuit. Daarbij zijn de draden op de juiste plaatsen verbonden en is het lampje goed aangedraaid. De stroomkring is onderbroken wanneer de draad is aangesloten op de isolatie van de bibberdraad of wanneer het lampje niet goed is aangedraaid. Bij de start van de opdracht zijn alle snoertjes (klemmetje) los gehaald. Het spel wordt met losse draadjes, een losse volle batterij en een los lampje (geen lampje in de fitting) klaargezet.

Knikkerbaan

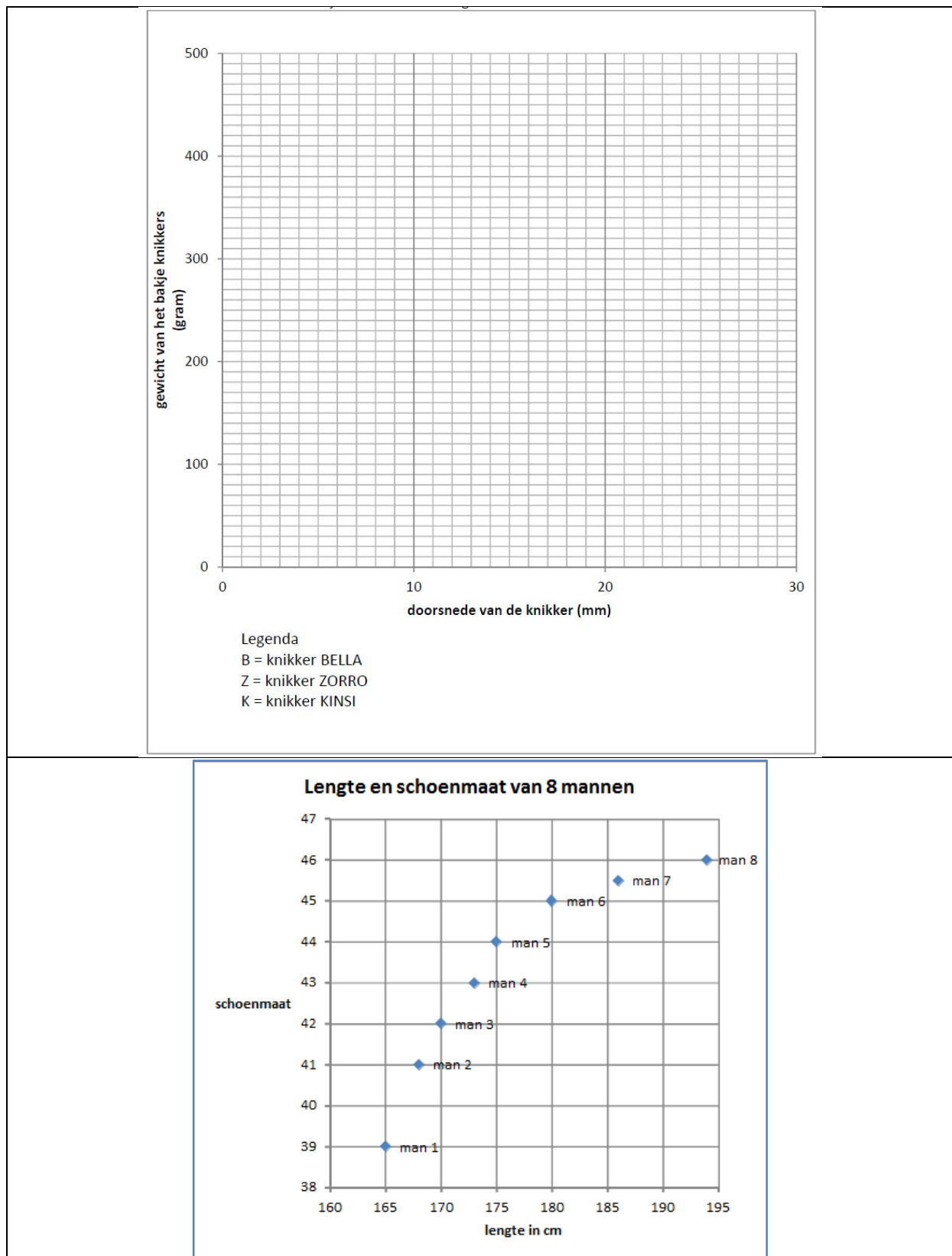
Bij de opdracht Knikkerbaan onderzoekt de leerling het effect van manipulatie van verschillende (onafhankelijke) variabelen op de het gedrag van knikkers op de knikkerbaan. De opdracht valt onder het inhoudelijke domein *Krachten*. Bij de gekozen deelopdrachten, het opzetten en uitvoeren van

twee experimenten, komen de volgende onderzoeksvaardigheden aan de orde: onderzoeksplan maken, experiment uitvoeren, onderzoekscondities neerzetten in experimentele opstelling, uitvoeren metingen, waarnemen, gegevens ordenen en bewerken, de kern weergeven, conclusies trekken en bediscussiëren, conclusies trekken, en verklaren resultaten.

Centraal in de opdracht staat de zogenoemde 'control of variables strategy' (Schwichow, Croker, Zimmerman, Höffler, & Härtig, 2016). Als een onderzoeker iets wil kunnen zeggen over het gevolg van verandering in een bepaalde variabele, dan moeten de overige variabelen die van invloed kunnen zijn op het resultaat, gelijk worden gehouden. De leerlingen voerden twee keer één volledig experiment uit. Voor de deelopdrachten van experiment 1 en 2 is resp. maximaal 20 en 25 minuten beschikbaar. Het materiaal en gereedschap voor de opdracht bestaat uit: twee knikkerbanen, vier knikkers, waarvan twee relatief lichtere en twee relatief zwaardere; een aanduidingskaartje "Baan 1" en kaartje "Baan 2". Bij aanvang van beide experimenten kunnen leerlingen de twee knikkerbanen instellen op de volgende variabelen: helling (steil-flauw), oppervlak (hard-zacht), startpositie (laag-hoog).

Knikkers te koop

In de opdracht Knikkers te koop onderzoekt de leerling het verband tussen de grootte van knikkers en het gewicht van een bakje knikkers. In de deelopdracht die in deze studie is betrokken wordt van de leerling verwacht dat deze de resultaten duidelijk weergeeft in een (lijn)diagram en daarbij een passende conclusie trekt. De opdracht valt onder het inhoudelijke domein *Eigenschappen van materialen*. Bij deelopdracht 3, metingen weergeven in een diagram, komen de vaardigheden 'gegevens ordenen en bewerken' en 'de kern weergeven' aan de orde. Voor de deelopdracht is maximaal 20 minuten beschikbaar. Het materiaal en gereedschap voor de volledige opdracht bestaat uit: een plastic bakje met knikkers van het relatief kleinst formaat, een bakje met knikkers van een relatief middelgroot formaat, een bakje met knikkers van het relatief grootst formaat, twee blokjes met afmetingen 10 x 4 x 4 cm, een liniaal met lengte van 15 cm. Voor deze deelopdracht 3 beschikt de leerling bovendien over een leeg diagram, potlood en pen (zie Figuur 1 boven). De toetsleider beschikt over een voorbeelddiagram waarin het verband tussen lengte en schoenmaat wordt weergegeven (zie Figuur 1 onder).



Figuur 1: In te vullen leeg diagram (boven) en voorbeelddiagram uit andere context (onder) bij de opdracht Knikkers te koop.

2.1.2 Afname, begeleiding en scoring van de praktische opdrachten

De begeleiding en scoring van de praktische opdrachten werd uitgevoerd door 40 getrainde derdejaars studenten van een lerarenopleiding basisonderwijs in het oosten van het land. Tijdens de afname was in een speciaal ingerichte ruimte van de school plaats voor zes praktische opdrachten voor evenzoveel leerlingen. In de ruimte waren twee toetsleiders aanwezig die elk drie leerlingen begeleidden bij een opdracht. Tussentijds scoorden de toetsleiders de opdrachten.

Per opdracht is een protocol voor begeleiding en scoring ontworpen en beproefd dat stoelt op het interventiemodel van dynamisch toetsen. Hierbij werden leerlingen actief gevolgd bij de oplossing van deelopdrachten en indien nodig geholpen. Getrainde toetsleiders gaven daarbij op grond van observaties gevraagd en ongevraagd (als leerlingen vastlopen) hulp. Bij de begeleiding is ervan uitgegaan dat verschillende soorten hulp een verschillende functie hebben bij de aanpak van de opdrachten door de leerling. Er werden vier vormen van hulp onderscheiden, die afhankelijk van de voortgang van de leerling kon worden verstrekt: leeshulp, proceshulp, inhoudelijke hulp en modeling (voordoen, - zeggen).

Leeshulp omvatte het opheffen van belemmeringen die de leerlingen ondervonden bij het begrijpen van de opdracht ten gevolge van leesproblemen of het niet begrijpen van woorden in de tekst of illustraties op het werkblad. Deze vorm van hulp moest het leerlingen mogelijk maken zich een beeld te vormen van de opdracht. Deze hulp beperkte zich tot het verduidelijken wat er op het werkblad staat en bevatte geen informatie die verder gaat dan dat. Anders gezegd, de toegankelijkheid van de opdracht werd hierdoor vergroot.

Proceshulp omvatte met name regulatieve steun met aanwijzingen als: “Lees het nog eens goed”, “Kun je verder?”, “Begrijp je de opdracht?”, “Heb je het goed op kunnen schrijven?”, “Heb je gecontroleerd of alles goed staat?”. Ook het geven van tijdsindicaties viel onder deze regulatieve steun: “Je hebt nog 5 minuten”. *Proceshulp* omvatte bovendien affectieve steun door middel van bemoedigende opmerkingen als “Ga zo door”, “Je doet het prima”, “Dat is duidelijk”. *Proceshulp* was algemeen van aard en niet gebonden aan een specifieke opdracht. Vanuit meetoptiek was *proceshulp* noodzakelijk om voldoende responsen te ontlokken aan leerlingen binnen de beschikbare toetstijd.

Inhoudelijke hulp betrof het voorbereiden of deels uitleggen van de denkstappen die de leerling moest maken om tot een antwoord te komen. *Inhoudelijke hulp* was een specifieke aanwijzing, die alleen bij een bepaalde taak hielp. Bijvoorbeeld een aanwijzing als: “Heb je nu alle mogelijke verschillen tussen de twee knikkerbanen opgeschreven?” was alleen relevant als een leerling verschillen moest vaststellen en stuurde de leerling in de richting van het zoeken van nog meer verschillen. Een voorbeeld geven was ook een vorm van inhoudelijke hulp, die de leerling een idee gaf van de vorm van het antwoord dat werd verwacht. Ook aanwijzingen wat voor soort informatie er op specifieke plaatsen in een grafiek of tabel moest worden opgeschreven viel hieronder. *Inhoudelijke hulp* moest worden gegeven als duidelijk werd dat leerlingen zonder deze hulp niet verder kwamen in het bedenken en beschrijven van een antwoord. *Inhoudelijke hulp* betrof nooit het (volledige) antwoord zelf. De meeste te verwachten interventies in de praktische opdrachten waren uitgeschreven voor de toetsleider. Het was essentieel dat toetsleiders zich hielden aan de voorgeschreven interventies of althans geen verdergaande hulp boden, die zou neerkomen op het weggeven van het antwoord of het voordoen (wat ‘modeling’ zou betekenen).

Modeling is een vorm van hulp waarbij de toetsleider het volledige antwoord gaf of de gevraagde handeling voordeed. Deze vorm van hulp zorgde ervoor dat leerlingen verder konden met een (volgende stap in de) opdracht als ze er niet in slaagden om met inhoudelijke hulp tot een oplossing te komen.

De eerste drie minuten gaf een toetsleider geen inhoudelijke hulp of modeling. In de beginfase mocht een toetsleider alleen proceshulp of leeshulp bieden. Dit om de leerlingen de kans te bieden de opdracht zelfstandig uit te voeren.

Na tweederde van de tijdslimiet vroeg de toetsleider: ‘Heb je hulp nodig?’

- Als de leerling aangaf hulp te willen, gaf de toetsleider *inhoudelijke* hulp.
- Wanneer een leerling aangaf geen hulp te willen, liet de toetsleider hem/haar nog 1 minuut verder werken, en mocht de leerling dan zichtbaar nog niet verder komen met de deelopdracht, dan gaf de toetsleider *inhoudelijke* hulp.
- Als de tijdslimiet voor de deelopdracht was verlopen en de leerling geen oplossing gevonden had of wanneer duidelijk was dat de leerling de opdracht niet kon voltooien werd de oplossing gemodelleerd.

De scoring van de opdrachten vond voor de leerling onopvallend plaats in de kantlijn van het leerlingenblad op een daartoe aangegeven tabelletje (zie Figuur 2). Gescoord werd op het wel (score 1) of niet volbracht hebben (score 0) van de onderscheiden prestatie-elementen, eerst in de conditie zonder hulp (onder de kolom met code 0) en eventueel in de conditie na inhoudelijke hulp (onder de kolom met code 3). Uit deze scoring kan direct het verlenen van inhoudelijke hulp worden afgeleid. Een score 0 in de eerste kolom betekent per definitie dat de toetsleider inhoudelijke hulp geeft over het gescoorde element. In de eerste twee tabelletjes van Figuur 2 zijn twee equivalente wijzen van scoring weergegeven. Daarbij bedragen de scores op de eerste twee prestatie-indicatoren in de conditie zonder hulp 1 en op de volgende twee indicatoren 0 zonder hulp en 1 na hulp. In het derde tabelletje blijven de scores 0 op indicatoren 3 en 4, nadat hulp is gegeven. In dat geval wordt de oplossing gedemonstreerd of verteld. Vervolgens wordt modeling gescoord door 4 aan te kruisen op de onderste regel. Ook zien we in de tabelletjes dat proceshulp is gegeven, ergens tijdens de opdracht; op de onderste regel is 2 aangekruist.

Het geven van lees- en proceshulp, zie respectievelijk codes 1 en 2 in de onderste rij van de tabel, kon gedurende de hele opdracht en werd apart van inhoudelijke hulp genoteerd. Per deelopdracht werd de beschikbare tijd aangegeven (zie linksboven tabelletjes figuur 2).

<table> <tr><td>T5</td><td>0</td><td>3</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>2</td><td>1</td><td>1</td></tr> <tr><td>3</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>2</td><td>4</td></tr> </table>	T5	0	3	1	1	1	2	1	1	3	0	1	4	0	1	1	2	4	<table> <tr><td>T5</td><td>0</td><td>3</td></tr> <tr><td>1</td><td>1</td><td></td></tr> <tr><td>2</td><td>1</td><td></td></tr> <tr><td>3</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>2</td><td>4</td></tr> </table>	T5	0	3	1	1		2	1		3	0	1	4	0	1	1	2	4	<table> <tr><td>T5</td><td>0</td><td>3</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>2</td><td>1</td><td>1</td></tr> <tr><td>3</td><td>0</td><td>0</td></tr> <tr><td>4</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>2</td><td>4</td></tr> </table>	T5	0	3	1	1	1	2	1	1	3	0	0	4	0	0	1	2	4
T5	0	3																																																						
1	1	1																																																						
2	1	1																																																						
3	0	1																																																						
4	0	1																																																						
1	2	4																																																						
T5	0	3																																																						
1	1																																																							
2	1																																																							
3	0	1																																																						
4	0	1																																																						
1	2	4																																																						
T5	0	3																																																						
1	1	1																																																						
2	1	1																																																						
3	0	0																																																						
4	0	0																																																						
1	2	4																																																						

Figuur 2: Tabelletjes waarin prestatie-elementen werden gescoord in de conditie zonder hulp ("0") en met hulp ("3")

2.1.3 Kennistoets

Kennis van Natuur en Techniek is gemeten door middel van een schriftelijke toets die bestond uit 25 of 26 opgaven. De opgaven zijn afkomstig uit een daartoe ontwikkelde opgavenbank van in totaal 180 opgaven, verdeeld over de te meten domeinen van N&T (zie tabel 4). Een deel van de opgaven was gelijk aan de Periodieke Peiling van 2010 om zodoende vergelijkbaarheid van de toetsscores te waarborgen. Er is gebruik gemaakt van in totaal 14 verschillende toetsversies, die via een gelinkt ontwerp overlappen in samenstelling. Hierdoor konden de toetsscores achteraf op één vaardigheidsschaal worden uitgedrukt door middel van IRT-analyses (zie voor meer informatie: Cito, 2016). Alle leerlingen in één klas kregen dezelfde toetsversie.

Tabel 4: Verdeling opgaven over de kerndoelen.

Kerndoel	Anker opgaven	Nieuwe opgaven	Totaal
40 Leefomgeving	14	17	31
41 Vorm en functie	18	13	31
42 Onderzoeken	18	17	35
43 Weer en klimaat	3	19	22
44 Werking, vorm en materiaal	17	20	37
46 Aarde	10	14	24
Eindtotaal	80	100	180

2.1.4 Leerlingvragenlijsten: attitude, buitenschoolse activiteiten en zelfbeoordeling

Met behulp van leerlingvragenlijsten is informatie ingewonnen over de attitude van leerlingen ten aanzien van Natuur en Techniek en de activiteiten, die zij in de eigen vrije tijd ondernemen ten aanzien van N&T. Bij de meting van attitude ten aanzien van N&T is gebruik gemaakt van een bestaand instrument (Denessen e.a., 2011, Van Langen & Vierke; 2009). Hierbij kreeg de leerling twaalf stellingen voorgelegd over vier aspecten die elk een (betrouwbare) subschaal vormen: a) de mate van plezier die de leerling ervaart met N&T, b) het ingeschatte belang van N&T voor zichzelf en de samenleving, c) plannen om zelf iets te gaan doen met N&T in de eigen toekomst, en d) de ervaren moeilijkheid van N&T.

Daarnaast zijn aan de leerlingen 9 vragen voorgelegd over hoe vaak ze sinds het begin van het schooljaar een aantal activiteiten in hun vrije tijd hadden ondernomen die te maken hebben met Natuur en Techniek (Zie tabel 5). De antwoordmogelijkheden waren: 1=nooit, 2= één of twee keer, 3 = vaker dan twee keer. De schaa score Vrijtijdsbesteding aan N&T is tot stand gekomen door het gemiddelde op de itemscores te berekenen. Deze schaa score heeft een acceptabele betrouwbaarheid ($\alpha = .73$).

Tabel 5: Vragen in de schaal voor Vrijtijdsbesteding op het terrein van Natuur en Techniek

Uitgangsvraag: Welke van de onderstaande activiteiten heb je sinds het begin van het schooljaar in je vrije tijd ondernomen?	
1.	Bezoek aan Museon, Nemo, Naturalis, Corpus of een ander museum over Natuur & techniek
2.	Apparaten of andere dingen in elkaar zetten / repareren
3.	Nieuwe uitvindingen doen die te maken hebben met Natuur & techniek
4.	Spelen met Natuur & techniek speelgoed (bijvoorbeeld bouwplaat, bouwdoos, (technisch) Lego, robot, scheikundendoos/proefjes doen, modelauto, modelvliegtuig, microscoop)
5.	Zelf programmeren op de computer
6.	Knutselen met materialen (bijv. Hout, metaal, karton)
7.	De natuur in of natuurfilms kijken om meer te leren over Natuur & techniek
8.	Boeken of tijdschriften lezen over Natuur & techniek
9.	Plantjes zaaien of stekken

N.B. Antwoordcategorieën: 1=nooit, 2= één of twee keer, 3 = vaker dan twee keer; Betrouwbaarheid van de samengestelde schaal (gemiddelde op itemscores) via Cronbach's $\alpha = .73$

Tot slot zijn per praktijkopdracht vragenlijsten ontworpen waarbij leerlingen werd gevraagd zichzelf te beoordelen op de onderliggende ontwerp -en onderzoeksvaardigheden behorend bij de opdracht die ze hadden uitgevoerd. De voorgelegde uitspraken zijn weergegeven in tabel 6a (onderzoekopdrachten) en 6b (ontwerp opdrachten). De leerlingen konden antwoorden op een vier-puntsschaal: 1= Daar gaat nog veel mis mee; 2= Dat lukt, maar met enkele fouten; 3= Dat lukt goed, heel soms een fout; 4= Dat lukt perfect. Alle schalen hebben een redelijke (Knikkers te Koop: $\alpha = .68$) tot goede betrouwbaarheid (α 's overige opdrachten tussen .80 en .84).

Tabel 6a: Uitspraken voor zelfbeoordeling van onderzoeksvaardigheden per praktische opdracht

Uitspraak	Onderzoeksvaardigheid	Knikker- baan	Knikkers te koop
Uitgangstelling: Geef nu aan hoe goed je bent in de volgende onderdelen van ontwerpen:			
Goede redenen geven waarom iets zal gebeuren (voorspelling doen)	I1.5 Verwachtingen formuleren: voorspellen en verklaren	X	
Goed klaarzetten van de twee knikkerbanen bij de experimenten	I2.2a Onderzoekscondities neerzetten in experimentele opstelling	X	
Het verzamelen van voldoende aantal metingen	I2.2b Uitvoeren metingen	X	X
Nauwkeurig meten	I2.2b Uitvoeren metingen	X	X
Het duidelijk opschrijven van de metingen in een tabel	I3.2 Gegevens ordenen en bewerken	X	
Het duidelijk opschrijven van de metingen in grafiek	I3.2 Gegevens ordenen en bewerken		X
Het duidelijk opschrijven van de metingen in de tabel	I3.2 Gegevens ordenen en bewerken	X	
Het kunnen rekenen met de getallen uit de grafiek	I3.2 Gegevens ordenen en bewerken		X
Het trekken van een conclusie uit de resultaten van de metingen in een experiment	I4.1a Conclusies trekken	X	
Het kunnen uitleggen wat de oorzaak is van de verschillen die zijn vastgesteld	I4.1b Verklaren resultaten	X	
Betrouwbaarheid van de samengestelde schaal (gemiddelde op itemscores; Cronbach's alpha)		.80	.68

N.B. Antwoordcategorieën: 1= Daar gaat nog veel mis mee; 2= Dat lukt, maar met enkele fouten; 3= Dat lukt goed, heel soms een fout; 4= Dat lukt perfect

Tabel 6b: Uitspraken voor zelfbeoordeling van ontwerpvaardigheden per praktische opdracht

Uitspraak	Ontwerpvaardigheid	Brug	Fietsbel	Bibberspel
Uitgangstelling: Geef nu aan hoe goed je bent in de volgende onderdelen van ontwerpen:				
Het begrijpen aan welke eisen [de brug] moet voldoen.	D1. Een probleem herkennen en verkennen: probleem verkennen	X		
Het bedenken van mogelijke oplossingen voor het maken van [de brug].	D2 Ontwerpen van de oplossing: D2.1 oplossing kiezen	X		
Het kiezen van het juiste gereedschap voor het maken van [de brug, de fietsbel, het bibberspel].	D3 Het ontwerp uitvoeren: D3.1 materialen en gereedschap gebruiken	X	X	X
Het kiezen van het juiste materiaal voor het maken van [het bibberspel].				X
Het kunnen maken van [de brug, de fietsbel, het Bibberspel].		X	X	X
Het oplossen van problemen bij het maken van [de brug, de fietsbel, het Bibberspel].	D3. Het ontwerp uitvoeren: D3.2 maken van het ontwerp	X	X	X
Het begrijpen van de problemen die ontstonden bij het maken [de brug, de fietsbel, het Bibberspel].	D4. Het ontwerp uittesten en verbeteren: D4.2 trouble shooting	X	X	X
Testen of [de brug, de fietsbel, het Bibberspel] voldoet aan de gestelde eisen.	D4. Het ontwerp uittesten en verbeteren: D4.1 testen van het ontwerp	X	X	X
Betrouwbaarheid van de samengestelde schaal (gemiddelde op itemscores; Cronbach's alpha)		.84	.82	.81

N.B. Antwoordcategorieën: 1= Daar gaat nog veel mis mee; 2= Dat lukt, maar met enkele fouten; 3= Dat lukt goed, heel soms een fout; 4= Dat lukt perfect

2.1.5 Subjecten

Leerlingen

De dataverzameling heeft plaatsgevonden in november 2015. Ten behoeve van de voorliggende deelstudie zijn de data gebruikt van 127 basisscholen die representatief zijn naar locatiekenmerken (regio, verstedelijking), denominatie, schoolgrootte, en percentage leerlinggewichten (zie ITS, Expertisecentrum Nederlands, 2016). In totaal zijn voor het doel van deze deelstudie de data gebruikt van 1485 leerlingen die deelnamen aan een van de geselecteerde vijf praktische opdrachten. Van de leerlingen was 51,5% meisje en 48.5% jongen. Onder de steekproef bevonden zich 0.2% 9-jarigen, 6.2% 10-jarigen, 71.2% 11-jarigen, 21.6% 12-jarigen en 0.8% 13-jarigen. De leerlingen zaten in groep 8 van de basisschool.

Proefleiders

De begeleiding en scoring van de praktische opdrachten is uitgevoerd door in totaal 40 derdejaarsstudenten aan de lerarenopleiding basisonderwijs. Deze studenten hadden inmiddels al eerste leservaring opgedaan in het kader van een (LIO)stage. Voor het doel van de peilingsstudie, waarvan deze studie deel uitmaakt, is een training verzorgd waarin de essentiële vaardigheden van begeleiding en scoring centraal stonden.

3 Resultaten

In deze paragraaf beschrijven we de resultaten van de studie. Omdat anders het aantal te rapporteren deelopdrachten te groot wordt voor één tabel of figuur, worden de bevindingen van de opdracht Knikkerbaan waar nodig gerapporteerd in afzonderlijke tabellen en figuren.

3.1 De verleende hulp aan leerlingen

Tabel 7a geeft een beeld van de combinaties van hulp die leerlingen hebben ontvangen bij het uitvoeren van vier verschillende praktische opdrachten. Tabel 7b geeft het beeld weer voor alle onderzochte deelopdrachten van de opdracht Knikkerbaan. Verder zijn twee maten voor beoordelaarsovereenstemming weergegeven voor de vier basale soorten van hulp: kappa en het proportie overeenkomstige coderingen. In sommige gevallen kon geen kappa berekend worden omdat niet alle categorieën voorkwamen (bijvoorbeeld bij de opdracht Brug kregen alle leerlingen die door de twee beoordelaars gemeenschappelijk waren geobserveerd inhoudelijke hulp). In dat geval kan alleen het proportie overeenstemming worden berekend.

Kijkend naar de beoordelaarsovereenstemming op de eerste vier opdrachten (Brug, Fietsbel, Bibberspel en Knikkers te koop) dan komt het volgende beeld naar voren:

- De beoordelaarsovereenstemming over inhoudelijke hulp en modeling is redelijk tot goed bij de opdrachten Brug Fietsbel, Bibberspel en Knikkers te koop, afgaand op de kappa-waarden.
- De beoordelaarsovereenstemming over proceshulp is alleen goed bij Knikkers te koop en matig bij Brug, Fietsbel, en Bibberspel. Vanwege de kleine aantallen leerlingen kan ook gekeken worden naar de proportie overeenstemmende observaties. Dan blijkt de overeenstemming goed.
- Kijkend naar leeshulp is op de opdrachten Fietsbel en Knikkers te koop sprake van een hoge beoordelaarsovereenstemming en op twee opdrachten van een matige beoordelaarsovereenstemming, afgaand op de kappa-waarden. Als echter ook hier wordt uitgeweken naar de proportie overeenstemmende observaties, dan is de overeenstemming goed.

Kijkend naar de beoordelaarsovereenstemming op de deel opdrachten voor het experiment Knikkerbaan dan komt het volgende beeld naar voren:

- De beoordelaarsovereenstemming is bij alle vormen van geobserveerde hulp goed, met uitzondering van proceshulp bij de deelopdracht Metingen in experiment 2.

Kijkend naar de resultaten wat betreft de hoeveelheid verleende hulp, komt het volgende beeld naar voren voor de eerste vier opdrachten (zie Tabel 7a). De aard en frequentie van verstrekte hulp loopt sterk uiteen per (deel)opdracht. Veruit het vaakst werd hulp verleend bij de constructie van de papieren brug (96.2%). Bij het in elkaar zetten van een Fietsbel, het aanleggen van een gesloten elektrisch circuit voor het Bibberspel, en het invullen van het diagram bij Knikkers te koop ontvingen in totaal respectievelijk 67.3, 85.9, en 80.8 % van de leerlingen een vorm van hulp.

De leerlingen ontvingen bij de verschillende deelopdrachten van de Knikkerbaan beduidend minder vaak hulp dan bij de eerste vier besproken opdrachten (zie Tabel 7b). Bovendien zijn er grote verschillen in frequentie van ontvangen hulp tussen de deelopdrachten van de Knikkerbaan onderling. Bij het klaarzetten van de experimentele opstelling van de Knikkerbaan (experiment 2) ontving 57.4% van de leerlingen hulp. Bij het verrichten van metingen in twee experimenten kreeg resp. 55.4 en 47.7 % van de leerlingen hulp. Bij het trekken van conclusies op grond van de metingen werd in experiment 1 en 2 aan 18.7 en 12.4% van de leerlingen hulp gegeven.

Er zijn aanzienlijke verschillen in hoe vaak een oplossing wordt voorgedaan (gemodelleerd). Bij Brug (42.4%), Bibberspel (33.1%) en Knikkers te koop (25.3%) wordt de oplossing vaker gemodelleerd dan bij Fietsbel (13.4%) en de deelopdrachten van Knikkerbaan. Bij Knikkerbaan lopen de percentages modellering uiteen van 3.2% bij conclusies trekken in experiment 2 tot 14.3% bij de deelopdracht Klaarzetten van de experimentele opstelling.

Tabel 7a: Ontvangen hulpvormen bij vier praktische opdrachten: beoordelaars-overeenstemming en resultaten in percentages leerlingen

	Brug (n=315)	Fietsbel (n=342)	Bibberspel (n=311)	Knikkers te koop (n=281)
Beoordelaarsovereenstemming in registratie van hulpsoort:				
kappa en proportie overeenstemming				
1. Leeshulp (ja-nee)	.32 (.77)	1.00 (1.00)	.37 (.70)	1.0 (1.00)
2. Proceshulp (ja-nee)	.33 (.81)	.36 (.76)	.38 (.70)	.78 (.96)
3. Inhoudelijke hulp (ja-nee)	NB (.89)	1.00 (1.00)	.42 (.76)	.76 (.92)
4. Modeling (ja-nee)	.62 (.89)	.80 (0.96)	.42 (.83)	.60 (.88)
Verkregen mate van hulp				
Geen hulp gehad	3.8	32.7	14.1	19.2
Wel hulp gehad	96.2	67.3	85.9	80.8
Na hulp oplossing gemodelleerd	42.4	13.4	33.1	25.3

N.B. Tussen haakje is de proportie overeenstemming weergegeven. NB= Kappa niet te berekenen; Interpretatie van waarden: .0 - .2: gering, .2 - .4: matig, .4 - .6: redelijk, >.6: voldoende tot perfect

Tabel 7b: Ontvangen hulpvormen bij de opdracht Knikkerbaan: beoordelaarsovereenstemming en resultaten in percentages leerlingen

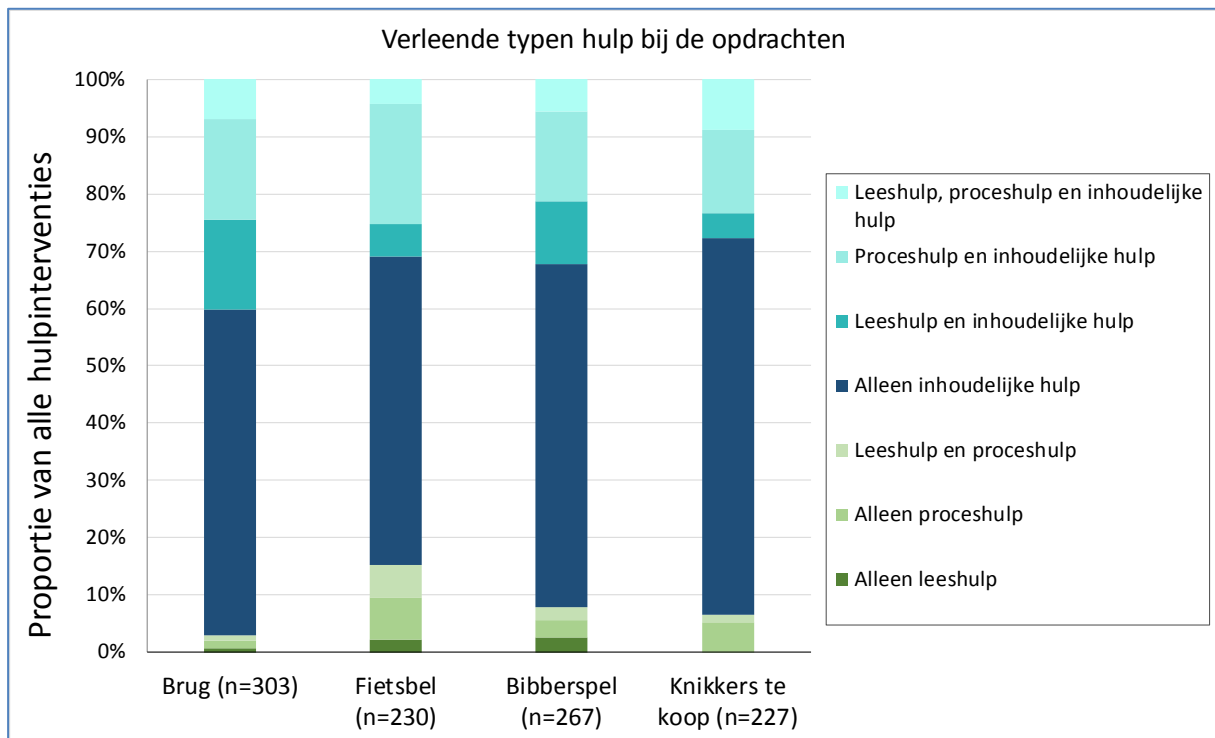
	Knikkerbaan experiment 1		Knikkerbaan experiment 2		
	Metingen verrichten (n=222)	Conclusies trekken (n=219)	Klaarzetten experiment (n=223)	Metingen verrichten (n=220)	Conclusies trekken (n=218)
Beoordelaarsovereenstemming in registratie van hulpsoort:					
kappa en proportie overeenstemming					
1. Leeshulp (ja-nee)	NB (1.00)	NB (1.00)	NB (.88)	.64 (.94)	NB (.94)
2. Proceshulp (ja-nee)	.67 (.87)	NB (.94)	NB (.94)	.45 (.88)	NB (1.00)
3. Inhoudelijke hulp (ja-nee)	.89 (.94)	1.00 (1.00)	.81 (.88)	.81 (.88)	1.00 (1.00)
4. Modeling (ja-nee)	NB (.73)	1.00 (1.00)	.82 .94)	.77 (.94)	1.00 (1.00)
Verkregen mate van hulp					
Geen hulp gehad	44.6	81.3	42.6	52.3	87.6
Wel hulp gehad	55.4	18.7	57.4	47.7	12.4
Na hulp oplossing gemodelleerd	8.1	5.4	14.3	8.2	3.2

N.B. Tussen haakje is de proportie overeenstemming weergegeven. NB= Kappa niet te berekenen; Interpretatie van waarden: .0 - .2: gering, .2 - .4: matig, .4 - .6: redelijk, >.6: voldoende tot perfect

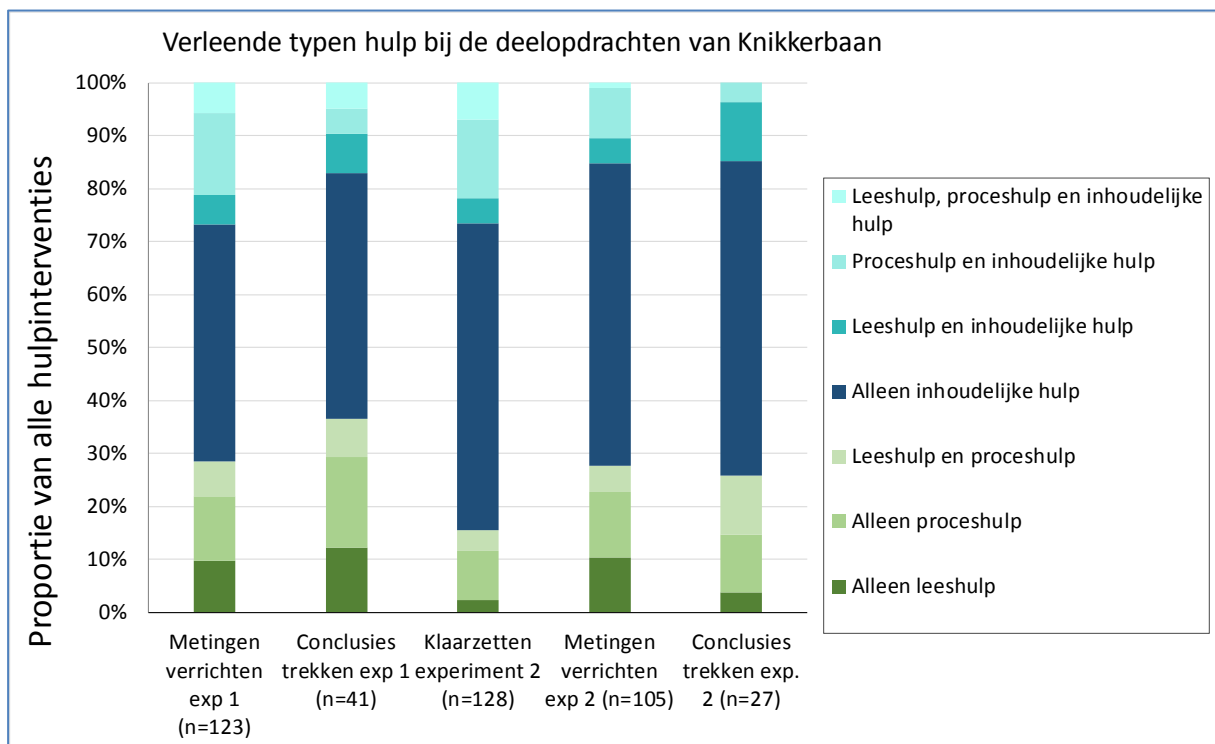
De Figuren 3a en 3b bieden inzicht in de verhoudingen tussen de soorten verleende hulp. De exacte percentages kunnen worden ontleend aan de tabellen in bijlage 1. De aantallen leerlingen die hulp kregen variëren zoals aangegeven in de figuren. Hiernaar kijkend zien we het volgende beeld.

- Pure inhoudelijke hulp (zie donkerblauwe balk) maakt bij de meeste opdrachten rond 50% van alle hulpinterventies uit. Bij Bibberspel (onder meer uitleg over principe van een gesloten stroomkring) en Knikkers te koop bedraagt dit aandeel ongeveer 60%.
- Bij de opdracht Brug zien we dat inhoudelijke hulp het vaakst gepaard gaat met leeshulp en proceshulp (40.3%), terwijl een zeer klein percentage van de hulp uitsluitend lees- en/of proceshulp betreft (3.0%).
- Bij de verschillende deelopdrachten van Knikkerbaan is veel vaker sprake van uitsluitend lees- en proceshulp (variërend tussen 15.6% en 36.6%). Ditzelfde beeld geldt in iets mindere mate ook voor de opdracht Fietsbel (15.2%).

- Bij Knikkerbaan wordt bij de deelopdracht Verrichten van metingen (beide experimenten) en de deelopdracht Conclusies trekken (experiment 1) wat vaker uitsluitend leeshulp verstrekt (zo'n 10 %) dan bij de andere praktische opdrachten (variërend tussen 0 en 2.6%).



Figuur 3a: Verhoudingen tussen verschillende gegeven hulpsoorten binnen vier praktische opdrachten (alleen de leerlingen die hulp hebben ontvangen)



Figuur 3b: Verhoudingen tussen verschillende gegeven hulpsoorten bij de deelopdrachten van de praktische opdracht Knikkerbaan (alleen de leerlingen die hulp hebben ontvangen)

In tabel 8a en 8b zijn verschillen in frequentie van verkregen hulp naar achtergrondkenmerken weergegeven.

Uit de resultaten van Tabel 8a blijkt dat er geen verschillen optreden in de frequentie van hulpcombinaties naar de achtergrondkenmerken leeftijd en leerlinggewicht.

Bij de opdracht Bibberspel doen meisjes de opdracht vaker zonder hulp dan de jongens (39% versus 26%; Chi-kwadraat=21.0; $df=5$; $p= .001$). De opdracht fietsbel wordt door jongens iets vaker zonder hulp uitgevoerd dan door meisjes. Dit verschil is echter net niet significant (39% versus 26%; Chi-kwadraat=10.0; $df=5$; $p= .07$).

Tabel 8a: Verschillen in verkregen hulp bij vier praktische opdrachten naar achtergrondkenmerken

	Geslacht (jongen- meisje)	Leeftijd (10-11-12-13)	Leerlinggewicht (0; 0,3; 1,2)
Brug			
n	Jongens (n= 163), Meisjes (n= 150)	10 jaar (n=22) 11 jaar (n=210) 12jaar (n=62) 13 jaar (n=6)	Gewicht 0 (n=266) Gewicht 0.3 (n=28) Gewicht 1.2 (n=6)
Chi-kwadraat significant?	nee	nee	nee
Fietsbel			
N	Jongens (n= 185), Meisjes (n= 157)	10 jaar (n=18) 11 jaar (n=225) 12jaar (n=81) 12jaar (n=2)	Gewicht 0 (n=297) Gewicht 0.3 (n=22) Gewicht 1.2 (n=13)
Chi-kwadraat significant?	nee	nee	nee
Aard verschil	39% van de jongens doen de opdracht zonder hulp tegen 26% van de meisjes		
Bibberspel			
N	Jongens (n= 133), Meisjes (n= 177)	10 jaar (n=18) 11 jaar (n=221) 12jaar (n=59) 12jaar (n=1)	Gewicht 0 (n=269) Gewicht 0.3 (n=21) Gewicht 1.2 (n=8)
Chi-kwadraat significant?	ja	nee	nee
Aard verschil	20% van de meisjes doen de opdracht zonder hulp tegen 6% van de jongens		
Knikkers te koop			
N	Jongens (n= 134), Meisjes (n= 147)	10 jaar (n=17) 11 jaar (n=195) 12jaar (n=57)	Gewicht 0 (n=194) Gewicht 0.3 (n=14) Gewicht 1.2 (n=8)
Chi-kwadraat significant?	nee	nee	nee

Tabel 8b: Verschillen in verkregen hulp bij de opdracht Knikkerbaan naar achtergrondkenmerken

	Geslacht	Leeftijd	Leerlinggewicht
	Jongens (n= 107)	10 jaar (n=12)	Gewicht 0 (n=194)
	Meisjes (n= 117)	11 jaar (n=152)	Gewicht 0.3 (n=14)
		12jaar (n=46)	Gewicht 1.2 (n=8)
Knikkerbaan experiment 1 Significantie Chi-kwadraat en aard verschil			
Uitvoeren metingen	nee	Nee	nee
Conclusies trekken	nee	nee	Ja. 0.3 en 1.2 leerlingen krijgen vaker hulp dan 0.0-leerlingen: resp. 28.6, 42.9 en 16.8%)
Knikkerbaan experiment 2: Significantie Chi-kwadraat en aard verschil			
Experiment klaarzetten	nee	nee	nee
Uitvoeren metingen	nee	nee	nee
Conclusies trekken	nee	nee	Ja 0.3 en 1.2 leerlingen krijgen vaker hulp dan 0.0-leerlingen: resp. 23.1, 47.1 en 10.5%

Uit de resultaten van tabel 8b blijkt dat er bij de deelopdrachten van Knikkerbaan geen verschillen optreden in de frequentie van hulpcombinaties aan jongens en meisjes aan leerlingen met verschillende leeftijden. Wel blijkt bij de deelopdracht conclusies trekken dat aan leerlingen met een leerlinggewicht van 0.3 en 1.2 wat vaker hulp wordt gegeven dan leerlingen met leerlinggewicht 0. Bij experiment 1 bedragen de percentages hulp voor de gewichtsgroepen 0, 0.3 en 1.2 resp. 16.8, 28.6, 42.9 (Chi-kwadraat=18.8; $df=10$; $p=0.04$). Voor experiment 2 bedragen deze percentages resp. 10.5, 23.1, en 47.1% (Chi-kwadraat=49.7; $df=8$; $p=.001$).

Tabel 9a: Product-moment correlaties tussen kennisniveau (vaardigheidsscore op de kennistoets Natuur en Techniek) en het al dan niet verkrijgen van hulp

	Brug	Fietsbel	Bibberspel	Knikkers te koop
1. Leeshulp (ja/nee ⁺)				
2. Proceshulp (ja/nee)		-.15*		
3. Inhoudelijke hulp (ja/nee)		-.21**	-.19**	-.17*
3b. Aantal ontvangen inhoudelijke aanwijzingen		-.22**		-.16*
4. Modeling (ja/nee)	-.13*		-.13*	-.22**

N.B. * $p<.05$; ** $p<.01$. N.B. ⁺: Ja/nee gescoord als 0 en 1.

Tabel 9b: Product-moment correlaties tussen de vaardigheidsscore op de Kennistoets Natuur en Techniek en het al dan niet verkrijgen van hulp bij de opdracht Knikkerbaan (twee experimenten)

	Experiment 1		Experiment 2	
	Metingen verrichten	Conclusies trekken	Experiment Klaarzetten	Metingen verrichten
1. Leeshulp (ja/nee ⁺)				
2. Proceshulp (ja/nee)		-.14*		
3. Inhoudelijke hulp (ja/nee)	-.13*	-.16*	-.20**	
3b. Aantal ontvangen inhoudelijke aanwijzingen	-.16*	NB	-.24**	NB
4. Modeling (ja/nee)	-.14*		-.17*	-.16*

N.B. * $p<.05$; ** $p<.01$. N.B. ⁺: Ja/nee gescoord als 0 en 1. NB = niet beschikbaar, omdat maximaal één aanwijzing kan worden gegeven.

In de Tabellen 9a en 9b is het verband tussen het kennisniveau op het terrein van Natuur en Techniek enerzijds en het al of niet ontvangen van vormen van hulp anderzijds weergegeven. Alleen significante correlaties zijn in de tabel afgedrukt.

Allereerst blijkt dat bij geen van de praktische opdrachten een statistisch significant verband is gevonden tussen het kennisniveau en het al of niet ontvangen van leeshulp.

Voorts is bij alle (deel)opdrachten, behalve bij de opdracht Brug en de deelopdracht Metingen verrichten bij experiment 2 van Knikkerbaan, sprake van een negatief verband tussen het kennisniveau en het al of niet ontvangen van inhoudelijke hulp: hoe groter de kennis van Natuur en Techniek, des te minder hulp wordt ontvangen. De waarden liggen tussen $-.13$ (deelopdracht Metingen verrichten, experiment 1 Knikkerbaan) en $-.22$ (Fietsbel). De correlaties zijn weliswaar laag, maar significant. Een vergelijkbaar negatief verband is gevonden tussen het kennisniveau en het aantal ontvangen inhoudelijke aanwijzingen voor de deelopdrachten Fietsbel, Knikkers te koop, en Knikkerbaan (Metingen verrichten en Experiment klaarzetten). Naarmate het kennisniveau van de leerlingen lager is, blijkt de frequentie van ontvangen inhoudelijke aanwijzingen hoger te zijn. Het grootst is hierbij het verband tussen kennis en het aantal aanwijzingen dat leerlingen krijgen bij het klaarzetten van de experimentopstelling ($r=-.24$).

Verder geldt bij twee opdrachten een negatief verband tussen kennisniveau en het vóórkomen van proceshulp. Bij de opdracht Fietsbel ($r=-.15$) en de deelopdracht Conclusies trekken bij experiment 1 Knikkerbaan ($r=-.14$) geldt dat een lager kennisniveau vaker gepaard gaat met proceshulp.

Tot slot zien we dat bij de meeste opdrachten een lager kennisniveau vaker gepaard gaat met modellering van de oplossing door de proefleider. De correlaties liggen tussen $-.13$ (Brug) en $-.22$ (Knikkers te koop). Het verband geldt niet voor Fietsbel en Conclusies trekken bij experiment 1 van Knikkerbaan.

Tabel 10a: Product-moment correlaties tussen aspecten van attitude ten aanzien van Natuur en Techniek en de mate van ontvangen hulp

Hulpsoort	Aspect van attitude NT	Brug	Fietsbel	Bibberspel	Knikkers te koop
Leeshulp	NT vrije tijd				
	Belang NT				
	NT plezier				
	NT moeilijk		.19**		
	Toekomst NT		-.14*		
Proceshulp	NT vrije tijd				
	Belang NT				
	NT plezier				
	NT moeilijk				
	Toekomst NT				
Inhoudelijke hulp	NT vrije tijd				
	Belang NT				
	NT plezier				
	NT moeilijk				
	Toekomst NT			-.14*	-.12*
Voordoen (Modeling)	NT vrije tijd				
	Belang NT		-.12*		
	NT plezier				
	NT moeilijk				
	Toekomst NT				

N.B. Significante positieve correlaties zijn groen gearceerd en negatieve rood. Niet significante correlaties zijn weggelaten; * $p < .05$, ** $p < .01$

Tabel 10b: Product-moment correlaties tussen aspecten van attitude ten aanzien van Natuur en Techniek en de mate van ontvangen hulp

Hulpsoort	Aspect van attitude NT	Experiment 1			Experiment 2	
		Metingen verrichten	Conclusies trekken	Klaarzetten experiment	Metingen verrichten	Conclusies trekken
Leeshulp	NT vrije tijd					
	Belang NT			-.12*		
	NT plezier					
	NT moeilijk				.12*	
	Toekomst NT					
Proceshulp	NT vrije tijd					
	Belang NT			-.12*		
	NT plezier	.16*				
	NT moeilijk		.14*			.14*
	Toekomst NT			-.13*		
Inhoudelijke hulp	NT vrije tijd	.16*				
	Belang NT			.18*		
	NT plezier					
	NT moeilijk					
	Toekomst NT					
Voordoen (Modeling)	NT vrije tijd					
	Belang NT	-.13*		.13*		
	NT plezier					
	NT moeilijk				.15*	
	Toekomst NT			-.17*		

N.B. Significante positieve correlaties zijn groen gearceerd en negatieve rood. Niet significante correlaties zijn weggelaten; * $p < .05$

In de Tabellen 10a en 10b zijn de correlaties tussen attitude-aspecten ten aanzien van Natuur en Techniek enerzijds en de frequentie van ontvangen hulpsoorten anderzijds weergegeven. Positieve correlaties zijn groen gearceerd en negatieve rood. De resultaten laten een grillig beeld zien. De volgende samenhangen worden gevonden.

- Er is geen enkel verband tussen attitude-aspecten en hulp bij de opdracht Brug.
- Meer doen aan *NT in de vrije tijd* hangt samen met
 - méér inhoudelijke hulp bij de deelopdracht metingen verrichten bij Knikkerbaan ($r = .16$).
- Inzien van het *belang van NT* voor het leven van mensen hangt samen met
 - een lagere frequentie waarin de opdracht moet worden voorgedaan bij Fietsbel ($r = -.12$.)
 - Minder leeshulp ($r = -.12$), minder proceshulp ($r = -.12$), meer inhoudelijke hulp ($r = .18$) en meer modeling ($r = .13$) bij het klaarzetten van het tweede experiment Knikkerbaan
 - Minder voordoen bij het trekken van conclusies bij het eerste experiment Knikkerbaan ($r = -.13$);
- Naarmate leerlingen *Plezier in NT* hebben.
 - is de frequentie van proceshulp bij de deelopdracht van Knikkerbaan hoger ($r = .16$).
- *Moeilijk vinden van NT* opdrachten gaat samen met
 - meer leeshulp bij de opdracht Fietsbel ($r = .19$).
 - meer leeshulp bij de deelopdracht Verrichten van metingen bij het tweede experiment Knikkerbaan ($r = .12$)

- meer proceshulp bij de deelopdracht Conclusies trekken bij zowel experiment 1 ($r=.14$) als 2 van Knikkerbaan ($r=.14$).
- meer voordoen bij de deelopdracht Verrichten van metingen bij het tweede experiment Knikkerbaan ($r=.15$)
- Naarmate leerlingen in de *toekomst meer met NT* (beroepsmatig) bezig zeggen te willen zijn,
 - is de frequentie van ontvangen leeshulp bij de Fietsbel ($r=-.14$) lager
 - is de frequentie van ontvangen inhoudelijke hulp lager bij de opdrachten Bibberspel ($r=-.14$) en Knikkers te koop ($r=-.12$)
 - is frequentie van proceshulp bij de deelopdracht Klaarzetten experiment Knikkerbaan lager ($r=-.13$) en de mate van voordoen bij diezelfde deelopdracht lager ($r=-.17$)

3.2 Leerlingenmerken en prestaties in de conditie zonder hulp

De mate van groei die leerlingen kunnen doormaken als gevolg van hulp hangt deels af van de beginscore op de opdrachten, die wordt toegekend voordat eventuele hulp is gegeven. Daarbij beginnen leerlingen bij iedere deelopdracht steeds onder de voorwaarde dat de oplossing van de voorafgaande deelopdracht zelf is gevonden of is gedemonstreerd door de toetsleider.

Te verwachten valt dat de beginscore samenhangt met leerlingkenmerken. Daarom is nagegaan hoe de beginscore op de verschillende praktijkopdrachten correleert met essentiële leerlingkenmerken. In Tabel 11 is het resultaat van deze analyse weergegeven. Opnieuw zijn alleen significante correlaties in de tabel weergegeven. Het volgende beeld komt naar voren.

Wat betreft de *achtergrondkenmerken* blijkt het volgende:

- Jongens scoren zonder hulp wat hoger dan meisjes op de opdrachten Brug ($r=.11$) en Fietsbel ($r=.15$).
- Andersom scoren meisjes zonder hulp hoger op de opdracht Bibberspel ($r=-.19$). Dit laatste verschil is nog wat groter dan het sekseverschil bij de opdracht Brug en Fietsbel.
- Leeftijd hangt op geen van de opdrachten samen met de prestaties zonder hulp.
- De etnische achtergrond van leerlingen, uitgedrukt met behulp van het leerlinggewicht, maakt niet uit voor de score die leerlingen behalen op de opdrachten in de conditie zonder hulp.

Wat betreft Attitude ten aanzien van Natuur en Techniek (N&T) blijkt het volgende:

- een hogere ingeschatte moeilijkheid van Natuur en Techniek samen met een lagere score in de conditie zonder hulp op de opdracht Fietsbel ($r=-.19$)
- Naarmate leerlingen meer toekomstplannen hebben ten aanzien van Natuur en Techniek scoren ze hoger op de opdrachten Fietsbel ($r=.14$), Bibberspel ($r=.12$), Knikkerbaan ($r=.14$), vóórdat hulp wordt gegeven

Wat betreft kennis van Natuur en Techniek (N&T) blijkt het volgende:

- Naarmate het kennisniveau van leerlingen hoger is, liggen de scores zonder hulp op de opdrachten Fietsbel ($r=.20$), Knikkerbaan ($r=.22$), Knikkers te koop ($r=.18$) hoger. Er is geen significant verband gevonden tussen het kennisniveau en scores op de Brug en het Bibberspel.

Tabel 11: Samenhang tussen leerlingkenmerken en de prestaties op de praktijkopdrachten in de conditie vóór hulp

	Brug (n=315)	Fietsbel (n=342)	Bibberspel (n=311)	Knikkerbaan ⁺ (n=220)	Knikkers te koop (n=281)
Achtergrondkenmerken					
Geslacht (0=meisje; 1=jongen)	.11*	.15*	-.19**		
Leeftijd in jaren en maanden					
Leerlinggewicht (0; 0.3 of 1.2)					
Attitude ten aanzien van Natuur en Techniek (N&T)					
Belang N&T (schaalscore)					
N&T plezier (schaalscore)					
N&T moeilijk (schaalscore)		-.19**			
Toekomst N&T (schaalscore)		.14*	.12*	.14*	
Frequentie en aantal bezigheden N&T in vrije tijd (schaalscore)					.15*
Vaardigheidsscore kennistest N&T		.20**		.22**	.18**
Zelfbeeld onderzoeks- en ontwerpvaardigheden (na afloop opdracht)	.46**	.40**	.19**	.22**	.21**

N.B. . Significante positieve correlaties zijn groen gearceerd en negatieve rood. + Hierbij is de somscore op alle deelopdrachten van experiment 1 gehanteerd; * $p < .05$; ** $p < .01$

Het zelfbeeld van leerlingen over de eigen ontwerp- en onderzoeksvaardigheden hangt positief samen met de scores op alle praktische opdrachten. Opvallend hierbij is dat de correlaties tussen prestatie en zelfbeeld voor de meer technische doe-opdrachten Brug ($r = .46$) en Fietsbel ($r = .40$) hoger zijn dan bij de opdrachten waar gemeten en geredeneerd moet worden (Bibberspel: $r = .19$; Knikkerbaan: $r = .22$; Knikkers te koop: $r = .21$).

3.3 Effecten van hulp

3.3.1 Overall effecten van hulp

In de Figuren 4a en 4 b zijn de effecten van hulp op de opdrachtscores weergegeven. Gekeken is naar de leerlingen die daarbij hulp hebben ontvangen. Door de scores voorafgaand aan hulp te vergelijken met die na afloop van hulp, kunnen effecten worden vastgesteld. Per opdracht kon een verschillend aantal scorepunten worden behaald, afhankelijk van de gekozen prestatie-indicatoren. De totaalscores voor de opdrachten zijn omgezet naar proportionele scores tussen 0 en 1 om vergelijkbaarheid te realiseren.

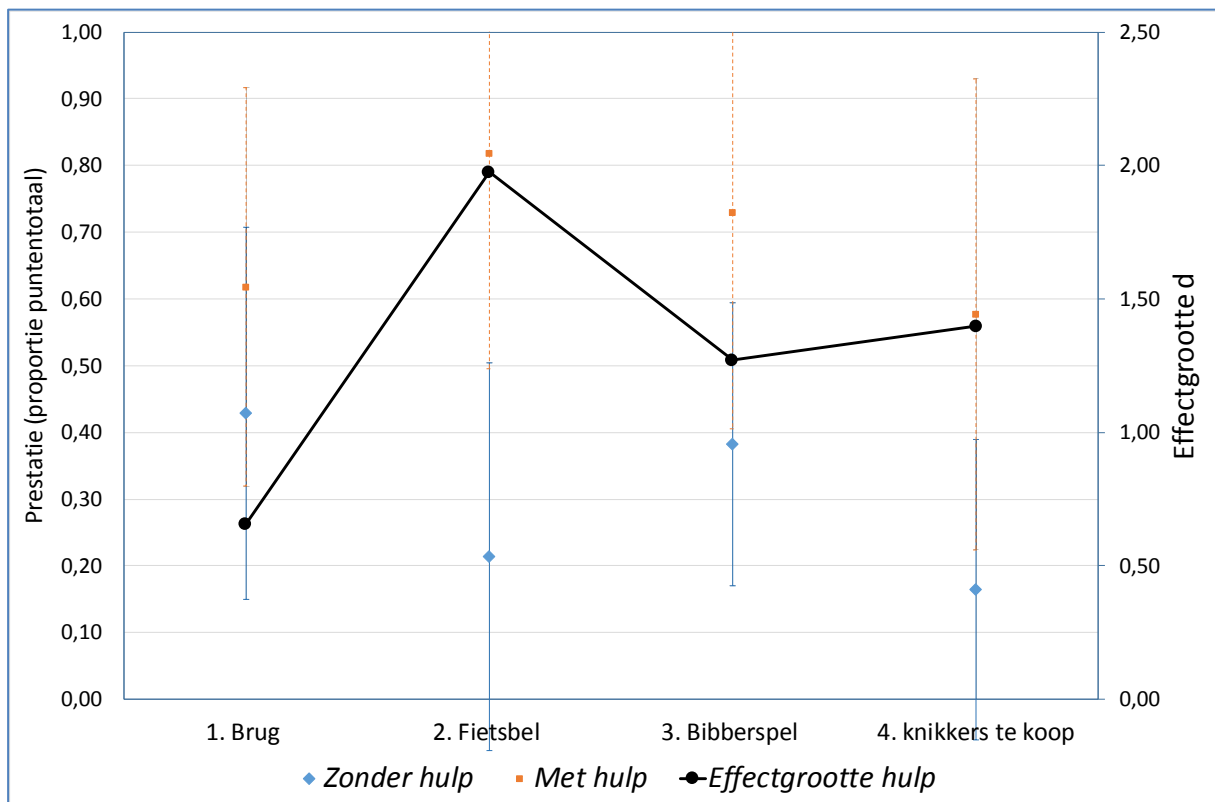
In de Figuren 4a en 4b kunnen de effectgroottes van hulp, afgebeeld met behulp van de bolletjes verbonden met behulp van een zwarte lijn, op de rechter verticale as worden afgelezen. Hieruit wordt duidelijk dat bij alle opdrachten, behalve de Brug, zeer hoge effectgroottes van hulp zijn gevonden, ruim boven de waarde .80. die wordt beschouwd als 'groot'. Daarnaast kunnen in Figuur 4a en 4b op de linker verticale as de gemiddelde scores worden afgelezen en ook de verdeling rondom de gemiddelden.

Hieruit wordt duidelijk dat bij alle opdrachten naast de gemiddelde scores zonder en met hulp ook een groot deel van de scores verschuift naar een hoger niveau. De blauwe lijn en oranje stippellijn vertegenwoordigen elk 68% van de scores rondom de gemiddelden. Kijkend naar deze verdeling, blijkt niet alleen dat leerlingen gemiddeld een sterke prestatieverbetering boeken na hulp, maar dat dat tevens geldt voor een groot deel van de leerlingen. Verder zien we dat op drie (deel)opdrachten de gemiddelde score na hulp, gevisualiseerd met behulp van een oranje rechthoekje, rond of boven de .80 ligt, een niveau dat vaak wordt gehanteerd om te spreken van beheersing van een

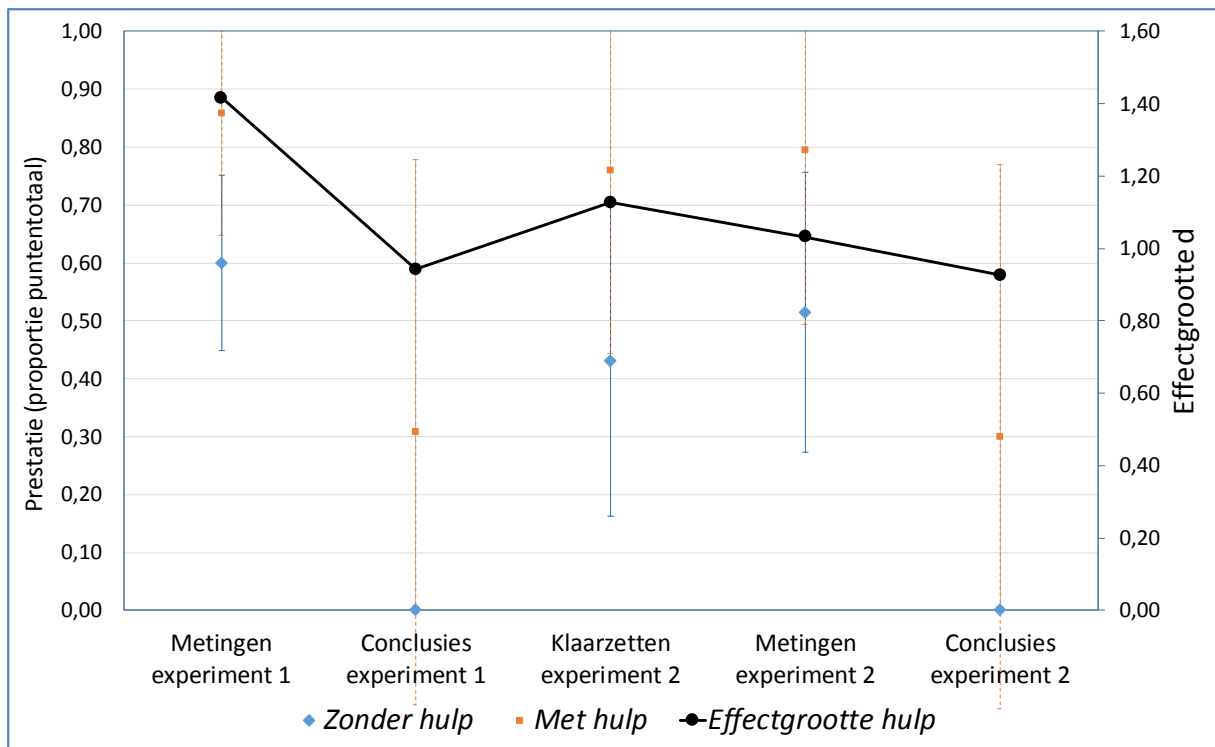
vaardigheid: het in elkaar zetten van de Fietsbel (.81, zie figuur 4a) en het verrichten van metingen bij resp. experiment 1 en 2 van de Knikkerbaan (.86 en .79, zie figuur 4b). Bij de opdracht Fietsbel bedroeg de score vóór hulp .22, hetgeen wijst op een grote vooruitgang. De scores op Knikkers te koop (.58), Brug (.62) en Bibberspel (.73) liggen na hulp onder dit niveau, maar de prestaties waren daarop ook lager voordat hulp was gegeven.

De effecten van hulp op de prestatieverbetering bij de opdrachten Fietsbel ($ES-v=1.98$), Bibberspel ($ES-v=1.27$), en Knikkers te koop ($ES-v=1.40$) zijn zeer groot (zie Tabel 12a).

Ook de effecten van hulp op de prestatieverbetering bij de deelopdrachten van de Knikkerbaan (zie Tabel 12b), verrichten van metingen experiment in 1 ($ES-v=1.42$) en 2 ($ES-v=1.03$), het klaarzetten van het experiment 2 ($ES-v=1.13$), conclusies trekken in experiment 1 ($ES-v=1.40$) en 2 ($ES-v=1.40$), zijn respectievelijk zeer groot en groot. De effecten op de prestatieverbetering bij de opdracht Brug ($ES-v=.66$) zijn middelgroot.



Figuur 4a: Prestaties op de praktische opdrachten en effectgroottes van hulp



Figuur 4b: Prestaties op de praktische opdracht Knikkerbaan en effectgroottes van hulp

Als we kijken naar de verschillen op de afzonderlijke prestatie-indicatoren (zie Tabellen 12a en 12b), dan valt op dat bij opdracht Fietsbel op alle indicatoren tegelijk een zeer groot effect wordt gevonden ($ES-v$ -waarden tussen 2.18 en 3.09). De inhoudelijke hint betrof hier het aanbieden van een andere fietsbel met een kleiner formaat.

Bij de opdracht Brug wordt alleen een kleine effect gevonden voor de indicator 'boogconstructie' ($ES-v=0.44$). Het grootste effect wordt gevonden op de indicator "de pijlers zijn goed opgesteld" ($ES-v=4.53$) en "er is één vel papier gebruikt" ($ES-v=2.39$). Bij deze laatste indicator werd de eis om één vel papier te gebruiken expliciet genoemd wat op zichzelf als modeling is aan te merken. Het juist neerzetten van de pijlers wordt via een verwijzing naar een foto ondersteund. Praktisch alle leerlingen doen dat goed. Op de overige prestatie-indicatoren van de opdracht Brug worden grote effecten gevonden, variërend tussen .86 (indicator "de brug draagt beide autootjes") en 1.18 (indicator "het bootje kan overal onder de brug door")

Bij het in elkaar zetten van het *Bibberspel* is het overall-effect van hulp op het juist monteren van de stroomdraden zeer groot ($ES-v= 1.27$). Bij deze opdracht werden successievelijk drie hints gegeven: 1. Toetsleider (TL) zegt: "de elektrische stroom moet rond gaan om het lampje te laten branden."; 2. TL zegt: "Volg met je vinger de stroom vanaf de batterij en ga zo na waar het probleem kan ontstaan"; 3. TL sluit één draad correct aan en zegt: "Ik sluit nu één draad goed aan. lukt het je nu?"). Met het goed vastdraaien van het lampje hadden maar 53 leerlingen problemen. Na hulp slagen 58% van hen hierin, wat gelijk staat met een effectgrootte van 1.68. Het juist monteren van de draden lukt 245 leerlingen niet zonder hulp. Met hulp slaagt 57% van hen wel in ($ES-v=1.63$).

Het overall-effect van hulp bij de opdracht *Knikers te koop* is zeer groot ($ES-v= 1.40$).

Wat betreft de afzonderlijke indicatoren zijn zeer grote effecten gevonden op het correct noteren van de meetpunten in het diagram ($ES-v$ -waarden resp. 1.89, 1.91, en 1.87). De effecten van hulp op het aanbrengen van labels (resp. $ES-v$ -waarden resp. 1.36, 1.32, en 1.33) zijn iets kleiner maar nog steeds zeer groot. Het effect van hulp bij het juist trekken van de trendlijn ($ES-v=.44$) door de meetpunten is klein. De inhoudelijke hulp betrof hier het aanbieden van een ingevuld diagram uit een andere context, waarbij de meetpunten *niet* verbonden waren door middel van een lijn.

Tabel 12a: Proporties behaalde scorepunten voor de praktische opdrachten vóór en na inhoudelijke hulp (alleen de leerlingen aan wie hulp is gegeven).

Prestatie-element	n	Prestatie voor inhoudelijke hulp		Prestatie na inhoudelijke hulp		Effect**	
		M	SD	M	SD	ES-v	t
Brug: een dragende brug maken van papier (n totaal = 315; n geholpen = 299; alpha somscore .80)	299	.43	.28	.62	.30	.66	13.31
1.1. De pijlers zijn goed opgesteld	45	.00	.00	.91	.29	4.53	21.24
1.2. Er is één vel papier gebruikt	116	.00	.00	.74	.44	2.39	18.16
1.3. Het papier is in profiel gevouwen	183	.00	.00	.33	.47	.99	9.42
1.4. Er is een boogconstructie bedacht	263	.00	.00	.09	.28	.44	5.01
1.5. Het bootje kan overal onder de brug door	153	.00	.00	.41	.49	1.18	10.32
1.6. De brug draagt het lichte autootje	180	.00	.00	.33	.47	1.00	9.46
1.7. De brug draagt het zware autootje	193	.00	.00	.32	.47	.97	9.53
1.8. De brug draagt beide autootjes	215	.00	.00	.27	.44	.86	8.89
Fietsbel: een fietsbel uit onderdelen in elkaar zetten (n totaal = 342; n geholpen = 197; alpha somscore .92)	197	0.21	0.29	0.82	0.32	1.98	22.39
2.1. De hendel om de bel te laten rinkelen en het kleine radertje zijn op de juiste manier aangebracht	145	0.00	0.00	0.82	0.38	3.03	25.67
2.2. Het balkje met de metalen ringetjes is op de juiste manier aangebracht	144	0.00	0.00	0.83	0.38	3.09	26.09
2.3. Het boutje om het balkje vast te zetten is op de juiste manier aangebracht	176	0.00	0.00	0.70	0.46	2.18	20.43
2.4. De dop is er op gedraaid; de bel rinkelt	143	0.00	0.00	0.80	0.40	2.80	23.63
Bibberspel: een gesloten stroomcircuit aanleggen, zodat spel functioneert (n totaal = 311; n geholpen = 246; alpha somscore .38)	246	0.38	0.21	0.73	0.32	1.27	16.26
3.1. lampje goed ingedraaid	53	0.00	0.00	0.58	0.50	1.68	8.56
3.2. De draden zijn goed aangesloten	245	0.00	0.00	0.57	0.50	1.63	18.04
Knikkers te koop: metingen in een diagram verwerken (n totaal = 281; n geholpen = 212; alpha somscore .90)	212	0.16	0.22	0.58	0.35	1.40	17.54
4.1. meetpunt kleine knikker op de juiste plaats in het diagram gezet	203	0.00	0.00	0.64	0.48	1.89	18.97
4.2. meetpunt middelgrote knikker op de juiste plaats in het diagram gezet	195	0.00	0.00	0.65	0.48	1.91	18.82
4.3. meetpunt grote knikker op de juiste plaats in het diagram gezet	192	0.00	0.00	0.64	0.48	1.87	18.25
4.4. Er is een trendlijn getrokken tussen de punten	261	0.00	0.00	0.09	0.28	.44	5.01
4.5. label BELLA op de juiste plaats in het diagram gezet	158	0.00	0.00	0.48	0.50	1.36	12.06
4.6. label ZORRO op de juiste plaats in het diagram gezet	153	0.00	0.00	0.46	0.50	1.32	11.47
4.7. label KINSI op de juiste plaats in het diagram gezet	155	0.00	0.00	0.47	0.50	1.33	11.71

** N.B. Betekenis van effectgroottes prestatieverbetering (ES-v): tussen 0 en .19; verwaarloosbaar effect; tussen .20 en .49; klein effect; tussen .50 en .79; middelgroot effect; tussen .80 en 1.29; groot effect; .1.3 en hoger: zeer groot effect

Tabel 12b: Proporties behaalde scorepunten voor de praktische opdracht Knikkerbaan vóór en na inhoudelijke hulp

Prestatie-element	n	Prestatie voor inhoudelijke hulp		Prestatie na inhoudelijke hulp		Effect**	
		M	SD	M	SD	ES-v	t
Klaarzetten experimentele opstelling experiment 2 (n totaal = 222; n geholpen =109; alpha totaal score=.69)	109	0.43	0.27	0.76	0.32	1.13	10.94
5.1 Helling gelijk ingesteld	68	0.00	0.00	0.74	0.44	2.36	13.64
5.2 Oppervlak gelijk ingesteld	58	0.00	0.00	0.50	0.50	1.41	7.55
5.3 Startpositie verschillend ingesteld	78	0.00	0.00	0.55	0.50	1.57	9.73
5.4 Gebruik knikker met gelijk gewicht	43	0.00	0.00	0.49	0.51	1.38	6.33
Uitvoeren metingen volgens vraagstellingen (experiment 1; (n totaal = 219; n geholpen =90; alpha totaalscore=.32)	90	0.60	0.15	0.86	0.21	1.42	12.39
5.5 Vijf metingen op beide banen	16	0.00	0.00	0.50	0.52	1.38	3.87
5.6 Bij <i>elke</i> meting zijn twee knikkers van gelijk gewicht gebruikt	16	0.00	0.00	0.56	0.51	1.56	4.39
5.7 In de tabel hebben de koppen een label (hoog of laag)	76	0.00	0.00	0.70	0.46	2.14	13.15
Uitvoeren metingen volgens vraagstellingen experiment 2; (n totaal = 223; n geholpen = 81; alpha totaalscore=.43)	81	0.51	0.24	0.79	0.30	1.03	13.54
5.5 Vijf metingen op beide banen	56	0.00	0.00	0.75	0.44	2.45	12.85
5.6 Bij <i>elke</i> meting zijn twee knikkers van gelijk gewicht gebruikt	8	0.00	0.00	0.38	0.52	1.10	2.05
5.7 In de tabel hebben de koppen een label (hoog of laag)	41	0.00	0.00	0.56	0.50	1.60	7.15
Passende conclusies trekken op grond van de verrichte metingen							
5.8 De conclusie is in overeenstemming met de gegevens uit de metingen (experiment 1; n totaal=218, n geholpen = 26)	26	0.00	0.00	0.31	0.47	0.94	3.33
5.8 De conclusie is in overeenstemming met de gegevens uit de metingen (experiment 1; n totaal=219, n geholpen = 20)	20	0.00	0.00	0.30	0.47	0.93	2.85
N.B.	** Betekenis van effectgroottes: tussen 0 en .19; verwaarloosbaar effect; tussen .20 en .49; klein effect; tussen .50 en .79; middelgroot effect; tussen .80 en 1.29; groot effect; 1.3 en hoger: zeer groot effect						

We vermeldde al eerder (paragraaf 3.1) dat bij de opdracht *Knikkerbaan* minder leerlingen hulp nodig hadden dan bij de andere praktische opdrachten. De gerapporteerde effectgroottes hebben bij sommige prestatie-indicatoren dan ook betrekking op relatief kleine groepen leerlingen.

Bij het klaarzetten heeft hulp een zeer groot positief effect op alle elementen, maar het sterkst op het instellen van de hellingshoek van de knikkerbaan ($ES-v=2.36$). Als leerlingen de startpositie (foutief) gelijk instelden, volgde de vraag: “als je wilt weten of de startpositie effect heeft, waar moet je de knikkers dan laten starten?” Bij (foutieve) verschillende instelling van een van de andere variabelen (helling, oppervlakte, of gebruik van knikkers van verschillend gewicht waar deze constant moesten blijven) werd de vraag gesteld: “Weet je nu of het alleen aan de startpositie ligt of een bal verder komt? Of kan het ook aan de andere instellingen liggen?”

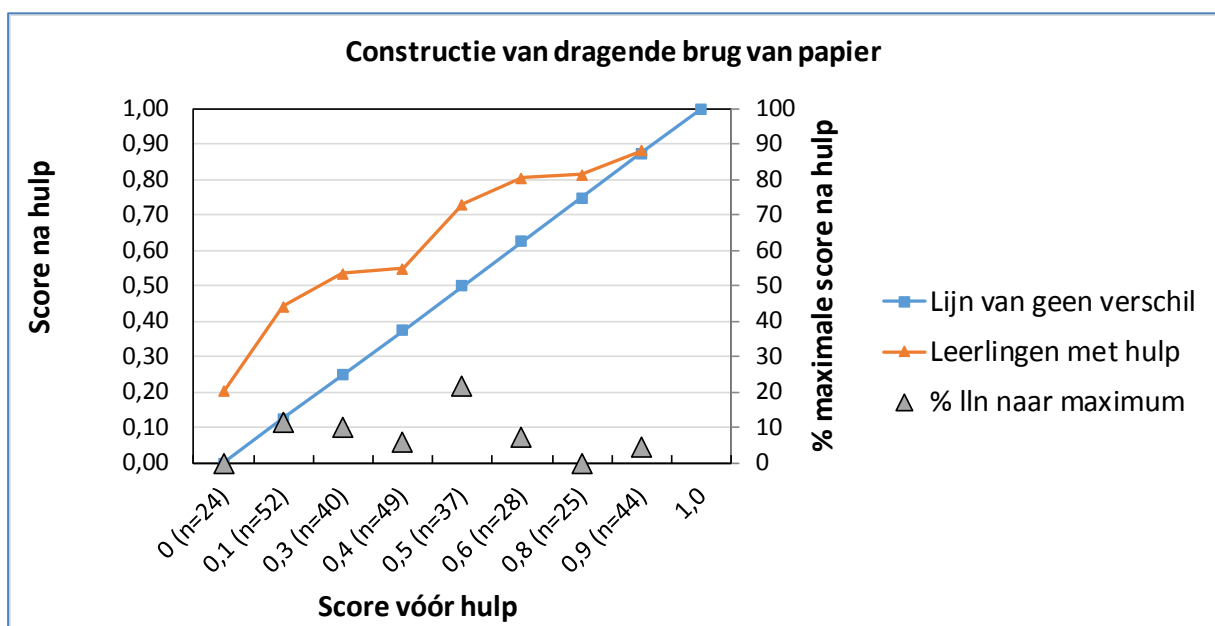
Daarnaast wordt bij de deelopdracht ‘uitvoeren metingen’ een zeer groot effect gevonden bij het correct vermelden van tabelkoppen ($ES-v=2.14$) in de datatabel bij experiment 1. De hint hierbij was de vraag: “Wat zou je achter ‘startpositie’ moeten invullen?”. Bij experiment 2 worden de grootste effecten geboekt op het consequent verrichten van vijf metingen ($ES-v=2.45$). Driekwart van de 56 geholpen leerlingen slaagden daarin na het verkrijgen van hulp (van score .00 naar .75.)

De effecten van hulp op het trekken van conclusies zijn iets minder groot dan bij de andere deelopdrachten, maar nog altijd groot (ES -v-waarden = 0.94 en 0.93). Opgemerkt moet worden dat slechts 26 resp. 20 leerlingen hulp nodig hadden bij deze deelopdracht. Dit betekent dat resp. 8 en 6 leerlingen er na hulp in slagen passende conclusies te formuleren.

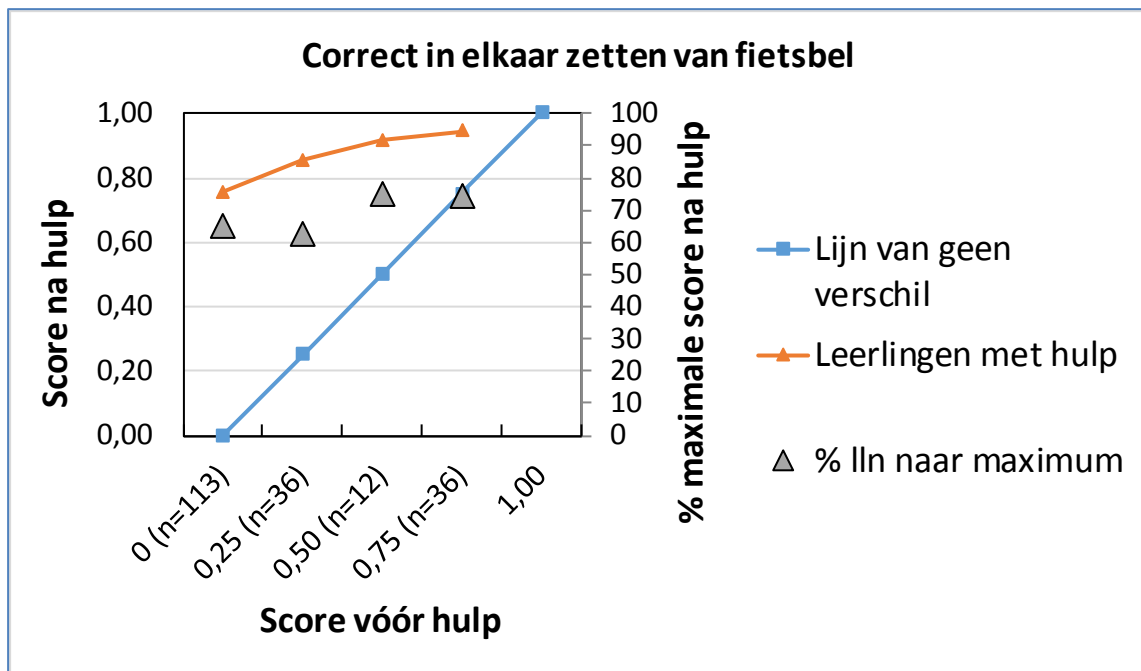
3.3.2 Differentiële effecten van hulp

In de Figuren 5 tot en met 9 is het verband weergegeven tussen de scores voorafgaand aan hulp en na hulp. Uit de figuren kan worden afgeleid of leerlingen met een verschillende beginscore (voorafgaand aan hulp) in verschillende mate profiteren van de hulp. Leerlingen met een maximale score zijn niet in de figuur opgenomen, omdat die uiteraard geen hulp nodig hadden. In de onderstaande figuren is de lijn “zonder hulp” gevormd door scorepunten in te vullen volgens de situatie wanneer leerlingen voor en na hulp dezelfde score zouden hebben behaald. Naarmate de lijn “met hulp” verder boven de lijn “van geen verschil”, is het effect van hulp dus groter. Ook is het percentage leerlingen afgebeeld dat na hulp de maximale score bereikt voor de opdracht. Dit zichtbaar gemaakt door middel van een driehoekje in de figuur.

Van alle praktische opdrachten blijken bij de opdracht Brug (Figuur 5) de verschillen in prestatie tussen leerlingen met uiteenlopende scores op de voormeting het minst afgenomen na afloop van de inhoudelijke hulpinterventies. De scores van leerlingen met een score vóór hulp van .40, .60 en .80, .90 zijn minder toegenomen dan leerlingen met een andere score. Bij de twee laatste genoemde leerlingen is er dan ook nog maar beperkt ruimte voor verbetering. De groepen met een andere score zijn in ongeveer gelijke mate vooruitgegaan (een winstscore van ongeveer .25). De correlatie tussen de scores voor en na hulp (van de geholpen leerlingen) bedraagt .68. De hulp gaat gepaard met percentages leerlingen tussen 0 en 20 die een brug vouwden die voldeed aan alle eisen van draagkracht (twee autootjes dragen) en doorvaarthoogte (bootje kan er onderdoor). Van de leerlingen die de minimale score behaalden zonder hulp, slaagde niemand erin een perfecte brug te vouwen. Leerlingen met een relatief hoge score (0.8 en 0.9) slaagden daar ook maar in resp. 3 en 4 % van de gevallen in.

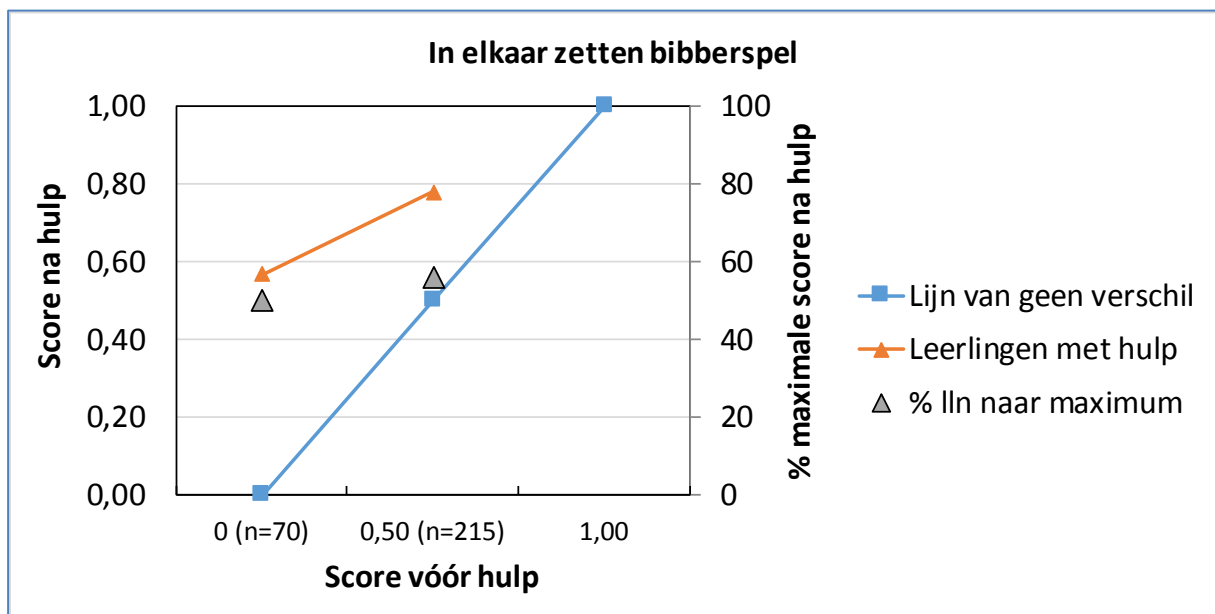


Figuur 5: Differentiële effecten van hulp op de opdrachtscore (opdracht Brug)



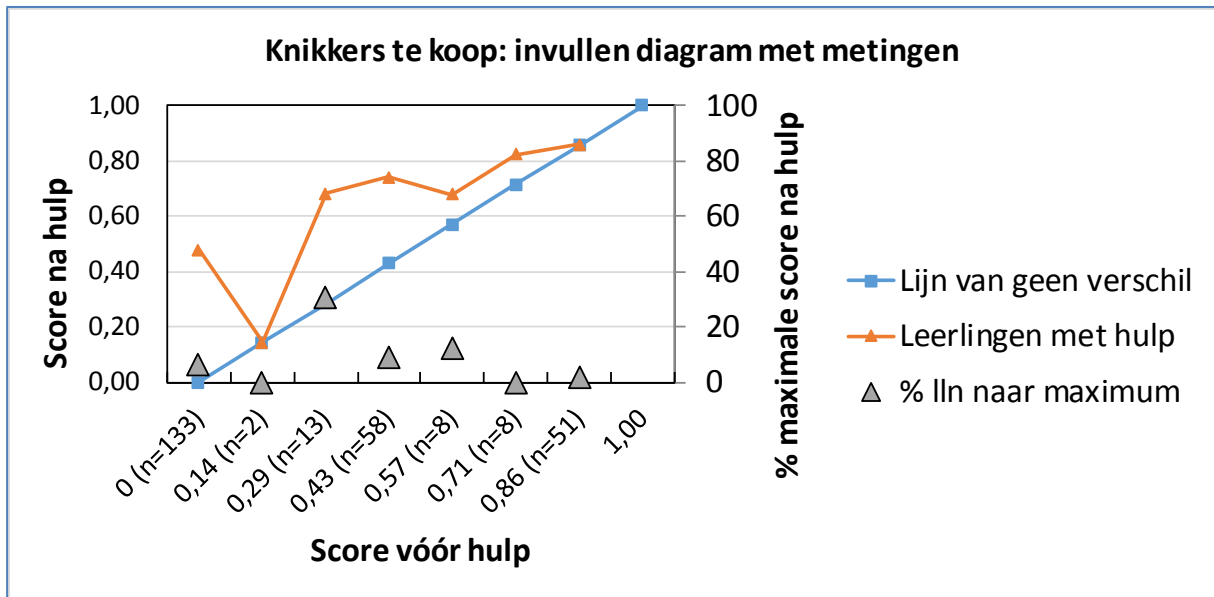
Figuur 6: Differentiële effecten van hulp op de opdrachtscore bij de opdracht Fietsbel

Kijkend naar de opdracht Fietsbel (Figuur 6) blijkt dat de scoreverschillen tussen leerlingen, die voorafgaand aan de hulp nog groot waren, sterk zijn afgenomen. De groep met beginscore 0, resp. 0.25, 0.50 en 0.75 scoren na de verkregen hulp 0.74, 0.86, 0.92 en 0.94. De correlatie tussen de score voor en na hulp bedraagt .24, hetgeen illustreert dat de rangschikking van leerlingen duidelijk gewijzigd is ten opzichte van de conditie zonder hulp. In alle onderscheiden scoregroepen bereikt meer dan 60% van de leerlingen na hulp de maximale score, hetgeen betekent dat de bel volledig correct in elkaar werd gezet.



Figuur 7: Differentiële effecten van hulp op de opdrachtscore bij de opdracht Bibberspel

Bij het Bibberspel (Figuur 7) nemen de scoreverschillen tussen de onderscheiden groepen (scores 0. en 0.5) sterk af na afloop van hulp. Resp. 50 en 56% van de leerlingen slaagt erin het elektrische circuit correct aan te sluiten. De correlatie tussen de score voor en na hulp bedraagt .28, hetgeen deze afname van verschillen en veranderingen in rangschikking illustreert.

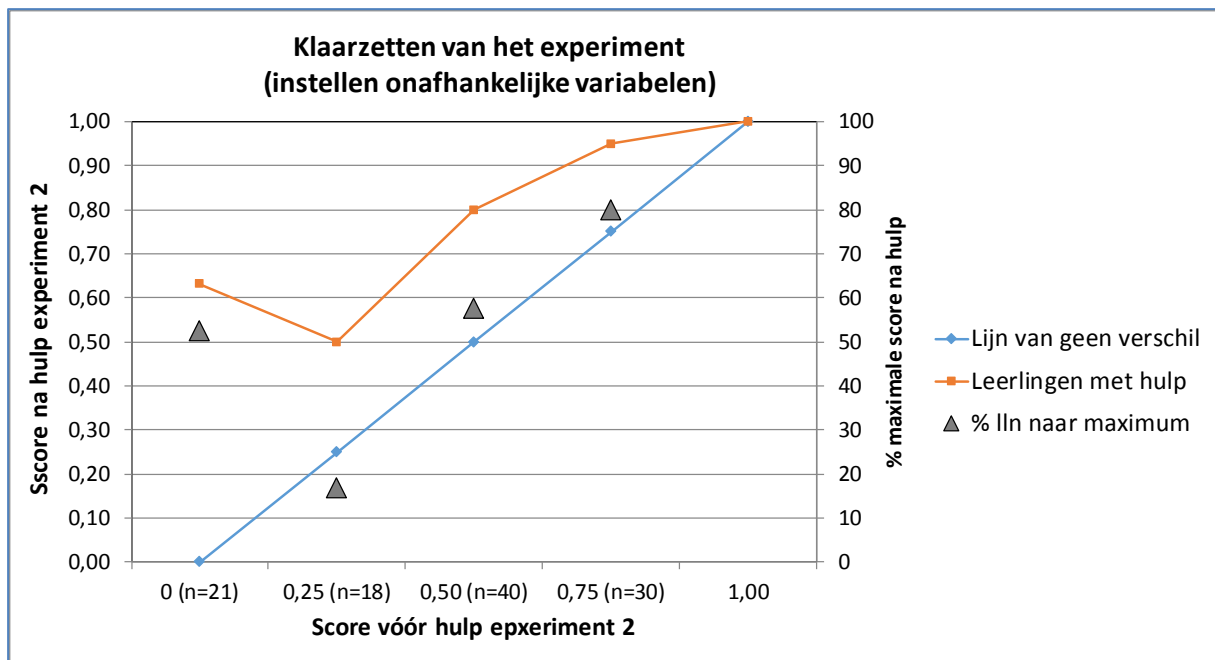


Figuur 8: Differentiële effecten van hulp op de opdrachtscore bij de opdracht Knikkers te koop

Ook bij de opdracht Knikkers te koop (Figuur 8) nemen de verschillen tussen de groepen af na hulp. Dit wordt bevestigd door de gevonden correlatie van .38 tussen score voor en na hulp.

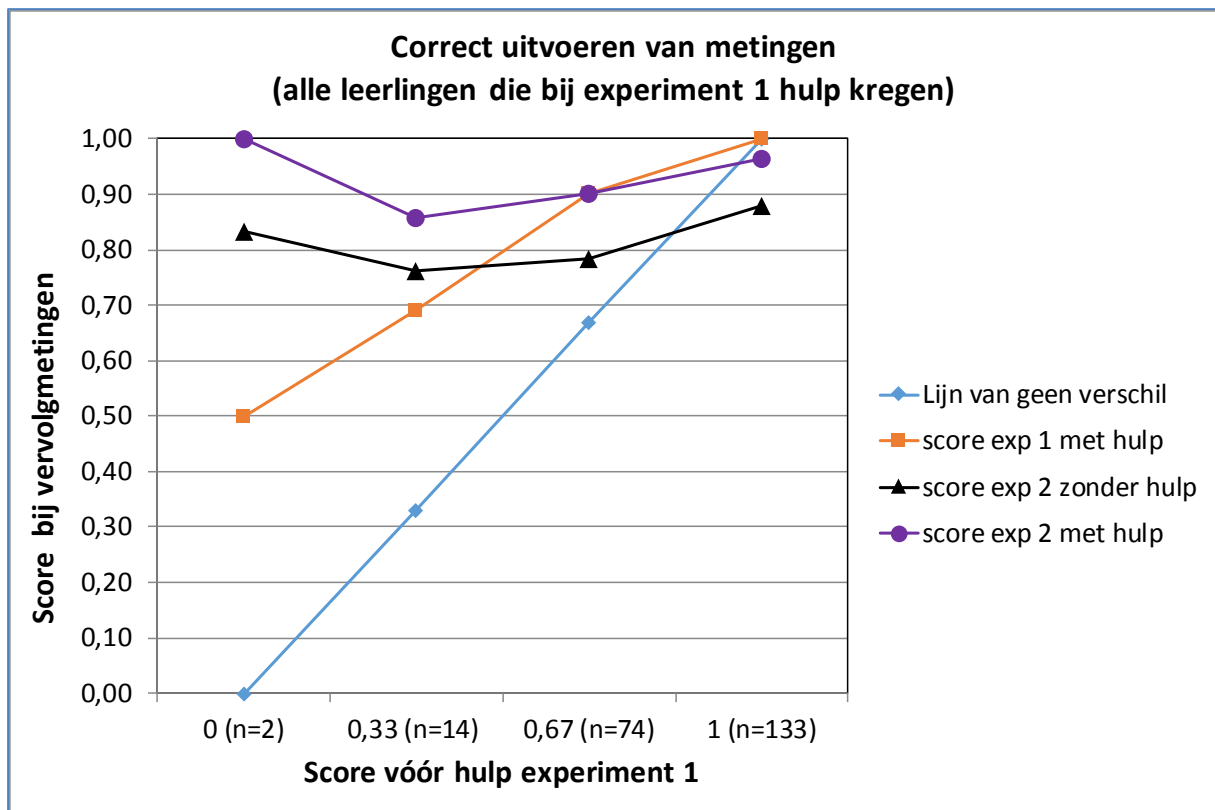
De scoregroepen profiteren in gelijke mate van de hulp. De prestatieverbetering van de scoregroepen 0.57, 0.71, en 0.86 blijft duidelijk achter (verbetering resp. .11, .11, en 0) bij die van 0, 0.29 en 0.43 (verbetering resp. .48, .40 en .31). Hierbij geldt wel een plafondeffect voor de twee groepen die het hoogst scores voordat hulp is gegeven.

Het aantal leerlingen dat na hulp het diagram correct en volledig heeft ingevuld is hoger in de scoregroep 0.29 dan in de overige groepen. Van de groep die zonder hulp 0.29 scoorde, behaalt 30% na hulp de maximale score. In de andere groepen liggen deze percentages tussen 0 en 11%.



Figuur 9a: Differentiële effecten van hulp op de opdrachtscore klaarzetten van het experiment bij de opdracht Knikkerbaan

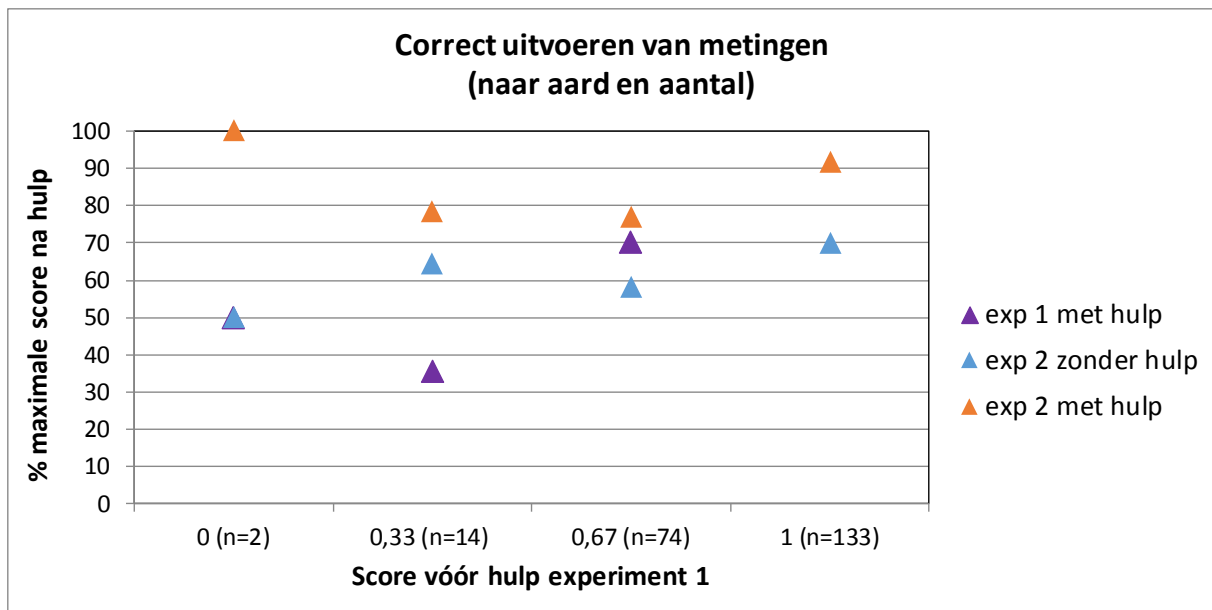
De figuren 9a tot en met 9c hebben betrekking op differentiële effecten van hulp bij de deelopdrachten van Knikkerbaan. Als we kijken naar de deelopdracht 'klaarzetten' van de experimentele opstelling, dan blijkt dat de rangschikking van leerlingen naar prestaties aanzienlijk is veranderd. De correlatie tussen scores vóór en na hulp bedraagt .43. De vier groepen die in de conditie zonder hulp scores behaalden van resp. 0, .25 en .50, .75 behalen na hulp scores van resp. .63, .50, .80 en .95. De prestatiegroep '.25' blijft daarmee in groei achter ten opzichte van de andere groepen. De groep 0 boekt een relatief grote vooruitgang naar een score van .62. Uiteindelijk bereikten de groepen resp. 52%, 17%, 58%, en 80% na hulp de maximale score.



Figuur 9b: Ontwikkeling in scores nadat in het eerste experiment hulp was gegeven bij het uitvoeren van metingen (transfer van eerste experiment)

Figuur 9b toont de ontwikkeling in scores op het verrichten van metingen van vier groepen leerlingen die op experiment 1 resp. 0, .33, .67 en 1.0 scoorden. De eerste groep bevat slechts 2 leerlingen. Uit deze figuur kunnen de volgende bevindingen worden afgelezen:

- De verschillen tussen de vier groepen nemen sterk af na hulp, getuige de scores op experiment 1 (scores resp. .50, .69, .90, 1.0 (de laatste score was al vóór hulp bereikt)). De correlatie tussen scores zonder en met hulp in experiment 1 bedraagt .44.
- De scores op de parallelopdracht in experiment 2 liggen in de condities zonder hulp nog dichter bij elkaar dan bij experiment 1: .83, .76, .78, .88. Daarbij valt het groepje met de maximale score zonder hulp in experiment 1 dus iets terug in prestatie. De correlatie tussen scores zonder hulp in experiment 1 en zonder hulp in het vervollexperiment 2 bedraagt .01.
- Na een tweede ronde van hulp verbeteren de scores in alle groepjes verder (resp. 1.0, .88, .90 en .96). De correlatie tussen scores zonder hulp in experiment 1 en met hulp in het vervollexperiment 2 bedraagt .02.



Figuur 9c: Percentages leerlingen die de maximale score halen onder verschillende hulpcondities bij uitvoeren van metingen voor de opdracht Knikkerbaan

Figuur 9c toont de percentages leerlingen die onder verschillende hulpcondities bij experiment 1 en 2 de maximale score behalen, en dus de metingen volledig correct uitvoeren. De resultaten laten het volgende zien:

- Van de scoregroepen die bij het eerste experiment hulp nodig hadden (met scores 0, .33, en .67), behaalt resp. 50%, 64% en 58% (blauwe driehoekjes) de maximale score nog vóórdit in experiment 2 hulp is gegeven.
- De hoogste prestatiegroep kent in experiment 2 zoals eerder beschreven een terugval. 30% van de leerlingen die in het eerste experiment de metingen correct uitvoerde, deed dat niet in het tweede experiment. Na hulp stijgt dat aantal weer naar 92%.
- De hulprondes in beide experimenten resulteren in toenames van het aantal leerlingen dat de metingen correct uitvoert (zie paarse driehoekjes en oranje driehoekjes voor hulp bij experiment 1 resp. 2). Dit geldt voor alle prestatiegroepen. Na hulp bij experiment twee stijgen deze percentages voor de vier groepen tot 100%, 79%, 77% en 92%.

3.3.3 Effecten van verschillende soorten hulp en hulpcombinaties

Tabel 13 geeft het verband weer tussen de verkregen hulpsoort en de prestatieverbetering (proportie scorepunten na hulp minus proportie scorepunten voor hulp) voor de praktische opdrachten. Het betreft hier dus alleen leerlingen die inhoudelijke hulp hebben ontvangen. Daar waar de inhoudelijke hulp uit meerdere hints kon bestaan, kon het verband worden gelegd tussen het aantal hints en de prestatieverbetering.

Het aantal ontvangen hints correleert positief met de prestatieverbetering op alle praktische opdrachten waarbij sprake kon zijn van meer dan één hint. De waarden variëren tussen .21 (Knikkers te koop) tot .41 (Knikkerbaan: klaarzetten opstelling experiment 2).

Daarboven blijkt er ook een positieve correlatie tussen het verkrijgen van proceshulp en de prestaties op Brug ($r=.40$), Fietsbel ($r=.18$), Bibberspel ($r=.50$), klaarzetten experimentele opstelling ($r=.28$) en conclusies trekken op grond van metingen op de Knikkerbaan ($r=.64$). Het kleine groepje dat er niet in slaagt conclusies te trekken heeft kennelijk baat bij proceshulp. Bij geen van de

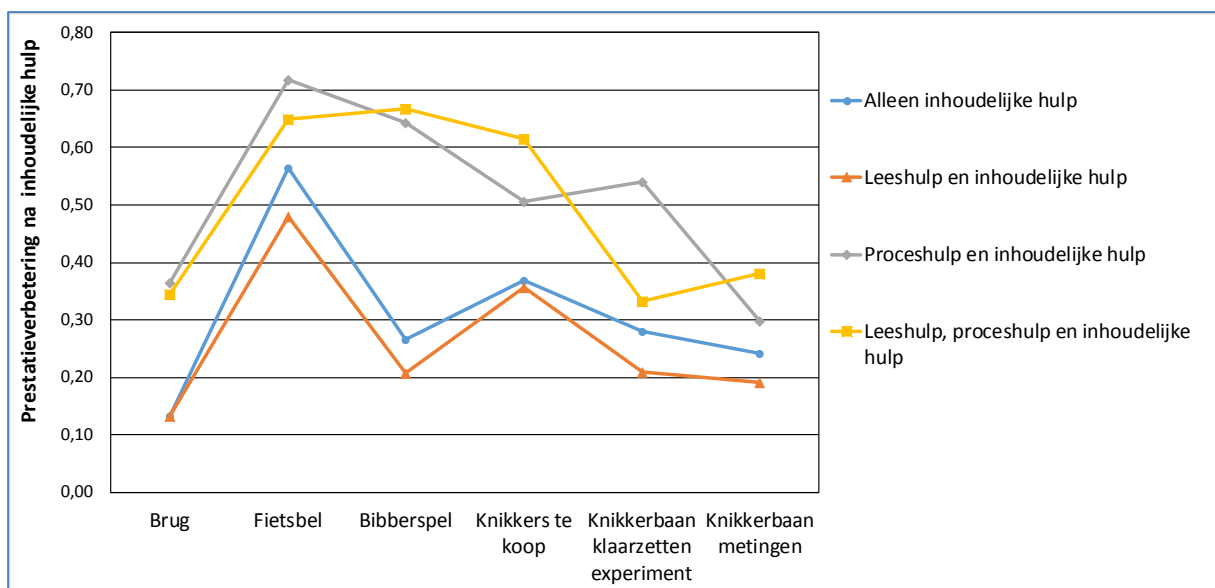
praktische opdrachten is sprake van een significante correlatie tussen leeshulp en de prestatieverbetering.

Tabel 13: Samenhang tussen de prestatieverbetering na het verkrijgen van inhoudelijke hulp en het al of niet verleend zijn van leeshulp, proceshulp en het aantal gegeven hints.

Prestatieverbetering op opdracht:	N	Leeshulp	Proceshulp	Aantal inhoudelijke hints
1. Brug	294		.40**	.38**
2. Fietsbel	195		.18*	
3. Bibberspel	246		.50**	.37**
4. Knikkers te koop	212		.23**	.21**
5a Knikkerbaan: klaarzetten opstelling experiment 2	109		.28**	.41**
5bc. Knikkerbaan: metingen doen, experiment 1	88			.29**
5bc. Knikkerbaan: metingen doen, experiment 2	76			.40**
5d. Conclusies trekken Knikkerbaan, experiment 1	26		.64**	NB
5d. Conclusies trekken Knikkerbaan experiment 2	20			NB

N.B. * $p < .05$; ** $p < .01$; NB: Niet beschikbaar, omdat hier sprake was van één inhoudelijke hint.

Figuur 10 biedt een exploratief inzicht in de samenhang tussen de combinaties van hulp en de prestatieverbetering op de opdrachten. Het volgende beeld komt naar voren. Bij alle opdrachten blijken leerlingen die een combinatie van inhoudelijke hulp en proceshulp krijgen een grotere prestatieverbetering te boeken dan leerlingen die uitsluitend inhoudelijke hulp ontvangen. Leerlingen die naast inhoudelijke hulp en proceshulp ook leeshulp ontvangen, bereiken bovendien op de opdracht Knikkers te koop een nog grotere prestatieverbetering dan de groep met inhoudelijke hulp en proceshulp (.64 versus .51). Op drie andere opdrachten (Brug, Fietsbel, Bibberspel) scoren leerlingen met deze hulpcombinatie ongeveer gelijk aan de 'inhoud en proces hulp'-groep. Bij de deelopdracht Klaarzetten van de experimentele opstelling verbetert de groep met deze hulpcombinatie juist beduidend minder dan de groep met de combinatie inhoudelijke en proceshulp (.34 versus .54). De groep die alle hulpvormen krijgt, is overigens steeds klein van omvang (n 's per opdracht resp. in volgorde op figuur 10: 21, 10, 15, 20, 7, 9). De groep leerlingen die leeshulp én inhoudelijke hulp ontvangt, laat op alle vijf opdrachten een iets lagere prestatieverbetering zien dan de groepen met andere hulpcombinaties.



Figuur 10: Prestatieverbetering bij verschillende hulpcombinaties voor de vijf praktische opdrachten

3.3.4 Verschillen tussen leerlingen bij wie wel en geen modeling plaatvond na inhoudelijke hulp

In een volgende analyse is gekeken naar verschillen in prestatieverbetering en het bereikte scoreniveau na hulp tussen leerlingen bij wie de oplossing wel en niet gemodelleerd is. (zie Tabel 14a en Tabel 14b). Deze analyse zegt iets over tot hoever leerlingen met hulp waren gevorderd met hun opdracht op het moment waarop al of niet besloten moest worden tot het modelleren van de volledige oplossing. Opgemerkt moet worden dat het al of niet vóórkomen van modellering is gescoord voor de hele opdracht en niet per prestatie-indicator. Het kan dus zo zijn dat leerlingen op één of meerdere prestatie-indicatoren achterbleven, voordat op modeling is overgegaan. De analyse geeft vooral trends weer, waarbij grote scoreverschillen op indicatoren aanduiden dat leerlingen die modeling kregen daarop achterblijven bij leerlingen zonder modeling.

Bij de opdracht Brug zetten leerlingen die later de oplossing gemodelleerd krijgen de pijlers goed neer (indicator 1.1) en gebruiken ze desgevraagd één vel papier (indicator 1.2). De prestatieverbetering na hulp verschilt niet van die van de groep zonder modeling (*ES-v-waarden* -.23 en .10). Het vouwen van een profiel (indicator 1.3) dat bovendien sterk genoeg is om zowel een licht (indicator 1.6) als een zwaar autootje (indicatoren 1.7 en 1.8) te dragen zijn activiteiten waarop deze groep leerlingen na hulp minder vorderen en laag presteren (*ES-v-waarden* resp. .64, .60, .79, en .91). Opmerkelijk is verder dat slechts een klein aandeel leerlingen die inhoudelijke hulp ontvangt, erin slaagt een boogconstructie te vouwen (indicator 1.4; *M*=.28 en *M*=.09. resp. wel en geen modeling). Het verschil in prestatieverbetering tussen de groepen die hierin wel en niet slaagde na het ontvangen van inhoudelijk hulp is klein (*ES-v*=.33. zie Tabel 14).

Bij de opdracht Fietsbel zijn de verschillen in bereikte score en de omvang van de prestatieverbetering tussen de groepen zonder modeling en met modeling op alle prestatie-indicatoren groot tot zeer groot (*ES-v-waarden* tussen .81 en 1.58). De groep die aangewezen is op modeling vordert met name in geringere mate op de montage van de bel waarbij het balkje met metalen ringetjes goed is vastgezet middels het bijgeleverde boutje (indicator 2.3, prestatieverbetering *M*= .15 tegen .78 zonder modeling).

Bij de opdracht Bibberspel profiteren leerlingen die modeling nodig hebben niet van de eerder gegeven inhoudelijke hulp bij het maken van een gesloten stroomcircuit (indicator 3.2; prestatieverbetering *M*= .02 tegen .97 zonder modeling). Ze blijven daarmee ver achter bij de groep die wel profiteert (*ES-v-waarde* 5.29). Wel draaien ze na hulp in meerderheid het lampje van het Bibberspel goed vast (scores na hulp resp. *M*=.99 en *M*=.79. zonder en met modeling).

Bij de opdracht Knikkers te koop moeten leerlingen de gewichten van knikkers met verschillende diameters invullen in een diagram. Op zes van de zeven prestatie-indicatoren blijft de groep die modeling ontving in prestatieverbetering sterk achter bij hen die geen modeling ontving (*ES-v-waarden* variërend van .80 tot 1.30). Resp. 29%, 41% en 38% van de modeling-groep slaagt erin het meetpunt van drie knikkers op de juiste plaats te zetten (indicatoren 4.1, 4.2, en 4.3; scores na hulp resp. *M*= .29, *M*=.41, en *M*= .38). Iets kleinere tot vergelijkbare aandelen van de modeling-groep slagen erin de labels van de knikkers op de juiste plaats te zetten (indicatoren 4.5, 4.6, en 4.7; scores na hulp resp. *M*= .25, *M*=.32, en *M*= .29). Opmerkelijk is dat in beide groepen een gering aandeel (*M*= .15 en *M*=.13. resp. zonder en met modeling) een trendlijn door de punten trekt. Mogelijk komt dit doordat de trendlijn ook ontbreekt in het voorbeelddiagram (zie Figuur 1) over het verband tussen lengte en schoenmaat, dat als hulp werd gepresenteerd.

Tabel 14a: Verschillen in winstcores tussen leerlingen bij wie de oplossing wel en niet is gemodelleerd nadat de voorgeschreven inhoudelijke hulp was gegeven.

Opdracht en prestatie-indicator	Eindoplossing niet gemodelleerd				Eindoplossing wel gemodelleerd				Effect-groottes*	
	Score na hulp		Prestatie-verbetering		Score na hulp		Prestatie-verbetering		Effect-groottes*	
	M	SD	M	SD	M	SD	M	SD	ES-p	ES-v
Brug: een dragende brug maken van papier (n=192; n=152)										
1.1. De pijlers zijn goed opgesteld	1.00	.00	.11	.31	.97	.17	.18	.39	.27	-.23
1.2. Er is één vel papier gebruikt	.98	.14	.31	.46	.80	.40	.27	.44	.63	.10
1.3. Het papier is in profiel gevouwen	.86	.35	.31	.46	.25	.44	.07	.25	1.57	.64
1.4. Er is een boogconstructie bedacht	.28	.45	.12	.32	.09	.29	.03	.17	.50	.33
1.5. Het bootje kan overal onder de brug door	.96	.21	.33	.47	.38	.49	.07	.25	1.59	.67
1.6. De brug draagt het lichte autootje	.94	.24	.31	.46	.19	.39	.07	.26	2.37	.60
1.7. De brug draagt het zware autootje	.92	.26	.35	.48	.12	.33	.04	.21	2.75	.79
1.8. De brug draagt beide autootjes	.82	.39	.35	.48	.06	.24	.01	.12	2.31	.91
Fietsbel: een fietsbel uit onderdelen in elkaar zetten (n=174; n=55)										
2.1. Het kleine radertje is op de juiste manier aangebracht	.97	.16	.70	.46	.52	.51	.33	.47	1.60	.81
2.2. Het balkje met de metalen ringetjes is op de juiste manier aangebracht	.98	.14	.71	.46	.52	.51	.30	.47	1.68	.88
2.3. Het boutje om het balkje vast te zetten is op de juiste manier aangebracht	.89	.31	.78	.41	.22	.42	.15	.36	2.00	1.58
2.4. De dop is er op gedraaid; de bel rinkelt	.98	.14	.69	.46	.42	.50	.24	.43	2.06	.99
Bibberspel: een gesloten stroomcircuit aanleggen. zodat spel functioneert (n=158; n=127)										
3.1. lampje goed ingedraaid	.99	.09	.21	.41	.79	.41	.02	.14	.72	.60
3.2. De draden zijn goed aangesloten	.97	.18	.97	.18	.03	.17	.02	.14	5.29	5.73
Knikkers te koop: metingen in een diagram verwerken (n=163; n=86)										
4.1. meetpunt kleine knikker op de juiste plaats in het diagram gezet	.83	.38	.79	.41	.29	.46	.25	.44	1.31	1.30
4.2. meetpunt middelgrote knikker op de juiste plaats in het diagram gezet	.80	.40	.74	.44	.41	.50	.29	.46	.89	1.01
4.3. meetpunt grote knikker op de juiste plaats in het diagram gezet	.80	.40	.72	.45	.38	.49	.28	.45	.99	.98
4.4. Trendlijn getrokken tussen de punten	.15	.35	.13	.33	.13	.34	.04	.21	.03	.27
4.5. label BELLA op juiste plaats in diagram	.83	.38	.46	.50	.25	.44	.10	.31	1.45	.81
4.6. label ZORRO op juiste plaats in diagram	.79	.41	.44	.50	.32	.47	.09	.29	1.09	.80
4.7. label KINSI op juiste plaats in diagram	.80	.40	.45	.50	.29	.46	.09	.29	1.22	.83

N.B. ES-p= effectgrootte prestatieverschil; ES-v: effectgrootte verschil in prestatieverbetering. Betekenis van effectgroottes: tussen 0 en .19; verwaarloosbaar effect; tussen .20 en .49; klein effect; tussen .50 en .79; middelgroot effect; tussen .80 en 1.29; groot effect; 1.3 en hoger: zeer groot effect

Tabel 14b: Verschillen in winstscores tussen leerlingen bij wie de oplossing (deelopdrachten Knikkerbaan) wel en niet is gemodelleerd nadat de voorgeschreven inhoudelijke hulp was gegeven.

	Eindoplossing niet gemodelleerd				Eindoplossing wel gemodelleerd					
	Score na hulp		Prestatie-verbetering		Score na hulp		Prestatie-verbetering		Effect-groottes*	
Opdracht en prestatie-indicator	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>ES-p</i>	<i>ES-v</i>
Klaarzetten experimentele opstelling (alleen experiment 2; n= 75; n= 32)										
5.1 Helling gelijk ingesteld	.96	.20	.55	.50	.53	.51	.25	.44	1.35	.62
5.2 Oppervlak gelijk ingesteld	.86	.35	.40	.57	.44	.50	.28	.68	1.04	.20
5.3 Startpositie verschillend ingesteld	.83	.38	.67	.53	.31	.47	.25	.72	1.28	.71
5.4 Gebruik knikker met gelijk gewicht	.89	.31	.32	.55	.56	.50	.00	.72	.89	.53
Uitvoeren metingen volgens vraagstellingen van een experiment (experiment 1; n=71; n=18)										
5.5 Vijf metingen op beide banen	.96	.20	.10	.30	.72	.46	.06	.24	.87	3.27
5.6 Bij <i>elke</i> meting zijn twee knikkers van gelijk gewicht gebruikt	.93	.26	.11	.32	.89	.32	.06	.24	.15	.57
5.7 In de tabel hebben de koppen een label (hoog of laag)	.85	.36	.70	.46	.33	.49	.11	.24	1.32	4.96
Uitvoeren metingen volgens vraagstellingen van een experiment (experiment 2; n=59; n=17)										
5.5 Vijf metingen op beide banen	.86	.35	.49	.50	.94	.24	.76	.44	-.24	-.56
5.6 Bij <i>elke</i> meting zijn twee knikkers van gelijk gewicht gebruikt	.93	.26	.04	.19	.94	.24	.06	.24	-.05	-.12
5.7 In de tabel hebben de koppen een label (hoog of laag)	.83	.38	.33	.47	.51	.51	.24	.44	.77	.20

N.B. ES-p= effectgrootte prestatieverschil; ES-v: effectgrootte verschil in prestatieverbetering. Betekenis van effectgroottes: tussen 0 en .19; verwaarloosbaar effect; tussen .20 en .49; klein effect; tussen .50 en .79; middelgroot effect; tussen .80 en 1.29; groot effect; 1.3 en hoger: zeer groot effect

De deelopdrachten bij Knikkerbaan hebben betrekking op het instellen van de onderzoeksvariabelen, het uitvoeren van metingen van de afgelegde afstand van twee knikkers en het trekken van conclusies die in overeenstemming zijn met de verrichte metingen.

Leerlingen aan wie de volledige oplossing van de opdracht moest worden gemodelleerd, scoren vlak voor modellering beduidend lager op de instelling van alle onafhankelijke variabelen (*ES-p-waarden*: 1.35, 1.04, 1.28, .89). De prestatieverbetering van deze groep (resp. .25, .28, .25, .00) blijft achter ten opzichte van de groep zonder modeling (resp. .55, .40, .67, .32). De effectgrootte voor prestatieverbetering op indicator 5.2 "Oppervlak gelijk ingesteld" is klein (*ES-v*: .20). De effectgroottes voor de overige indicatoren zijn matig (*ES-v*-waarden: .62, .71, .53).

Wat betreft uitvoeren van metingen volgens de vraagstellingen ontstaat het volgende beeld. Op de prestatie-indicatoren 5.5 (minimaal vijf metingen) en 5.7 (het gebruiken van passende tabelkoppen) zijn in experiment 1 duidelijke prestatieverschillen tussen de groep die wel en geen modeling kreeg (*ES-p-waarden*: .87 en 1.32). De modeling-groep verbetert in experiment 1 duidelijk minder op deze prestatie-indicatoren (*ES-v*: 3.27 en 4.96). De verschillen in prestatieverbetering tussen de groepen zijn veel kleiner als het gaat om indicator 5.6 (knikkers van gelijk gewicht gebruiken, (*ES-v*: .57), waarop ook nauwelijks prestatieverschillen bestaan tussen de groepen (*ES-p*: .15).

In experiment 2 daarentegen, zijn de verschillen tussen de groepen die wel en geen modeling ontvingen bij het uitvoeren van metingen volgens de vraagstellingen veel kleiner. Beide groepen scoorden na hulp hoog op indicator 5.5 (minimaal 5 metingen uitvoeren: *M* resp. .86 en .94; *ES-p*: -

.24) en indicator 5.6 (knikkers van gelijk gewicht gebruiken: M resp. .93 en .94; $ES-p$: -.05). De prestatieverbetering is in de modeling-groep zelfs iets groter dan in de niet-modeling groep ($ES-v$ -waarden: -.56 en -.12). Alleen op indicator 5.7 (het gebruiken van passende tabelkoppen) zien we dat de modeling-groep achterblijft bij de niet-modeling groep ($ES-p$ =.77). Het verschil in prestatieverbetering is gering ($ES-v$: .20).

Voor prestatie-indicator 5.8 (trekken van juiste conclusies) hoeft deze analyse niet te worden verricht, omdat leerlingen die hierop 0 scoren per definitie modeling ontvangen en leerlingen die geen modeling ontvangen 1. Van de 26 leerlingen in experiment hadden 18 leerlingen modeling nodig en 8 niet. Bij experiment 2 bedroegen deze aantallen 14 en 6.

3.3.5 Leerlingenmerken en effecten van hulp

In Tabel 15 is via correlatiecoëfficiënten zichtbaar gemaakt in hoeverre de omvang van de bereikte prestatieverbetering na hulp varieert met leerlingenmerken. Omdat de te boeken prestatieverbetering bij lagere beginscores groter is dan bij hoge beginscores, zijn de analyses afzonderlijk verricht voor groepen leerlingen met een zelfde beginscore. Groepjes met minder dan 30 leerlingen zijn hierbij buiten beschouwing gelaten, omdat bij deze aantallen de power van significantietoetsing te gering is. Uiteraard zijn ook groepjes die zonder hulp al de maximale score hadden buiten beschouwing gelaten

Bij de meeste opdrachten en bij de meeste prestatiegroepen is er geen verband tussen geslacht en leeftijd enerzijds en de mate van bereikte prestatieverbetering anderzijds. Leerlingen van verschillende leeftijden profiteren over het algemeen in gelijke mate van de gegeven hulp.

De volgende bevindingen vormen hierop een uitzondering:

- Meisjes met een beginscore van 4 (uit 7) op de opdracht Knikkerbaan boeken iets grotere prestatieverbetering dan jongens met dezelfde beginscore (r =-.28).

Ook verschillen in attitude, te weten vrijetijdsbesteding aan N&T, het toegekende belang aan NT. plezier hebben in N&T en N&T moeilijk vinden vertonen geen duidelijke samenhang met prestatieverbetering na hulp in de verschillende prestatiegroepen.

De volgende significante verbanden worden gevonden:

- Naarmate leerlingen met een beginscore 3 (uit 8) op de Brug aangaven N&T moeilijk te vinden was de prestatieverbetering geringer (r =-.43)
- Omgekeerd, naarmate leerlingen met een beginscore 3 (uit 7) op Knikkers te koop aangaven N&T moeilijk te vinden was de prestatieverbetering juist groter (r =.28).

Een meer consistent verband is gevonden tussen het zelfbeeld van leerlingen over de eigen ontwerp- en onderzoeksvaardigheden met betrekking tot de opdracht en de mate van prestatieverbetering op die opdracht. Bij de opdrachten Brug (beginscores 1, 2, 3, 4 uit 8; r -waarden .51, .48, .40, en .33). Fietsbel (beginscores 0 en 3 uit 4; r -waarden .28 en .54), Bibberspel (beginscores 0 en 1 uit 2; r -waarden .32 en .18) en Knikkerbaan (beginscore 6 uit 8; r = .51) zien we dat naarmate leerlingen een grotere prestatieverbetering bereiken ze een positiever beeld rapporteren over de eigen ontwerp- en onderzoeksvaardigheden die nodig zijn bij het uitvoeren van deze opdrachten. Aangezien het om een correlatie gaat, kan ook andersom worden gesteld dat naarmate leerlingen een hoger zelfbeeld hebben, ze een grotere prestatieverbetering bereiken.

Bij Knikkers te koop is het verband tussen zelfbeeld en prestatieverbetering niet consistent. Bij leerlingen met beginscore 0 (uit 7) geldt een positief verband ($r = .21$). Bij leerlingen met een beginscore van 6 (uit 7) is zelfbeeld negatief gecorreleerd met prestatieverbetering ($r = -.37$). Het verband tussen het kennisniveau van Natuur en Techniek en prestatieverbetering, uitgedrukt in de vaardigheidsscore op de kennistest, is niet consistent.

- Bij Knikkers te koop bestaat in de hoogste prestatiegroep met hulp (score 6 uit 7) een negatief verband tussen kennisniveau en prestatieverbetering ($r = -.30$).
- Bij Knikkerbaan geldt in twee prestatiegroepen (resp. 3 uit 8 en 6 uit 8) een positief verband tussen het kennisniveau en de bereikte prestatieverbetering ($r = .36$ en $r = .59$). Was dat kennisniveau hoger, dan was hun prestatieverbetering groter.

De analyse op de prestatieverbetering in verschillende scoregroepen laat geen uitspraken toe over het verband tussen sociaal-economische afkomst (leerlinggewicht) en prestatieverbetering, omdat de aantallen daarvoor te gering waren.

Tabel 15: Verband tussen leerlingkenmerken en prestatieverbetering op de opdrachten per prestatiegroep

Beginscore per opdracht	N	Geslacht	Leeftijd	Vrije tijd NT	Belang NT	Plezier NT	Moeilijk NT	Toekomst NT	Zelfbeeld vaardig- heden	Vaardigheids- score op NT kennistest
De Brug (0-8)										
1	47								.51**	
2	37								.48**	
3	48						-.43**		.40**	
4	35								.33*	
7	44									
Fietsbel (0-4)										
0	104								.28**	
1	36									
3	36								.54**	
Bibberspel (0-2)										
0	55								.32*	
1	181								.18*	
Knikkers te koop (0-7)										
0	128								.21*	
3	56						.28*			
6	49								-.37**	-.30*
Knikkerbaan (hele experiment (0-8)										
3	44									.36*
4	49	-.28*								
5	42									
6	30								.51**	.59**

N.B. * $p < .05$; ** $p < .01$; Significante positieve correlaties zijn groen gearceerd en negatieve rood.

4 Conclusies

4.1 Soorten hulp (onderzoeksvraag 1)

Bij de eerste onderzoeksvraag stond centraal welke vormen van hulp verstrekt worden om leerlingen een stap verder te brengen in de oplossing van een onderzoeks- of ontwerpprobleem. De volgende conclusies kunnen worden getrokken. Bij de eerste onderzoeksvraag stond centraal welke vormen van hulp verstrekt worden om leerlingen een stap verder te brengen in de oplossing van een onderzoeks- of ontwerpprobleem. De volgende conclusies kunnen worden getrokken.

1. **De aard en frequentie van verstrekte hulp** loopt sterk uiteen per opdracht. Veruit het vaakst werd hulp verleend bij de constructie van de papieren brug (96.2%). Veruit het minst bij de diverse deelopdrachten van de opdracht Knikkerbaan in experiment 1 en 2 (metingen 55.4% en 47.7%, conclusies trekken 18.7 en 12.4%; klaarzetten experiment 2: 57.4%).
2. **De frequentie van specifieke combinaties van hulpsoorten** varieert sterk per opgave: Bij de opdracht Knikkerbaan wordt het vaakst uitsluitend leeshulp en/of proceshulp verstrekt, los van inhoudelijke hulp. Bij de brug komt deze hulpvorm afzonderlijk nauwelijks voor.
3. **Inhoudelijke hulp** maakt bij de meeste opdrachten rond 50% van alle interventies uit. Bij Bibberspel (onder meer uitleg over principe van een gesloten stroomkring) en Knikkers te koop (resultaten van metingen in een diagram verwerken) is dit 60%.
4. **Modellering:** Vergeleken met de andere opdrachten zien we dat bij Brug, Bibberspel en Knikkers te koop de oplossing het vaakst wordt gemodelleerd; bij Knikkerbaan en Fietsbel wordt het vaakst proceshulp gegeven en bij Knikkerbaan het vaakst leeshulp.

4.2 Soorten hulp naar leerlingkenmerken (onderzoeksvraag 2)

In het kader van de tweede onderzoeksvraag is nagegaan in hoeverre verschillen in verleende typen hulp enerzijds samenhangen met demografische kenmerken (leeftijd, geslacht, leerlinggewicht) en psychologische leerlingkenmerken van leerlingen (attitude ten aanzien van Natuur en Techniek, kennisniveau ten aanzien van Natuur en Techniek) anderzijds.

5. **Sekse en leerlinggewicht:** Er treden nauwelijks verschillen op in de frequentie van hulpcombinaties tussen jongens en meisjes en hulpfrequenties aan leerlingen met een verschillend leerlinggewicht. Uitzondering hierop vormt het Bibberspel waar jongens iets vaker hulp krijgen dan meisjes.
6. **Kennis van Natuur en Techniek:** Het kennisniveau van de leerlingen hangt samen met het al of niet verkrijgen van inhoudelijke hulp en modeling. Naarmate dit kennisniveau hoger is, wordt bij de praktische opdrachten minder inhoudelijke hulp verstrekt en worden oplossingen minder vaak gemodelleerd. Bij Fietsbel en de deelopdracht Conclusies trekken bij Knikkerbaan geldt dat een hoger kennisniveau samenhangt met minder proceshulp. Er is voorts geen samenhang tussen het kennisniveau en de mate van ontvangen leeshulp.
7. **Attitude en vrijetijdsbesteding:** Er is geen consistent verband gevonden tussen de attitude van leerlingen ten aanzien van N&T en vrije tijdsbesteding aan N&T enerzijds en de soort verkregen hulp anderzijds.
8. **Beleving van taakmoeilijkheid Natuur en Techniek:** Hoewel het verband niet sterk is, blijkt dat naarmate leerlingen Natuur en Techniek opdrachten als moeilijker beschouwen, ze meer leeshulp ontvingen. Bij de deelopdracht conclusies trekken Knikkerbaan blijkt tevens dat een hoge moeilijkheidsbeleving gepaard gaat met meer proceshulp.
9. **Toekomstplannen met Natuur en Techniek:** Naarmate leerlingen in sterke mate aangeven in de toekomst iets te willen doen met Natuur en Techniek, blijken ze minder vaak

verschillende vormen van hulp te ontvangen (leeshulp bij Fietsbel; proceshulp en modeling bij Klarzetten Knikkerbaan, inhoudelijke hulp bij Bibberspel en Knikkers te koop.

4.3 Leerlingkenmerken en prestaties op de praktische opdrachten in de conditie zonder hulp (onderzoeksvraag 3. deelvraag i)

Ter voorbereiding op de beantwoording van de vraag naar de effecten van hulp op de prestaties op praktische opdrachten (onderzoeksvraag 3. deelvraag i) is nagegaan op welke manier de scores voorafgaand aan hulp samenhangen met leerlingkenmerken. De volgende conclusies kunnen worden getrokken.

10. **Geslacht:** Jongens scoren zonder hulp wat beter dan meisjes op de opdrachten Fietsbel en Brug. Andersom scoren meisjes zonder hulp beter op de opdracht Bibberspel.
11. **Sociaal-economische achtergrond:** de sociaal-economische achtergrond van leerlingen, uitgedrukt met behulp van het leerlinggewicht, maakt niet uit voor de score zonder hulp op de opdrachten.
12. **Toekomstplannen met Natuur en Techniek:** naarmate leerlingen meer toekomstplannen hebben ten aanzien van Natuur en Techniek (Fietsbel, Bibberspel, Knikkerbaan) scoren ze hoger vóórdat hulp wordt gegeven.
13. **Beleving van taakmoeilijkheid Natuur en Techniek:** naarmate leerlingen Natuur en Techniek opdrachten als moeilijker beschouwen boeken ze een lagere score voor Fietsbel in de conditie zonder hulp.
14. **Kennis van Natuur en Techniek:** Naarmate het kennisniveau van leerlingen hoger is, liggen de scores zonder hulp op de opdrachten Fietsbel, Knikkerbaan, en Knikkers te koop hoger. Bij de andere opdrachten bestaat dit verband niet.
15. **Zelfbeeld ontwerp- en onderzoeksvaardigheden:** Naarmate leerlingen zonder hulp een hogere score halen op de praktische opdrachten, rapporteren ze een positiever beeld van de eigen ontwerp- en onderzoeksvaardigheden. Bij de doe-opdrachten Brug en Fietsbel is het verband beduidend sterker dan bij de andere opdrachten. Kennelijk is de zelfinschatting accurater op deze opdrachten dan op de overige opdrachten.

4.4 Overall effecten van hulp (onderzoeksvraag 3. deelvraag ii)

In het kader van de derde onderzoeksvraag (deelvraag ii) is nagegaan in hoeverre gegeven hulp gepaard gaat met prestatieverbetering. Bovendien is gekeken naar verschillen in effecten naar het type verleende hulp of de combinatie van hulptypen.

16. **Omvang effecten:** Bij alle opdrachten, op enkele prestatie-indicatoren na, worden zeer hoge effectgroottes van hulp gevonden, vaak ver boven de waarde .80. die wordt beschouwd als 'groot'.
17. **Verdeling van prestaties:** Behalve een positief effect op de gemiddelde scores blijkt dat een groot deel van de leerlingen een hogere score behaalt na inhoudelijke hulp.
18. **Effect per prestatie-indicator:** De effecten van hulp variëren per prestatie-indicator. Er zijn aanwijzingen dat de grootte van het hulpeffect samenhangt met de mate waarin de hint een strategie of stap naar een oplossing modelleert.
19. **Deelvaardigheden bij experimenteren:** Bij de opdracht Knikkerbaan worden grote effecten gevonden, maar verschillend voor de deelopdrachten. Zeer grote effecten worden gevonden op klarzetten van de opstelling en op het verrichten van metingen. Grote effecten worden gevonden op het trekken van juiste conclusies op grond van de resultaten.

20. **Meest effectieve interventie:** Het meest effectief zijn de hulpinterventies waarbij een voorbeeld wordt gegeven van de oplossing uit een andere context (zie samenvattende tabellen 16 en 17 over de effectiviteit van specifieke hulpinterventies). Herhaling van een prestatie-eis is zeer effectief, maar strikt genomen geen inhoudelijke hulp, maar proceshulp.
21. **Kwaliteit voorbeeld als hulpinterventie:** Daar waar in een gegeven voorbeeldoplossing uit een andere context specifieke informatie over een onderdeel van de oplossing ontbreekt, zien we dat op dit onderdeel niet beter wordt gescoord (Opdracht Knikkers te koop: invullen metingen in lijndiagram).
22. **Experiment 1 Knikkerbaan als voorbeeld voor een experiment 2:** Het blijkt dat (al dan niet ondersteunde) deelname aan experiment 1 ervoor zorgt dat de verschillen in prestaties op experiment 2 sterk verminderen. De algehele prestatieverbetering bij het uitvoeren van metingen tijdens experiment 2 wijst op een (nabij) transfereffect van het eerste experiment.

4.5 Differentiële effecten van hulp (Onderzoeksvraag 3. deelvraag iii)

Bij de vierde onderzoeksvraag (deelvraag iii) is gekeken of door middel van hulp de verschillen tussen leerlingen zouden afnemen. Bovendien was een meer specifieke vraag of leerlingen die zonder hulp relatief zwak scoorden na hulp ook in staat bleken de maximale score voor de opdrachten te behalen. De volgende conclusies kunnen worden getrokken.

23. **Verschillen in prestatieverbetering tussen leerlingen:** Bij drie van de vijf opdrachten nemen de scoreverschillen tussen leerlingen na hulp sterk af na hulp: Fietsbel in elkaar zetten, Bibberspel (gesloten stroomcircuit maken), en Knikkerbaan (deelopdrachten Klaarzetten opstelling en Verrichten metingen). Tussen 40 en 60% van de leerlingen behaalt na hulp de maximale score. Bij drie andere opdrachten (vouwen Brug en invullen van een diagram bij Knikkers te koop) blijven de aanvankelijke scoreverschillen tussen leerlingen grotendeels bestaan.

4.6 Effectiviteit van specifieke typen hulp en combinaties van hulp (Onderzoeksvraag 3. deelvraag iv)

De resultaten van het onderzoek geven aanleiding tot de volgende conclusies over de effectiviteit van verschillende typen en combinaties van hulp.

24. **Effect aantal inhoudelijke hulpinterventies:** Bij opdrachten waar meer dan één inhoudelijke hulpinterventie kon worden gegeven (Brug, Bibberspel, Knikkers te koop, en Knikkerbaan) blijkt dat naarmate het aantal inhoudelijke hints groter is, de prestatieverbetering groter is op de totale opdracht.
25. **Effect hulpcombinaties:** Leerlingen die behalve inhoudelijke hulp ook proceshulp ontvingen, bereiken een grotere prestatieverbetering dan leerlingen die uitsluitend inhoudelijke hulp ontvingen.
26. **Effect leeshulp:** Zoals beoogd, bereiken leerlingen die naast inhoudelijke hulp ook leeshulp ontvingen geen grotere prestatieverbetering dan leerlingen die alleen inhoudelijk hulp ontvingen. Kennelijk blijft de intrinsieke taaklast constant, terwijl de opdracht wel toegankelijker wordt gemaakt.
27. **Effect bij leerlingen die achterblijven:** Bij bijna alle deelopdrachten geldt dat leerlingen aan wie de oplossing moet worden voorgedaan (gemodelleerd) al ver achterlagen op leerlingen die wel hulp nodig hadden, en die geen modeling nodig hadden. Zowel de algehele prestatie als de prestatieverbetering waren na hulp over het algemeen laag. Specifiek ten aanzien van

de control-of-variables-strategie blijft de prestatie op de juiste instelling van alle onafhankelijke variabelen relatief ver achter. Het leerpotentieel van deze leerlingen is mogelijk lager op deze gebieden.

4.7 Samenhang tussen leerlingkenmerken en prestatieverbetering (onderzoeksvraag 4)

Bij de vierde onderzoeksvraag stond het verband tussen demografische kenmerken en psychologische kenmerken van leerlingen enerzijds en de bereikte prestatieverbetering na hulp anderzijds centraal. De volgende conclusies kunnen worden getrokken.

28. **Effect leeftijd en geslacht:** Bij de meeste opdrachten en bij de meeste prestatiegroepen (op grond van de scores voorafgaand aan hulp) is er geen verband tussen geslacht en leeftijd enerzijds en de mate van bereikte prestatieverbetering anderzijds. Jongens en meisjes en leerlingen van verschillende leeftijden profiteren in gelijke mate van de gegeven hulp.
29. **Effect attitude:** Daar waar wel samenhang is en attitude-elementen enerzijds en prestatieverbetering na hulp anderzijds, verschilt deze per opdracht en het scoreniveau voordat hulp was ontvangen. Het beeld is lastig te interpreteren.
30. **Relatie prestatieverbetering en zelfbeeld:** Naarmate leerlingen na hulp een grotere prestatieverbetering bereiken, rapporteren ze een positiever beeld over de eigen ontwerp- en onderzoeksvaardigheden die nodig zijn bij het uitvoeren van deze opdrachten.

Tabel 16: Samenvatting effecten van specifieke hulpinterventies op prestatieverbetering voor vier praktische opdrachten

Opdracht en prestatie-element	Specifieke hulpinterventie	Score na hulp	Effect-grootte	Aanduiding
Brug		0.62	0.66	Middelgroot
1.1. De pijlers zijn goed opgesteld	Model met ander materiaal: Foto van brug met pijlers	0.91	4.53	Zeer groot
1.2. Er is één vel papier gebruikt	Uitleg: je mag maar één vel papier gebruiken (Model)	0.74	2.39	Zeer groot
1.3. Het papier is in profiel gevouwen	Probeer nog een steviger profiel te vouwen (proceshulp)	0.33	0.99	Groot
1.4. Er is een boogconstructie bedacht	Geen hint	0.09	0.44	Klein
1.5. Het bootje kan overal onder de brug door	Alleen als eis vermeld (proceshulp)	0.41	1.18	Groot
1.6. De brug draagt het lichte autootje	Alleen als eis vermeld (proceshulp)	0.33	1.00	Groot
1.7. De brug draagt het zware autootje	Alleen als eis vermeld (proceshulp)	0.32	0.97	Groot
1.8. De brug draagt beide autootjes	Alleen als eis vermeld (proceshulp)	0.27	0.86	groot
Fietsbel		0.82	1.98	Zeer groot
2.1. De hendel om de bel te laten rinkelen en het kleine radertje zijn op de juiste manier aangebracht	Eén hint: een voorbeeldbel van kleiner formaat	0.82	3.03	Zeer groot
2.2. Het balkje met de metalen ringetjes is op de juiste manier aangebracht		0.83	3.09	Zeer groot
2.3. Het boutje om het balkje vast te zetten is op de juiste manier aangebracht		0.70	2.18	Zeer groot
2.4. De dop is er op gedraaid; de bel rinkelt		0.80	2.80	Zeer groot
Bibberspel		0.73	1.27	Groot
3.1. lampje goed ingedraaid	1. Uitleg principe: de elektrische stroom moet rond gaan om het lampje te laten branden.	0.58	1.68	Zeer groot
3.2. De draden zijn goed aangesloten	2. Uitleg strategie trouble shooting: Volg met je vinger de stroom vanaf de batterij en ga zo na waar het probleem kan ontstaat 3. Half afgemaakt voorbeeld: Ik sluit nu één draad goed aan. lukt het je nu?	0.57	1.63	Zeer groot
Knikkers te koop		0.50	1.40	Zeer groot
4.1. meetpunt kleine knikker op de juiste plaats in het diagram gezet	Eén interventiehint: een voorbeelddiagram uit een andere context (lengte en schoenmaat). zónder trendlijn	0.64	1.89	Zeer groot
4.2. meetpunt middelgrote knikker op de juiste plaats in het diagram gezet		0.65	1.91	Zeer groot
4.3. meetpunt grote knikker op de juiste plaats in het diagram gezet		0.64	1.87	Zeer groot
4.4. Trendlijn getrokken tussen de punten		0.09	.44	Klein
4.5. label BELLA op juiste plaats in diagram		0.48	1.36	Zeer groot
4.6. label ZORRO op juiste plaats in diagram		0.46	1.32	Zeer groot
4.7. label KINSI op juiste plaats het diagram		0.47	1.33	Zeer groot

Tabel 17: Samenvatting effecten van specifieke hulpinterventies op prestatieverbetering voor de opdracht Knikkerbaan

Deelopdracht en prestatie-element Knikkerbaan	Specifieke hulpinterventie	Score na hulp	Effect-grootte	Aanduiding
Klaarzetten experimentele opstelling	Totaalscore	0.76	1.13	Groot
5.1 Helling gelijk ingesteld	Bij (foutieve) verschillende instelling: "Weet je nu of het alleen aan de startpositie ligt of een bal verder komt? Of kan het ook aan de andere instellingen liggen?"	0.74	2.36	Zeergroot
5.2 Oppervlak gelijk ingesteld	Zie 5.1	0.50	1.41	Zeergroot
5.3 Startpositie verschillend ingesteld	Als leerlingen de startpositie (foutief) gelijk instelden, volgde de vraag: "als je wilt weten of de startpositie effect heeft, waar moet je de knikkers dan laten starten?"	0.55	1.57	Zeergroot
5.4 Gebruik knikker met gelijk gewicht	Zie 5.1	0.49	1.38	Zeergroot
Uitvoeren metingen volgens vraagstellingen van een experiment	Experiment 1 totaalscore	0.86	1.42	Zeergroot
	Experiment 2 totaalscore	0.79	1.03	Groot
5.5. Vijf metingen op beide banen verricht	Hier vul je in bij welk cijfer op de knikkerbaan de knikker stil ligt.	0.50	1.38	Zeergroot
Idem experiment 2	Idem + experiment 1 is model voor experiment 2	0.75	2.45	Zeergroot
5.6. Bij elke meting zijn twee knikkers van gelijk gewicht gebruikt	Is het eerlijk als je verschillende knikkers gebruikt?	0.56	1.56	Zeergroot
Idem experiment 2	Idem + experiment 1 is model voor experiment 2	0.38	1.10	Groot
5.7. In de tabel hebben de koppen een label (hoog of laag)	Wat zou je achter 'startpositie' moeten invullen?	0.70	2.14	Zeergroot
Idem experiment 2	Idem + experiment 1 is model voor experiment 2	0.56	1.60	Zeergroot
Passende conclusies trekken op grond van de verrichte metingen				
5.8 De conclusie is in overeenstemming met de gegevens uit de metingen)	Wijzen op een patroon in de data. "Kijk eens naar wat je hebt opgeschreven. Maakt het wat uit of de startpositie hoog of laag is?"	0.31	0.94	Groot
Idem experiment 2	Idem + experiment 1 is model voor experiment 2	0.30	0.93	Groot

5 Discussie en aanbevelingen

We sluiten dit rapport af met een terugblik op de resultaten, enkele aanbevelingen inzake de verdere ontwikkeling van toetsen met hulp, en een schets van perspectieven op verdere theorievorming. Tot slot noemen we enkele beperkingen van de studie en de gevolgen daarvan voor vervolgonderzoek.

5.1 Een terugblik op de resultaten

In deze studie is voor het eerst in het kader van de periodieke peiling N&T gebruik gemaakt van toetsen met hulp. Er is nog een lange weg te gaan in de ontwikkeling van dit type toetsen, maar deze studie heeft een aantal eerste inzichten opgeleverd over de hulp die feitelijk wordt gegeven en de effecten daarvan op prestatieverbetering op praktische opdrachten voor Natuur en Techniek.

5.1.1 Aan leerlingen verstrekte hulp

Een eerste algemene observatie was dat de frequentie van inhoudelijke hulp sterk verschilt naar de moeilijkheid van de opdracht (Brug, Bibberspel, Knikkers). Dit is geen verassend resultaat, maar het geeft wel aan dat niet of nauwelijks inhoudelijke hulp is gegeven, waar dat niet nodig was. Een tweede observatie is dat basiskennis van Natuur en Techniek een voorspellende waarde heeft voor de frequentie van hulp bij de praktische opdrachten. Naarmate dit kennisniveau hoger is, is bij de praktische opdrachten minder inhoudelijke hulp nodig en worden oplossingen minder vaak gemodelleerd. Het verband is echter niet sterk, wat inhoudt dat er diverse andere factoren meespelen in het voorspellen van ontwerp- en onderzoekprestaties. Een van die andere factoren betreft de attitude van leerlingen ten aanzien van N&T. Toekomstplannen die leerlingen hebben met N&T blijkt positief samen te hangen met de prestaties op sommige opdrachten (Fietsbel, Bibberspel, Knikkerbaan). Echter, ook dit verband is niet heel sterk, wat betekent dat er nog diverse ongespecificeerde factoren een rol spelen bij het vragen en geven van hulp.

5.1.2 Effecten van hulp op prestatieverbetering

Effecten van inhoudelijke hulp

Het effect van inhoudelijke hulp op prestatieverbetering blijkt groot te zijn, maar verschilt per opdracht. Ook zijn er differentiële effecten: leerlingen met een lage beginscore profiteren soms sterk, maar niet op alle opdrachten.

Veder zijn er aanwijzingen dat de grootte van het hulpeffect samenhangt met de aard van de gegeven aanwijzing. Aanwijzingen die een strategie of stap naar een oplossing modelleren, hebben een groter effect dan het noemen van een onderliggend principe bij een opdracht, en die aanwijzing is weer effectiever dan het herhalen van een prestatie-eis. De laatstgenoemde aanwijzing valt achteraf eerder onder proceshulp dan inhoudelijke hulp te rangschikken. Bij de opdracht Knikkerbaan bleek dat het doorlopen hebben van een experimenteeropdracht in feite werkte als een model voor het vervolgexperiment met een (iets) gewijzigde onderzoeksvraag, die verschilde op de te manipuleren onafhankelijke variabele.

Effecten van proceshulp in combinatie met inhoudelijke hulp

Leerlingen die behalve inhoudelijke hulp ook proceshulp ontvingen, bereikten een grotere prestatieverbetering dan leerlingen die uitsluitend inhoudelijke hulp ontvingen. Proceshulp ondersteunt en stimuleert vermoedelijk de regulatievaardigheden van de leerlingen die nodig zijn bij het uitvoeren van de opdracht. We vonden grote effecten op prestatieverbetering van proceshulp naast inhoudelijke hulp bij taken waarbij leerlingen aan verschillende taakelementen tegelijkertijd aandacht moesten besteden. Hiervan was sprake bij drie opdrachten: 1) het invullen van metingen in een diagram (rekening houden met assen, labels, trendlijn), 2) het aanleggen van een gesloten stroomcircuit (verschillende problemen kunnen het oplichten van het lampje bij het Bibberspel

verhinderen), en 3) het klaarzetten van een experimentele opstelling (naast de aanpassing in de onafhankelijke variabele, moet de leerling de andere instellingen constant houden).

Leeshulp: geen effect

Als onderdeel van de hulp stond het de toetsleiders vrij om leerlingen bij alle deelopdrachten leeshulp te bieden. Deze hulpvorm beoogde de leerling toegang tot de opgave te verschaffen, als dat niet zo was door lees- of informatieverwerkingsproblemen. De resultaten wijzen uit dat de leeshulp, in tegenstelling tot inhoudelijke hulp niet gepaard gaat met prestatieverbetering. De studie was echter niet zo opgezet dat we konden nagaan wat het effect van het weglaten van leeshulp zou zijn geweest op het verdere verloop van de opdracht. Wanneer een leerling namelijk om welke reden dan ook niet verder kwam met de opdracht, werd altijd overgeschakeld op inhoudelijke hulp, die mogelijk in combinatie met leeshulp werd gegeven.

Verschillen naar de aard van de opdrachten en de doelvaardigheden

De aard van de opdracht speelde een duidelijke rol bij verschillen in prestaties en prestatieverbetering. Zo lijkt de Brug een op zichzelf staande opdracht. Bij deze opdracht, waarbij leerlingen een stevige (boogvormige) constructie van papier moesten vouwen, werd veruit het vaakst hulp geboden. Verschillen in kennis van Natuur en Techniek konden de prestatieverschillen tussen leerlingen niet verklaren.

De control-of-variables strategie bleek een vaardigheid te zijn waarop verschillen tussen leerlingen met hulp niet voor iedereen waren te overbruggen. Bij het klaarzetten van een experimentele opstelling had hulp weliswaar een positief effect op alle prestatie-elementen, maar het sterkst op het instellen van de onafhankelijke variabele, waarop een effect moet worden vastgesteld. Bij experiment 2 betrof dat de startpositie (hoogte) van waaruit de knikker werd losgelaten. Effecten op het juist instellen van de constant te houden variabelen (hellingshoek, oppervlak van de glijbaan, gewicht knikker) waren beduidend kleiner.

Een daarmee samenhangende bevinding was dat leerlingen die bij deze deelopdracht bijna niet van inhoudelijke hulp bleken te profiteren, voornamelijk op de juiste instelling van alle onafhankelijke variabelen relatief ver achterbleven. Waarschijnlijk is de vaardigheid control-of-variables bij een aantal leerlingen nog niet ontwikkeld en vraagt dit meer dan een éénmalige ondersteuning door een toetsleider.

5.2 Perspectieven op verdere ontwikkeling en theorievorming

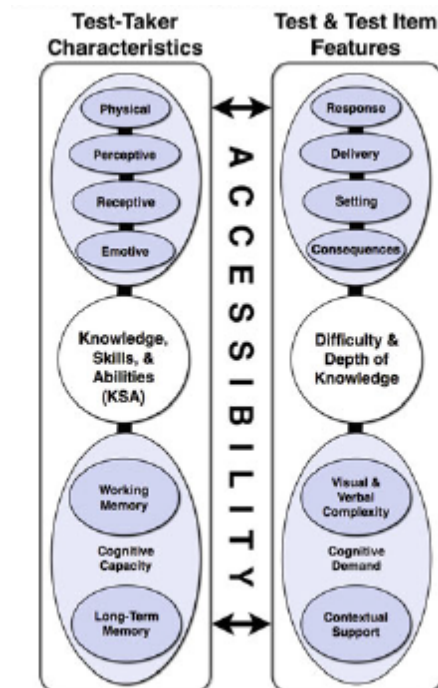
Voor het ontwerp van toetsen met hulp vraagt om verdere theorievorming, waarbij verschillende onderzoekslijnen worden gecombineerd.

Een eerste onderzoekslijn is die van dynamisch toetsen, waarbij het meten van leerpotentieel centraal staat. Het leerpotentieel wordt gezien als het verschil tussen waartoe een leerling zelfstandig in staat is en wanneer hij/zij hulp krijgt van een deskundige (volwassene). Hierbij voert een leerling toetstaken uit, begeleid door een getrainde toetsleider, die waar nodig hulp op maat biedt. In deze studie is het interventiemodel gehanteerd, waarbij alle hulpinterventies vooraf zijn gespecificeerd. Dit in tegenstelling tot het interactiemodel waarbij het onderscheid tussen instructies, leren en beoordeling vervaagt.

Om verder te kunnen komen in de ontwikkeling van toetsen met hulp is een verfijnder inzicht nodig in de aard van de hulpinterventies op inhoudelijk en procesniveau en de manier waarop die aangrijpen op de taakuitvoering door leerlingen, de leerstappen die leerlingen als gevolg daarvan zetten. Door meer kennis te vergaren over deze kenmerken, kan geleidelijk toegewerkt worden naar

de beschrijving van hulpconfiguraties. Voor een belangrijk deel zal hier gebruik gemaakt worden van onderzoek naar leren en instructie bij Natuur en Techniek.

Een tweede onderzoekslijn betreft toegankelijkheid van toetsopgaven. Toegankelijkheid is het resultaat van de interactie tussen leerlingkenmerken en toetskenmerken (zie Figuur 11).



Figuur 11: Toegankelijkheid als resultaat van de interactie tussen leerlingkenmerken en Test en opgave-kenmerken (figuurfragment ontleend aan Beddow, Kurz, & Frey, 2011)

Leerlingkenmerken betreffen met name toegangsvaardigheden (de paarse ovalen links in Figuur 11) en vaardigheid op het doeldomein (de witte cirkel links). Toets- en opgavekenmerken betreffen onder meer de moeilijkheid van de inhoud (de witte cirkel rechts) en ook de wijze waarop opgaven worden gepresenteerd (paarse ovalen rechts).

In de peilingsstudie is ervoor gekozen aan leerlingen die bij de toetsafname bepaalde toegangsvaardigheden blijken te missen, die op zich geen onderdeel vormen van de te beoordelen doelvaardigheid, hulp te bieden om hen toegang tot een opdracht te laten krijgen. In deze studie werd leeshulp gegeven, als de leerlingen zich geen goed beeld van de opdracht konden vormen, bij het verwerken van de opdrachtinformatie en de opdrachtstelling. Het toegankelijk maken van opgaven betekent dat onnodige taaklast (of: extrinsieke taaklast (Sweller, 2010), zoveel mogelijk wordt weggenomen, om zo uitsluitend de voor de taak essentiële taaklast in stand te houden. In vervolgactiviteiten blijft de vraag van belang welke taakelementen gerekend moeten worden tot de doelvaardigheid en welke tot toegangsvaardigheden. Daarbij dienen richtlijnen te worden ontwikkeld om te bepalen of bijvoorbeeld leesvaardigheid wel of niet tot de doelvaardigheid moet worden gerekend.

Een toetstaak kan niet toegankelijk zijn voor een leerling of zelf ongeschikt, als deze veel te complex is voor het bereikte ontwikkelingsniveau. Dit brengt ons op een derde onderzoekslijn. Deze betreft studies naar leerprogressies en ontwikkelingstrajecten, in dit geval op het terrein van ontwerp- en onderzoeksvaardigheden binnen Natuur en Techniek. Voor de ontwikkeling van toetsen met hulp zijn met name inzichten van belang over de voortschrijding van leerlingen in het uitvoeren van steeds complexere taken, die een steeds geavanceerder niveau van vaardigheid vereisen. Bij het uitwerken

van leerlingmodellen zouden uitgewerkte leerprogressies een nadrukkelijker rol moeten spelen (Pellegrino, Wilson, Koenig, & Beatty, 2014; Pellegrino, 2014). Een onderdeel daarvan is het definiëren van (typische) taken met een bij het ontwikkelingsniveau zogeheten passende intrinsieke taaklast. Volgens de cognitieve lasttheorie wordt de intrinsieke last van een taak bepaald door de mate waarin verschillende taakelementen met elkaar interacteren. Een taakelement is daarbij alles dat moet worden beheerst (of: geleerd) om een taak goed uit te voeren, zoals een concept, procedure, een principe, of een feit. Als een taak wordt ervaren als een veelvoud van met elkaar interacterende elementen, waaraan bewust aandacht moet worden besteed, dan is sprake van hoge intrinsieke last. Worden die elementen door de leerling als één samenhangend geheel gezien, bijvoorbeeld als gevolg van ontwikkelde schemata, dan is voor deze leerling op dit ontwikkelingsniveau de intrinsieke taaklast relatief laag. Hoe sterk taakelementen als losse interacterende eenheden worden ervaren, hangt af van het ontwikkelingsniveau (kennis- en vaardigheden) van de leerling. Ervaren onderzoekers herkennen een experimentele opstelling als één geheel, terwijl (sommige) leerlingen van groep 8 de benodigde elementen van een experimentele opstelling mogelijk nog als losse onderdelen ervaren. We zagen bijvoorbeeld in deze studie dat een deel van de leerlingen wel in staat was de onafhankelijke variabele (na hulp) bij de Knikkerbaan juist in te stellen, maar niet de variabelen die constant moesten worden gehouden. Bij de verdere ontwikkeling van toetsen met hulp is van belang dat voor de achtereenvolgende ontwikkelingsniveaus van leerlingen in voor de doelvaardigheden passende taakkenmerken en cognitieve last bepalende elementen worden uitgewerkt.

5.3 Aanbevelingen: ontwerpen van ontwikkelingsgerichte prestatiematen

In deze studie is exploratief in kaart gebracht welke effecten kunnen optreden bij verschillende hulpinterventies en combinaties daarvan, om te kunnen rapporteren waartoe leerlingen zelfstandig in staat zijn en wat ze met hulp van een deskundige kunnen.

Aansluitend op de boven geschetste theoretische en toetstechnische ontwikkelingen, is het zaak toe te werken naar ontwikkelingsgerichte prestatiematen. We denken daarbij aan maten die aangeven:

- a) hoe een leerling presteert op een doelvaardigheid bij taken met een voor een bepaald ontwikkelingsniveau vastgestelde complexiteitsconfiguratie, waarbij duidelijk wordt wat leerlingen zelfstandig kunnen in complexiteitsconfiguratie A, maar niet meer in een eerstvolgende complexiteitsconfiguratie B.
- b) hoever de afstand is tussen wat leerlingen kunnen met een bepaalde hoeveelheid hulp en met minder hulp, uitgewerkt in bijvoorbeeld hulpconfiguraties A en hulpconfiguratie B. Uiteindelijk is het aan het onderwijsveld om te bepalen onder welke hulpconfiguraties leerlingen per ontwikkelingsniveau dienen te presteren.
- c) hoe een leerling presteert op een doelvaardigheid bij verschillende gradaties van ondersteuning op *toegangsvaardigheden* (zoals ordelijk en aandachtig werken, leesvaardigheid). Dit zijn vaardigheden die wel nodig zijn om te presteren, maar die op zichzelf geen onderdeel zijn van de te meten doelvaardigheid (cf. Beddow, Kurz, & Frey, 2011). Daarbij zal bepaald moeten worden wat toegangsvaardigheden zijn, waarop altijd hulp kan worden gegeven, en wat de te meten doelvaardigheden zijn.

Voor individuele leerlingen die dreigen achter te blijven kunnen dit soort maten uitmonden in uitspraken over hun leerpotentieel. De vraag die daarbij beantwoord wordt is, welke leerprogressies een leerling kan doormaken bij verschillende configuraties van hulp, en of die hulp bekijft in nieuwe taakcontexten.

5.4 Methodologische beperkingen en aanbevelingen

Dieper onderzoek naar hulpinterventies

In de huidige opzet konden de effecten van alle verschillende hulpsoorten niet onafhankelijk van elkaar worden geschat. Om redenen van praktische uitvoerbaarheid is er niet voor gekozen om de toepassing van afzonderlijke hulpinterventies te registreren. Bovendien werd het exacte moment van leeshulp en proceshulp niet vastgelegd. Beide hulpvormen kunnen voorafgegaan zijn aan inhoudelijke hulp maar ook tegelijk of ná inhoudelijk hulp zijn toegepast. Om een meer fijnmazig beeld te krijgen van effecten, verdient het aanbeveling om alle hulpinterventies en het moment ervan te registreren, zoals ook gedaan is in studies van Swanson (2011) en Fuchs et al. (2011). Verder is het aan te bevelen vervolgonderzoek op te zetten waarin gewerkt wordt met experimentele condities: een groep die de opdrachten uitvoert zonder hulp en een groep die dat met hulp doet.

Verschillen tussen toetsleiders

Aanvullende analyses hebben aanzienlijke verschillen tussen klassen in prestaties op de opdrachten aangetoond. Op grond van de beschikbare data kon niet worden vastgesteld wat de oorzaak van deze verschillen is: de leerkracht of de toetsleider.

Er zijn aanwijzingen dat toetsleiders in de rol van tweede beoordelaar redelijk overeenstemmen in hun beoordeling van diverse prestatie-elementen met de oorspronkelijke toetsleiders. Over de praktijk van hulp is echter weinig bekend. Er zijn geen gegevens verzameld over mogelijke knelpunten bij het signaleren van hulpbehoeften, het bieden van hulp en de registratie van gegeven hulp. Het verdient dan ook aanbeveling om het hulpgedrag door toetsleiders systematischer en meer fijnmazig te onderzoeken: op welke wijze signaleren toetsleiders hulpbehoeften, in welke mate realiseren ze de beoogde hulpinterventie daadwerkelijk en verstrekken ze bijvoorbeeld meer of minder vergaande hulp dan nodig is.

Een andere denkbare richting voor vervolgonderzoek is het werken met toetsleiders die een verschillende achtergrond hebben, zoals ervaren leraren, die wellicht beter in staat zijn om hulpbehoeften te signaleren, deelprestaties te scoren, en daarop afgestemde hulp te verlenen.

6 Referenties

- Beddow, P.A., Kurz, A., & Frey, J.R. (2011). Accessibility Theory: guiding the science and practice of test item design with the test-taker in mind. In: S.N. Elliott, R. J. Kettler, P. A. Beddow, & A. Kurz (Eds). *Handbook of Accessible Achievement Tests for All Students Bridging the Gaps Between Research, Practice, and Policy*. New York: Springer (pp. 163-182).
- Bennett, R. E. (2010). Cognitively Based Assessment of, for, and as Learning (CBAL): A Preliminary Theory of Action for Summative and Formative Assessment. *Measurement: Interdisciplinary Research & Perspective*, 8(2-3), 70-91.
- Cito (2016). *Natuur en Techniek. Technisch rapport over resultaten Peil. onderwijs in 2015*. Arnhem: Cito.
- Feuerstein, R., Rand, Y. & Hoffman, M. B. (1979). *The Dynamic Assessment of Retarded Performers: The Learning Potential Assessment Device, Theory, Instruments and Techniques*. Baltimore: University Park Press
- Fischer, K. W., & Bidell, T. R. (2007). *Dynamic development of action and thought*. Handbook of Child Psychology, doi:10.1002/9780470147658.chpsy0107
- Fuchs, L.S., Compton, D.L., Fuchs, D., Hollenbeck, K.N., Hamlett, C.L., & Seethaler, P.M. (2011). Two-Stage Screening for Math Problem-Solving Difficulty Using Dynamic Assessment of Algebraic Learning. *Journal of Learning Disabilities*, 44(4) 372–380.

- Hansen, E. G., Mislevy, R. J., Steinberg, L. S., Lee, M. J., & Forer, D. C. (2005). Accessibility of tests within a validity framework. *System: An International Journal of Educational Technology and Applied Linguistics*, 33, 107-133.
- Jonassen, D. (2010). *Research Issues in Problem Solving*. Paper presented at the 11th International Conference on Education Research New Educational Paradigm for Learning and Instruction September 29 – October 1, 2010
- Keulen, H. van, & Slot, E. (2013). *Excellentiebevordering door middel van onderzoekend en ontwerpend leren: Vaardigheden Rubrics Onderzoeken en Ontwerpen (VROO)*. Den Haag: School aan zet.
- Kneepkens, B., Van der Schoot, F., & Hemker, B. (2011). *Balans van het natuurkunde- en techniekonderwijs aan het einde van de basisschool 4*. Arnhem: Cito.
- Langen, A. van, & Vierke, H. (2009). *Wat bepaalt de keuze van een natuurprofiel? De invloed van de leerling, de school, de ouders en de peergroup*. Den Haag: Platform Bèta Techniek.
- Lidz, C. S., & Elliott, J. G. (Eds) (2000) *Dynamic Assessment: Prevailing models and applications* (New York. Elsevier).
- Martin Schwichow, M., Croker. S., Zimmerman. C., Höffler, T., & Härtig. H, (2016). Teaching the control-of-variables strategy: A meta-analysis. *Developmental Review*. 39. 37–63.
- Pellegrino, J. W., Wilson, M., Koenig, J. & Beatty, A. (Eds.) (2014). *Developing assessments for the next generation science standards*. Washington, DC: National Academies Press.
- Pellegrino, J.W. (2014). Assessment as a positive influence on 21st century teaching and learning: A systems approach to progress. *Psicología Educativa*, 20, 2, 65-77.
- Sternberg, R. J., & Grigorenko, E. L. (2002) *Dynamic Testing*. New York: Cambridge University Press.
- Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, 22, 123–138.

7 Bijlagen

Bijlage 7.1 Ontvangen hulpvormen in percentages leerlingen

7.1.1 Ontvangen hulpvormen bij vier praktische opdrachten

	Brug (n=315)	Fietsbel (n=342)	Bibberspel (n=311)	Knikkers te koop (n=281)
Verkregen mate van hulp				
Geen hulp gehad	3.8	32.7	14.1	19.2
Wel hulp gehad	96.2	67.3	85.9	80.8
Na hulp oplossing gemodelleerd	42.4	13.4	33.1	25.3
Verdeling in soorten verkregen hulp*	(n=303)	(n=230)	(n=267)	(n=227)
1. Alleen leeshulp en/of proceshulp	3.0	15.2	7.9	6.6
Alleen leeshulp	.7	2.2	2.6	0.0
Alleen proceshulp	1.3	7.4	3.0	5.3
Leeshulp en proceshulp	1.0	5.7	2.2	1.3
2. Alleen inhoudelijke hulp	56.8	53.9	59.9	65.6
3. Combinatie inhoudelijke hulp en andere hulp	40.3	30.9	32.2	27.8
Leeshulp en inhoudelijke hulp	15.8	5.7	10.9	4.4
Proceshulp en inhoudelijke hulp	17.5	20.9	15.7	14.5
Leeshulp, proceshulp en inhoudelijke hulp	6.9	4.3	5.6	8.8

N.B. * Percentages berekend over hen die hulp hebben ontvangen; NB= Niet te berekenen; Interpretatie van waarden: .0 - .2: gering, .2 - .4: matig, .4 - .6: redelijk, >.6: voldoende tot perfect.

7.1.2 Ontvangen hulpvormen bij deelopdrachten van de Knikkerbaan

	Knikkerbaan experiment 1		Knikkerbaan experiment 2		
	Metingen verrichten (n=222)	Conclusies trekken (n=219)	Klaarzetten experiment (n=223)	Metingen verrichten (n=220)	Conclusies trekken (n=218)
Verkregen mate van hulp					
Geen hulp gehad	44.6	81.3	42.6	52.3	87.6
Wel hulp gehad	55.4	18.7	57.4	47.7	12.4
Na hulp oplossing gemodelleerd	8.1	5.4	14.3	8.2	3.2
Verdeling in soorten verkregen hulp*	(n=123)	(n=41)	(n=128)	(n=105)	(n=27)
1. Leeshulp en/of proceshulp	28.5	36.6	15.6	27.6	25.9
Alleen leeshulp	9.8	12.2	2.3	10.5	3.7
Alleen proceshulp	12.2	17.1	9.4	12.4	11.1
Leeshulp en proceshulp	6.5	7.3	3.9	4.8	11.1
2. Alleen inhoudelijke hulp	44.7	46.3	57.8	57.1	59.3
3. Combinatie inhoudelijke hulp en andere hulp	26.8	17.1	26.6	15.2	14.8
Leeshulp en inhoudelijke hulp	5.7	7.3	4.7	4.8	11.1
Proceshulp en inhoudelijke hulp	15.4	4.9	14.8	9.5	3.7
Leeshulp, proceshulp en inhoudelijke hulp	5.7	4.9	7.0	1.0	0.0

N.B. * Percentages berekend over hen die hulp hebben ontvangen; NB= Niet te berekenen; Interpretatie van waarden: .0 - .2: gering, .2 - .4: matig, .4 - .6: redelijk, >.6: voldoende tot perfect.

7.2 Prestaties voor en na hulp van p10, p50 en p90 leerlingen

7.2.1 Toelichting op de totstandkoming van p-waarden.

In deze bijlage zijn tabellen opgenomen met daarin de p-waarden van prestatie-indicatoren op de praktische opdrachten, die onderdeel waren van de voorliggende deelstudie (Effecten van hulp, hierna afgekort als EVH). In de hoofdtekst worden deze p-waarden niet besproken. De reden waarom ze in deze bijlage zijn opgenomen is dat soortgelijke tabellen zijn gepresenteerd in het technisch rapport (aangeduid als TN) van de peilingsstudie (Cito, 2016). De voorliggende EVH en de oorspronkelijke TN tabellen hebben echter een verschillende grondslag, die hieronder kort wordt toegelicht.

Overeenkomst

De groep leerlingen waarover is geanalyseerd, is exact gelijk. Ook het vertrekpunt, scores op prestatie-indicatoren (steeds beschouwd als 'items') is hetzelfde.

Verschillen in databehandeling

Verschil 1: De analyses hebben bij TN en EVH niet steeds betrekking op dezelfde deelopdrachten. Voor het doel van de deelstudie EVH zijn gericht deelopdrachten gekozen (zie paragraaf 2.1.1 van dit rapport), om het effect van hulp te onderzoeken. Hierdoor hebben somscores vaak betrekking op minder onderdelen en kunnen dan ook verschuivingen optreden in de rangschikking van leerlingen, en daarmee in p-waarden.

Verschil 2: De scoring van prestaties is anders. Ik maak twee scores: voor en na hulp en score overall 0-1. De p-waarden verschillen daardoor per definitie, ook die na hulp. TN geeft 0 in geval dat het antwoord met hulp incorrect is, 1 bij correct na hulp en 2 bij correct met hulp. Onderstaande tabel kruist beide scoringswijzen met elkaar. P-waarden komen verschillend tot stand (zie onder).

		De pijlers zijn goed opgesteld zonder hulp (scoring EVH)		Totale n
		0	1	
De pijlers zijn goed opgesteld (scoring TN)	0: fout na inhoudelijke hulp	4	0	4
	1: goed na inhoudelijke hulp	41	0	41
	2: goed zonder inhoudelijke hulp	0	270	270
Totale n		45	270	315

De P-waarde voor items werd bij TN bepaald op grond van een gecombineerde score met en zonder hulp.

n1 goed na hulp	n2 goed beantwoord zonder hulp	puntentotaal (n1*1+n2*2)	Max aantal punten (n1+n2)*2	p-waarde
41	270	581	622	0.93

Bij EVH werd de P-waarde berekend voor de scores onder de condities 'zonder hulp'

n goed beantwoord zonder hulp	n totaal	p-waarde zonder hulp
270	315	0.86

Op soortgelijke wijze is bij EVH de p-waarde voor de condities met hulp worden berekend.

		De pijlers zijn goed opgesteld met hulp (scoring EVH)		Total
		0	1	
AB1_02._1 1. De pijlers zijn goed opgesteld (scoring TN)	0 fout na inhoudelijke hulp	4	0	4
	1 goed na inhoudelijke hulp	0	41	41
	2 goed zonder inhoudelijke hulp	0	270	270
Total		4	311	315

Bij EVH werd de P-waarde berekend voor de scores onder de condities 'met hulp'

n goed beantwoord: situatie na hulp	n totaal	p-waarde in situatie na hulp
311	315	0.99

Verschil 3: De verdeling van somscores is bij mij gebaseerd op de situatie zonder hulp. Die verdeling is wat anders dan die waarbij hulp in de score is verdisconteerd. Er komen verschuivingen van leerlingen in voor, omdat hulp eruit is gehaald. Daardoor zitten er deels andere leerlingen in p10, p50, p90.

Verschil 4: De somscore is bij TN iets anders samengesteld. Ten eerste is de betrouwbaarheidsanalyse op de items (prestatie-indicatoren) verricht op de 0,1,2 scores. Daarbij valt bijvoorbeeld item 1 bij brug wel uit bij TN en niet in mijn analyse met 0-1 scores zonder hulp. Ten tweede heeft TN een gecorrigeerd gemiddelde voor alle itemscores gebruikt, waardoor de somscores er anders uitzien dan de ruwe somscores, die bij EVH worden gebruikt.

Dit leidt tot een verschillende scorering. TN Brug: 7 items, met somscores gebaseerd op gecorrigeerde itemscores versus EVH-Brug: 8 items, ruwe somscores. Zo komt een score 1.55 voor bij TN op een schaal van 0-14. Bij EVH lopen de scores van 0-8, zonder gebroken getallen.

7.2.2 P-waarden van prestatie-indicatoren op de opdracht Brug voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore

	Vóór inhoudelijke hulp			Na inhoudelijke hulp		
	p10	p50	p90	p10	p50	p90
Somscore opdracht Brug: een dragende brug maken van papier	.06	.44	.94	.32	.64	.94
1. De pijlers zijn goed opgesteld	.44	.90	1.00	.89	1.00	1.00
2. Er is één vel papier gebruikt	.04	.77	.99	.60	.99	1.00
3. Het papier is in profiel gevouwen	.00	.41	.99	.24	.59	1.00
4. Er is een boogconstructie bedacht	.00	.18	.64	.09	.29	.64
5. Het bootje kan overal onder de brug door	.02	.56	.99	.28	.81	.99
6. De brug draagt het lichte autootje	.00	.35	.97	.22	.55	.97
7. De brug draagt het zware autootje	.00	.20	1.00	.22	.49	1.00
8. De brug draagt beide autootjes	.00	.12	.93	.18	.40	.93

N.B. groepsgrootte resp. n=76, 86, 65

7.2.3 P-waarden van prestatie-indicatoren op de opdracht Fietsbel voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore

	Vóór inhoudelijke hulp			Na inhoudelijke hulp		
	p10	p50	p90	p10	p50	p90
Fietsbel: een fietsbel uit onderdelen in elkaar zetten	.00	.63	1.00	.75	.93	1.00
1. Het kleine radertje is op de juiste manier aangebracht	.00	.68	1.00	.80	.99	1.00
2. Het balkje met de metalen ringetjes is op de juiste manier aangebracht	.00	.85	1.00	.79	1.00	1.00
3. Het boutje om het balkje vast te zetten is op de juiste manier aangebracht	.00	.24	1.00	.71	.78	1.00
4. De dop is er op gedraaid; de bel rinkelt	.00	.74	1.00	.79	.96	1.00

N.B. groepsgroottes resp. n=113, 48, 147

7.2.4 P-waarden van prestatie-indicatoren op de opdracht Bibberspel voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore

	Vóór inhoudelijke hulp			Na inhoudelijke hulp		
	p10	p50	p90	p10	p50	p90
Bibberspel: een gesloten stroomcircuit aanleggen, zodat spel functioneert	.00	.50	1.00	.57	.78	1.00
1. Lampje goed ingedraaid	.00	.99	1.00	.60	.99	1.00
2. De draden zijn goed aangesloten	.00	.01	1.00	.60	.56	1.00

N.B. groepsgroottes resp. n=58,188, 65

7.2.5 P-waarden van prestatie-indicatoren op de opdracht Knikkers te koop voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore

	Vóór inhoudelijke hulp			Na inhoudelijke hulp		
	p10	p50	p90	p10	p50	p90
Knikkers te koop: metingen in een diagram verwerken	.06	.44	.94	.32	.64	.94
1. Meetpunt kleine knikker op de juiste plaats in het diagram gezet	.44	.90	1.00	.89	1.00	1.00
2. Meetpunt middelgrote knikker op de juiste plaats in het diagram gezet	.04	.77	.99	.60	.99	1.00
3. Meetpunt grote knikker op de juiste plaats in het diagram gezet	.00	.41	.99	.24	.59	1.00
4. Er is een trendlijn getrokken tussen de punten	.00	.18	.64	.09	.29	.64
5. Label BELLA op de juiste plaats in het diagram gezet	.02	.56	.99	.28	.81	.99
6. Label ZORRO op de juiste plaats in het diagram gezet	.00	.35	.97	.22	.55	.97
7. Label KINSI op de juiste plaats in het diagram gezet	.00	.20	1.00	.22	.49	1.00

N.B. groepsgrootte resp. n=133, 15, 60

7.2.6 P-waarden van prestatie-indicatoren op de opdracht Knikkerbaan voor de groepen leerlingen met percentielscores p10, p50 en p90 op de totaalscore

	Vóór inhoudelijke hulp Experiment 1			Na inhoudelijke hulp Experiment 1			Vóór inhoudelijke hulp Experiment 2			Na inhoudelijke hulp Experiment 2		
	p10*	p50	p90	p10	p50	p90	p10	p50	p90	p10	p50	p90
Opdracht Knikkerbaan (totaalscore)	.29	.63	.94	.37	.66	.96	.67	.80	.86	.81	.90	.93
Opstellen Knikkerbaan (totaalscore)	.05	.33	.96	NB	NB	NB	.58	.73	.88	.78	.88	.96
5.1 Helling gelijk ingesteld	.02	.27	.94	NB	NB	NB	.48	.75	.82	.78	1.00	.94
5.2 Oppervlak gelijk ingesteld	.02	.20	1.00	NB	NB	NB	.62	.70	.93	.79	.84	.97
5.3 Startpositie verschillend ingesteld	.16	.62	.94	NB	NB	NB	.56	.64	.83	.79	.80	.94
5.4 Gebruik knikker met gelijk gewicht	.00	.24	.94	NB	NB	NB	.66	.82	.94	.81	.86	.97
Uitvoeren metingen volgens vraagstellingen van een experiment (totaalscore)	.55	.90	.92	.79	.99	.98	.76	.87	.84	.87	.95	.93
5.5 Vijf metingen op beide banen	.68	.96	.97	.80	1.00	1.00	.72	.76	.72	.94	.96	.87
5.6 Bij elke meting zijn twee knikkers van gelijk gewicht gebruikt	.76	.98	1.00	.85	1.00	1.00	.89	1.00	1.00	.92	1.00	1.00
5.7 In de tabel hebben de koppen een label (hoog of laag)	.22	.76	.78	.73	.96	.94	.73	.89	.85	.84	.93	.96
De conclusie is in overeenstemming met de gegevens uit de metingen	.55	.98	.94	.69	.98	.97	.84	.98	.85	.86	1.00	.91

N.B. * groepsgroottes resp. n= 46, 45, 53; NB= Niet beschikbaar, omdat bij deze deelopdracht geen hulp is gegeven.