

EFFECTIVITEIT VAN DE STIP-AANPAK

EEN ONDERZOEK NAAR KENNISUITWISSELING EN KENNISONTWIKKELING
BIJ LEERLINGEN VAN VERSCHILLENDE COMPETENTIENIVEAUS



Auteur: Elise Eshuis

Begeleiders: Dr. T.H.S. Eysink
M. Hulsbeek, MSc.

Faculteit: Behavioural, Management and Social Sciences

Master: Psychologie

Specialisatie: Instructie, Leren & Ontwikkeling

Datum: 21 december 2015

UNIVERSITEIT TWENTE.

Voorwoord

Deze these is het resultaat van het onderzoek dat ik heb uitgevoerd voor het STIP-project en dient als afsluiting van mijn masteropleiding Psychologie, richting Instructie, Leren en Ontwikkeling aan de Universiteit Twente. Ik kijk terug op een erg fijne en leerzame tijd, waarin ik met veel plezier aan mijn onderzoek heb gewerkt en de kans heb gekregen om mezelf te ontwikkelen. Dit was niet het geval geweest zonder de fijne begeleiding ik heb gehad, waarvoor ik Tessa Eysink en Manon Hulsbeek graag wil bedanken. De vele gesprekken en waardevolle adviezen en feedback die ze hebben gegeven, maakte dat ik steeds weer met een goed gevoel en nieuwe inzichten aan de slag kon. Daarnaast wil ik Hannie Gijlers graag bedanken voor haar hulp en adviezen met betrekking tot het analyseren van de data. Verder gaat mijn dank uit naar alle leerkrachten die hebben meegewerkt aan dit onderzoek en die mij het vertrouwen hebben gegeven om hun lessen te observeren. Ten slotte wil ik graag mijn familie, in het bijzonder mijn ouders, en vriendinnen bedanken voor hun betrokkenheid en steun, in welke vorm dan ook.

Elise Eshuis

Rijssen, December 2015

Samenvatting

In deze studie is onderzoek gedaan naar de effectiviteit van de STIP-aanpak (Samenwerken tijdens Taak-, Inhoud- en Procesdifferentiatie). Volgens deze aanpak doen leerlingen eerst op niveau kennis op door samen te werken in homogene groepen, waarna ze hun unieke kennis uitwisselen in heterogene groepen om daarmee een gezamenlijke eindopdracht te kunnen maken. Het doel van dit onderzoek was om uit te zoeken wat het effect is van deze aanpak op de kennisontwikkeling van leerlingen van verschillende competentieniveaus en hoe de mogelijke verschillen in inbreng tijdens de kennisuitwisseling in de heterogene groepen samenhangen met de individuele kennisontwikkeling. Daarnaast is onderzocht of er een verband bestaat tussen de kwaliteit van de opbrengsten van de homogene samenwerking, de inbreng van de individuele leerling tijdens de heterogene samenwerking, de kwaliteit van de opbrengsten van de heterogene samenwerking en de individuele kennisontwikkeling. Om dit te onderzoeken is gebruik gemaakt van twee condities: een experimentele conditie waarin leerlingen een leerinhoud aangeboden kregen op basis van de STIP-aanpak ($N = 95$) en een controle conditie waarin leerlingen dezelfde leerinhoud op reguliere wijze aangeboden kregen ($N = 149$). Uit de resultaten bleek dat leerlingen die lessen op basis van de STIP-aanpak volgden minder kennis ontwikkelden dan leerlingen die reguliere lessen volgden. Een verklaring hiervoor kan zijn dat, ondanks dat in beide condities de totale lestijd vergelijkbaar was, de tijd die leerlingen hebben gehad om kennis tot zich te nemen in de experimentele conditie korter was dan in de controle conditie. Daarnaast is een positieve samenhang gevonden tussen de hoeveelheid groepsinbreng tijdens de heterogene samenwerking en de individuele kennisontwikkeling. Ook werd gevonden dat hoe beter de opbrengst van de homogene samenwerking was, hoe groter de inbreng van de leerling, hoe beter de opbrengst van de heterogene samenwerking en hoe groter de kennisontwikkeling was. Concluderend kan worden gesteld dat de STIP-aanpak voorziet in een methodiek waarbij leerlingen op niveau uitgedaagd worden en waarbij, mits er goed samengewerkt wordt, de potentie aanwezig is om de kennisontwikkeling bij leerlingen verder te doen groeien.

Trefwoorden: samenwerkend leren; differentiatie; kennisuitwisseling; homogeen groeperen; heterogeen groeperen; kennisontwikkeling

Abstract

This study examined the effectiveness of the STIP approach (in Dutch: “Samenwerken tijdens Taak-, Inhoud- en Procesdifferentiatie”). According to this approach, learners construct knowledge by working together in homogeneous groups, in which instruction is adjusted to their ability level. After that, they have to exchange their unique knowledge in heterogeneous groups in order to accomplish a group assignment. The aim of this study was to examine the effect of this approach on the learning gains of learners with different ability levels and to determine how the different contributions of the heterogeneous groups relate to the individual learning gains. Besides that, it was examined how the results of the homogeneous cooperation, the contribution of the individual learner, the results of the heterogeneous cooperation and the individual learning gains relate to each other. In order to test this, there were two conditions: an experimental condition in which lessons were given according to the STIP approach ($N = 95$) and a control condition in which lessons were given in a more traditional way ($N = 149$). The results showed that learners in the control condition reached higher learning gains than learners in the experimental condition. Despite the fact that the total amount of time for the lessons was the same in both conditions, learners in the control condition had more time to construct knowledge than learners in the experimental condition, which could be an explanation for this result. The other results showed that there was a positive relationship between the amount of group contributions and the individual learning gains. Furthermore, it was found that the better the results of the homogeneous cooperation, the more learners contribute during knowledge exchange, the better the results of the heterogeneous cooperation and the higher the individual learning gains. To conclude, on condition that learners cooperate well, the STIP approach offers a useful methodology that has the capacity for learners to reach higher learning gains.

Keywords: cooperative learning; differentiation; knowledge exchange; homogeneous grouping; heterogeneous grouping; knowledge construction

Inhoudsopgave

Voorwoord	2
Samenvatting	3
Abstract.....	4
Inleiding.....	6
Samenwerkend leren	6
Groepssamenstelling	7
Jigsaw	9
Differentiëren met jigsaw	9
Huidig onderzoek	10
Methode.....	12
Design.....	12
Deelnemers.....	12
Domein	13
Materialen.....	13
Procedure.....	16
Data-analyse	18
Resultaten.....	22
Kennisonwikkeling	22
Kennisuitwisseling	24
Samenhang	29
Discussie en Conclusie.....	31
Literatuur.....	35

Inleiding

Basisscholen hebben, mede door de recente invoering van passend onderwijs (Van Bijsterveldt-Vliegenthart, 2011), te maken met grote cognitieve verschillen tussen leerlingen in één klas. Traditionele vormen van onderwijs, waarbij klassikale instructie wordt gegeven, bieden vaak niet de ruimte om leerlingen uit te dagen op niveau (Tomlinson et al., 2003), terwijl blijkt dat het toepassen van differentiatie veelal leidt tot betere leerprestaties van leerlingen (Reis & Renzulli, 2010). Daarnaast blijken veel docenten moeite te hebben om het onderwijs af te stemmen op het niveau van de leerlingen, vooral wanneer het lesmateriaal hier geen ondersteuning in biedt (Inspectie van het Onderwijs, 2015). Een manier om leerlingen op hun eigen niveau uit te dagen is om ze te groeperen op competentieniveau en ze in deze groepen samen te laten werken aan een taak (Cohen, 1994; Lou, Abrami, Spence, Poulsen, Chambers & d'Apollonia, 1996).

Samenwerkend leren

Uit meerdere studies blijkt dat samenwerkend leren, al dan niet in verschillende vormen, een positief effect heeft op kennisontwikkeling bij leerlingen (o.a., Bossert, 1988; Rohrbeck, Ginsburg-Block, Fantuzzo & Miller, 2003; Roseth, Johnson & Johnson, 2008). Onderzoek wijst daarnaast uit dat wanneer leerlingen samen in groepen werken, ze zelfs meer kennis ontwikkelen dan wanneer ze dat niet doen (Kulik & Kulik, 1982; Lou et al., 1996). Het stimuleert het uitvoerig behandelen en onthouden van (conceptuele) kennis (Van Boxtel, Van der Linden, & Kanselaar, 2000; Janssen, Kirschner, Erkens, Kirschner & Paas, 2010), waarbij leerlingen hun ideeën delen en hier uitleg over geven (Howe et al., 2007; Lou et al., 1996). Daarnaast stimuleert het samenwerken in groepsverband de groepsleden om elkaar feedback te geven, elkaar verantwoordelijk te houden en aan te moedigen om te leren (Johnson, Johnson, & Smith, 2007). Er kunnen echter een aantal factoren zijn die een negatieve invloed hebben op het samenwerkingsproces. Dit is het geval wanneer één of meerdere groepsleden het gevoel hebben geen waardevolle bijdrage te kunnen leveren aan het groepsproces of niet te kunnen bijdragen aan het groepsdoel (Shepperd, 1993). Een daarmee samenhangende factor is ongelijke inbreng. Dit is bijvoorbeeld het geval wanneer bovengemiddelde leerlingen zichzelf meer competent achten of als meer competent worden beschouwd door andere groepsleden, waardoor zij actiever deelnemen aan en meer inbrengen tijdens de samenwerking (Cohen, 1994). Dit kan ten koste gaan van de motivatie van de andere leerlingen in de groep, hetgeen niet bevorderlijk is voor het samenwerkingsproces (Shepperd,

1993). Het enkel bij elkaar plaatsen van leerlingen in een groep is dus geen garantie voor een goede samenwerking (Gillies, 2003; Kreijns, Kirschner & Jochems, 2003; Lou et al., 1996). Twee belangrijke voorwaarden die verbonden zijn aan een goede samenwerking en die de bovengenoemde negatieve factoren beperken, zijn positieve wederzijdse afhankelijkheid en individuele verantwoordelijkheid (Brody & Davidson, 1998; Johnson et al., 2007). Positieve wederzijdse afhankelijkheid wordt gestimuleerd wanneer leerlingen in kleine groepen leren, waarbij ze kennis met elkaar delen en gezamenlijk aan een einddoel of groepsproduct werken (Cohen, 1994, Johnson et al., 2007). Leerlingen ontwikkelen meer kennis wanneer ze wat betreft de uitkomst van het groepsproces afhankelijk van elkaar zijn (Lou et al., 1996). Dit stimuleert hen elkaar aan te sporen om een bijdrage aan het gezamenlijke doel te leveren, bijvoorbeeld door het delen van informatie (Johnson et al., 2007). Om individuele verantwoordelijkheid te creëren dient elk groepslid verantwoordelijk gehouden te worden voor een bepaalde bijdrage aan het groepssucces (Johnson et al., 2007). Beide voorwaarden dragen eraan bij de inbreng van elke leerling relevant te maken, hetgeen leerlingen het gevoel geeft een essentiële bijdrage te kunnen leveren (Shepperd, 1993).

Een belangrijke indicator voor een goede samenwerking en voor kennisontwikkeling is het geven en ontvangen van uitleg (Webb, 1982a, 1982b, 1982c). Door het geven van uitleg herhalen leerlingen hardop wat ze geleerd hebben (Lou et al., 1996), wat hen helpt het geleerde te verduidelijken en te structureren (Webb, 1982a, 1982b). Om een uitgebreide uitleg te geven dient de leerling de informatie voor zichzelf helder te hebben, deze te organiseren en, indien de informatie onvolledig of onjuist blijkt, te herorganiseren. Daarnaast kan het nodig zijn, dat wanneer een groepsgenoot de uitleg niet volledig begrijpt, de leerling de uitleg herformuleert en op een andere manier uitlegt (Webb, 1989). Daarnaast kunnen leerlingen, wanneer ze uitleg ontvangen, niet alleen nieuwe kennis en inzichten opdoen, maar stelt het ze ook in staat om mogelijke misconcepties bij te stellen (Webb, 1982a, 1982b). Omdat het geven en ontvangen van uitleg belangrijk is met betrekking tot kennisontwikkeling, is het van belang dat elke leerling uitleg geeft en ontvangt en dat dit niet te veel uit balans is (Webb, 1982a, 1982b). De hoeveelheid uitleg die gegeven en ontvangen wordt, hangt samen met de groepssamenstelling en het competentieniveau van de leerlingen (Van Boxtel et al., 2000).

Groepssamenstelling

Leerlingen van verschillende competentieniveaus hebben niet allemaal baat bij dezelfde groepssamenstelling, dus het lijkt belangrijk om leerlingen op verschillende manieren te

groeperen (Tomlinson et al., 2003). Wat betreft groepssamenstellingen op basis van competentieniveau kunnen leerlingen zowel samenwerken in homogene groepen, waarin leerlingen van hetzelfde niveau gegroepeerd zijn, als in heterogene groepen, waarin leerlingen van verschillende niveaus gegroepeerd zijn. Homogeen groeperen biedt de mogelijkheid om methodes en materialen aan te passen aan het niveau van de leerling (Esposito, 1973). Deze groepssamenstelling blijkt dan ook het meest effectief als gebruik wordt gemaakt van instructie op maat (Kulik & Kulik, 1991; Lou et al., 1996). Een voordeel van deze groepssamenstelling is dat leerlingen over het algemeen meer voortbouwen op elkaar en dat discussies ontstaan op basis van gelijke inbreng, hetgeen positief samenhangt met de prestaties van leerlingen (Saleh, Lazonder, & De Jong, 2005). Dit gaat echter niet op voor leerlingen van alle competentieniveaus. Ondergemiddelde leerlingen profiteren over het algemeen meer van het samenwerken in heterogene groepen, aangezien het meer waarschijnlijk is dat ze hulp en uitleg krijgen (Lou et al., 1996; Saleh et al., 2005). Een homogene groepssamenstelling blijkt met name gunstig voor de kennisontwikkeling van gemiddelde leerlingen (Lou et al., 1996; Saleh et al., 2005), aangezien zij gemiddeld genomen meer uitleg geven en ontvangen in een homogene groepssamenstelling dan in een heterogene groepssamenstelling (Webb, 1982a, 1982b, 1988). Een verklaring hiervoor is dat gemiddelde leerlingen in een heterogene groepssamenstelling, wat betreft actieve inbreng, vaak ondergeschikt zijn aan de leerkracht-leerling verhouding die ontstaat tussen bovengemiddelde en ondergemiddelde leerlingen (Lou et al., 1996; Van Boxtel, Van der Linden, & Kanselaar, 2000). In heterogene groepen is de inbreng, wat betreft het geven van en de vraag om uitleg, over het algemeen minder gelijkmatig dan in homogene groepen. Groepsleden richten zich vaak tot het meest competente groepslid als het gaat om (extra) uitleg, hetgeen niet bevorderlijk is voor de interactie (Lou et al., 1996). In tegenstelling tot ondergemiddelde en gemiddelde leerlingen, blijken bovengemiddelde leerlingen zowel baat te hebben bij homogene als heterogene groepssamenstellingen (Saleh et al., 2005; Webb, Nemer & Zuniga, 2002). Ondanks dat uit onderzoek blijkt dat het homogeen groeperen van bovengemiddelde leerlingen, vanwege de mogelijkheid om discussies te voeren en kennis uit te wisselen op niveau, gunstig blijkt voor hun kennisontwikkeling (Kulik & Kulik, 1991), zijn er ook studies waaruit blijkt dat ze beter presteren in heterogene groepssamenstellingen, aangezien ze vaak de rol van leerkracht op zich nemen en relatief veel uitleg geven aan groepsleden (Lou et al., 1996; Webb, 1982a). Daarnaast geven bovengemiddelde leerlingen zelf aan in homogene groepssamenstellingen meer cognitieve voordelen (zoals meer uitdaging, sneller verloop van de samenwerking en meer discussie) te ervaren, terwijl ze in heterogene groepen meer

sociaal-emotionele voordelen (onder andere het helpen van anderen) ondervinden (Adams-Byers, Whitsell, & Moon, 2004).

Jigsaw

Een manier om de voordelen van zowel homogene als heterogene groepen te combineren is het toepassen van de jigsaw methode (Aronson, Blaney, Stephan, Sikes & Snapp, 1978). Over het algemeen blijkt dat leerlingen die de jigsaw methodiek volgen gelijk of beter presteren dan leerlingen in traditionele klassensituaties (Aronson, 1978; Karacop & Doymus, 2013; Walker & Crogan, 1998). Het kernpunt van deze aanpak is dat leerlingen in kleine, heterogene groepen samenwerken en een gezamenlijk einddoel hebben. Dit doel kan alleen worden bereikt als elke leerling kennis inbrengt. De informatie die nodig is om het einddoel te behalen wordt onderverdeeld, waarbij elke leerling verantwoordelijk wordt voor een deelonderwerp. In homogene groepjes doen leerlingen gezamenlijk kennis op over dit onderwerp. Doordat leerlingen van gelijke competentieniveaus samenwerken, kunnen ze de snelheid waarin de samenwerking verloopt en het niveau van discussies per groep afstellen op hun niveau. Hierna keren de leerlingen terug naar de heterogene groepen waar de kennis van alle leerlingen samengevoegd dient te worden om het einddoel te behalen. Het creëren van expertrollen, waarbij uitleg van leerlingen onderling vereist wordt, is een manier om het geven van uitgebreide uitleg te stimuleren (Webb, 1989). Doordat de heterogene groep niet optimaal kan presteren zonder hulp en uitleg van groepsgenoten, wordt individuele verantwoordelijkheid en positieve wederzijdse afhankelijkheid (Bossert, 1988; Johnson et al., 2007) bevorderd. Aangezien de inbreng van elke leerling essentieel is, heeft dit een positieve invloed op het creëren van gelijkmatige inbreng. De jigsaw methode biedt hiermee een oplossing voor genoemde knelpunten die kunnen ontstaan tijdens het samenwerken.

Differentiëren met jigsaw

Een methodiek die de jigsaw combineert met differentiatie op niveau is de STIP-methodiek, wat staat voor Samenwerken tijdens Taak-, Inhoud- en Procesdifferentiatie. Bij deze aanpak worden leerlingen van gelijke competentieniveaus bij elkaar geplaatst in (homogene) expertgroepen. In deze groepssamenstelling wordt voorzien in differentiatie op zowel *taak* (taken die veelal een beroep doen op lagere orde denkvaardigheden voor ondergemiddelde leerlingen versus taken die vooral een beroep doen op hogere orde denkvaardigheden voor bovengemiddelde leerlingen), *inhoud* (concrete inhoud voor ondergemiddelde leerlingen

versus abstracte(re) inhoud voor bovengemiddelde leerlingen), als *proces* (directe procesbegeleiding voor ondergemiddelde leerlingen versus coachende procesbegeleiding voor bovengemiddelde leerlingen). Tijdens deze groepssamenstellingen leren leerlingen, door middel van experimenten, over een uniek deelonderwerp waarover ze expert worden. Vervolgens komen leerlingen van verschillende competentieniveaus samen in (heterogene) ontwerpgroepen, waarin ze kennis uitwisselen over de verschillende expertises. Deze informatie is nodig wanneer ze vervolgens werken aan een groepsopdracht, met een overkoepelend onderwerp, waarbij kennis over alle onderwerpen samenkomt.

Huidig onderzoek

In deze studie ligt de focus op het onderzoek naar het effect dat de STIP-aanpak heeft op de kennisontwikkeling van leerlingen van verschillende competentieniveaus. Daarnaast wordt gekeken naar de samenhang tussen de groepsinbreng tijdens de kennisuitwisseling in de heterogene groepen en individuele kennisontwikkeling. Ook wordt gekeken naar de relaties tussen de kwaliteit van de opbrengst van de homogene samenwerking, de inbreng van de individuele leerling tijdens de kennisuitwisseling in de heterogene groepen, de kwaliteit van de opbrengst van de heterogene samenwerking en de kennisontwikkeling. Dit heeft geleid tot een drietal onderzoeksvragen:

1. Wat is het effect van de STIP-aanpak op de kennisontwikkeling van leerlingen van verschillende competentieniveaus?
2. Zijn er verschillen tussen heterogene groepen wat betreft de groepsinbreng tijdens de kennisuitwisseling en hoe hangt dit samen met de individuele kennisontwikkeling van leerlingen?
3. Hoe hangt de kwaliteit van de opbrengst van de homogene samenwerking, de inbreng van de individuele leerling tijdens de kennisuitwisseling in de heterogene groepen, de kwaliteit van de opbrengst van de heterogene samenwerking en de kennisontwikkeling met elkaar samen?

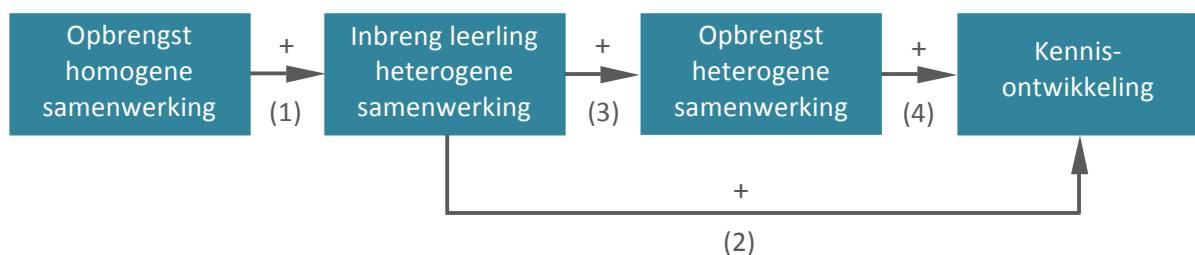
Om dit te onderzoeken worden voor de eerste vraag twee condities met elkaar vergeleken: een experimentele conditie waarin lessen volgens de STIP-methodiek worden gegeven en een controle conditie waarin reguliere lessen worden gegeven. Voor de tweede en derde vraag wordt aanvullend onderzoek gedaan binnen de experimentele conditie.

Gezien de cognitieve voordelen van het leren op niveau en van het samenwerken in verschillende groepssamenstellingen ten opzichte van klassikaal leren, is de verwachting dat

leerlingen in de experimentele conditie kennis zullen ontwikkelen en dat deze ontwikkeling, afhankelijk van de inbreng van de leerlingen tijdens de kennisuitwisseling in de heterogene groepen, gelijk of hoger is dan bij leerlingen in de controle conditie. Afhankelijk van de hoeveelheid uitleg die gegeven en ontvangen wordt per leerling, worden geen verschillen in kennisontwikkeling verwacht tussen leerlingen van verschillende competentieniveaus. Wel wordt verwacht dat de voorkennis van deze leerlingen onderling zal verschillen (oplopend met het competentieniveau).

Gebaseerd op het feit dat het geven en ontvangen van uitleg een belangrijke indicator is voor kennisontwikkeling, wordt een positieve relatie verwacht tussen de groepsinbreng tijdens de kennisuitwisseling in de heterogene groepen en de individuele kennisontwikkeling van leerlingen. Hoe meer kennis gegeven en ontvangen wordt binnen een groep, hoe meer kennis leerlingen zullen ontwikkelen.

In Figuur 1 is de verwachte samenhang tussen de verschillende componenten van de derde onderzoeksvraag weergegeven. Verondersteld wordt dat de mate waarin leerlingen expert zijn, hetgeen wordt weergegeven door middel van de kwaliteit van de opbrengsten van de homogene samenwerking, samenhangt met de individuele inbreng tijdens kennisuitwisseling in de heterogene groepen (1). Dit is gebaseerd op het feit dat het creëren van expertrollen het geven van uitgebreide uitleg stimuleert. Daarnaast wordt verwacht dat de individuele inbreng tijdens de kennisuitwisseling in de heterogene groepen zowel positief samenhangt met de kennisontwikkeling (2) als met de kwaliteit van de opbrengsten van de heterogene samenwerking (3). Ook wordt een positieve relatie verwacht tussen de kwaliteit van opbrengst van de heterogene samenwerking en de kennisontwikkeling (4). Dit is wederom gebaseerd op het feit dat leerlingen kennis ontwikkelen door het geven en ontvangen van uitleg. Dus hoe meer uitleg elke leerling geeft en ontvangt, hoe beter de kwaliteit van de opbrengst van de heterogene samenwerking en hoe groter de kennisontwikkeling naar verwachting zal zijn.



Figuur 1. Verwachte samenhang tussen de opbrengst van de homogene samenwerking, de inbreng van de leerling, de opbrengst van de heterogene samenwerking en de kennisontwikkeling.

Methode

Design

Deze studie is uitgevoerd in de context van een groter onderzoek. In dat onderzoek werd gekeken naar het effect van de STIP-aanpak op het differentiatiegedrag van leerkrachten. Daartoe werden de leerkrachten die met STIP gingen werken getraind. Als gevolg hiervan heeft de huidige studie een quasi-experimentele onderzoeksopzet, waarin klassen ofwel de STIP-lessen (experimentele conditie), ofwel reguliere lessen (controle conditie) hebben gevolgd.

Deelnemers

Aan deze studie hebben 244 leerlingen uit groep 6 deelgenomen (110 jongens, 134 meisjes; $M_{\text{leeftijd}} = 9.30$, $SD = .54$, variërend van 8 tot 11 jaar), afkomstig uit 15 klassen van 11 verschillende katholieke basisscholen uit Enschede en omgeving. Van deze leerlingen hebben 95 leerlingen (8 klassen) deelgenomen in de experimentele conditie (34 jongens, 61 meisjes; $M_{\text{leeftijd}} = 9.37$ jaar, $SD = .60$) en 149 leerlingen (7 klassen) in de controle conditie (76 jongens, 73 meisjes; $M_{\text{leeftijd}} = 9.26$ jaar, $SD = .48$). In eerste instantie deden 309 leerlingen mee aan het onderzoek, maar leerlingen die de voortoets of natoets niet hebben gemaakt, die afwezig waren bij een van de lessen of die in heterogene groepen hebben gezeten waarbij de leerkracht de rol van een afwezige leerling heeft overgenomen, zijn niet meegenomen in de analyse. Daarnaast is in één van de klassen de voortoets met de leerlingen besproken, wat ertoe heeft geleid dat ook deze leerlingen zijn uitgesloten van de analyse. Voorafgaand aan het onderzoek zijn de ouders van leerlingen in de experimentele conditie door middel van een brief op de hoogte gesteld over het STIP-project en het feit dat hun kind gefilmd zou worden tijdens één van de lessen. Naar aanleiding van deze brief hadden ze de mogelijkheid om hier geen toestemming voor te verlenen. Ouders van één leerling hebben deze toestemming niet verleend, wat ertoe heeft geleid dat de leerling in kwestie niet heeft deelgenomen aan de betreffende les.

Aan de leerkrachten van zowel de experimentele conditie als de controle conditie is gevraagd om de leerlingen te rangschikken op competentieniveau. Hierbij dienden ze rekening te houden met het algemene niveau van de leerling, het niveau op het gebied van het vakdomein (in dit geval 'Natuur & Techniek') en de CITO-score van de leerling. Vervolgens zijn de leerlingen op basis van de rangorde ingedeeld in drie categorieën, waarbij een grove

verhouding van 1:3:1 is aangehouden: ‘ondergemiddeld’ (experimentele groep: n = 17 (18%); controle groep: n = 30, (20%)), ‘gemiddeld’ (experimentele groep: n = 57 (60%); controle groep: n = 85, (57%)) of ‘bovengemiddeld’ (experimentele groep: n = 21 (22%); controle groep: n = 34 (23%)). Aan de leerkrachten in de experimentele conditie is ten slotte gevraagd om op basis van bovenstaande indeling minimaal vijf (homogene) expertgroepen met ieder vijf (of, bij een even totaal aantal leerlingen, vier of zes) leerlingen samen te stellen. Daarnaast is hen gevraagd om leerlingen in te delen in (heterogene) ontwerpgroepen, waarbij in elke ontwerpgroep minimaal één leerling uit elke expertgroep vertegenwoordigd zou zijn. Dit heeft geresulteerd in 22 heterogene ontwerpgroepen.

Domein

Er zijn voor het STIP-project zes modules ontwikkeld, die allemaal binnen het Natuur & Techniek domein vallen. In deze studie stond de module centraal die betrekking heeft op het onderwerp ‘Het weer’. Dit overkoepelende onderwerp is voor zowel de experimentele als de controle conditie onderverdeeld in vijf deelonderwerpen, namelijk ‘Temperatuur’, ‘Regenboog’, ‘Wind’, ‘Neerslag’ en ‘Onweer’. Deze onderwerpen waren gekoppeld aan één van de vijf leerdoelen (leerlingen leren: (1) hoe je temperatuur kunt meten; (2) hoe een regenboog ontstaat; (3) hoe wind ontstaat; (4) hoe neerslag ontstaat; en (5) hoe onweer ontstaat).

Materialen

Werkbladen les 1 STIP. Het werkblad voor de eerste les verschilde wat inhoud betreft per expertgroep. Afhankelijk van het competentieniveau, leerden leerlingen over één van de vijf deelonderwerpen. Op het werkblad van de ondergemiddelde leerlingen stond het deelonderwerp ‘Temperatuur’ centraal. Bij de gemiddelde leerlingen stonden ‘Regenboog’, ‘Wind’ en ‘Neerslag’ (oplopend in moeilijkheidsgraad) centraal en voor de bovengemiddelde leerlingen was ‘Onweer’ het centrale deelonderwerp.

Wat de opzet betreft, kwamen de werkbladen voor elke expertgroep overeen. Elk werkblad startte met een beschrijving van de voorbereiding, waarin stond wat de leerlingen nodig hadden voor de betreffende les en waar ze werden gevraagd om de taken te verdelen. Vervolgens stond op elk werkblad een experiment met bijbehorende opdrachten beschreven. Afhankelijk van het competentieniveau van de groep, verschilden deze opdrachten wat betreft inhoud (concreet versus abstract), type taak (lagere orde denkvaardigheden versus hogere

orde denkvaardigheden stimulerend) en type feedback (directe versus coachende procesbegeleiding). Alle opdrachten hadden gemeen dat het om een experiment ging, waarbij de leerlingen eerst gevraagd werd hun voorspellingen op te schrijven, vervolgens het bijbehorende proefje uit te voeren en daarna hun bevindingen te noteren. De opdrachten werden op elk werkblad ondersteund door informatieve teksten over het betreffende deelonderwerp, al dan niet aangevuld met een connectie naar het gedane proefje. Ook stond op het werkblad een link naar een informatief filmpje over het behandelde onderwerp, waarmee leerlingen hun bevindingen konden controleren. Onderaan stond de eindopdracht van de eerste les geformuleerd, waarbij leerlingen drie belangrijke leerpunten uit de les moesten opschrijven.

Werkbladen les 2 STIP. Het werkblad voor de tweede les was voor elke ontwerpgroep gelijk. Ook dit werkblad startte met een beschrijving van de voorbereiding, waarin de les kort werd toegelicht. Vervolgens werden de leerlingen, met ondersteuning van een stappenplan, aangespoord om hun kennis, opgedaan in de eerste les, met elkaar te delen. In het stappenplan werd de leerlingen gevraagd om één voor één te vertellen wat ze in de vorige les geleerd hadden. Aan de leerlingen die niet aan het woord waren werd de opdracht gegeven om drie belangrijke punten van hetgeen verteld werd, op te schrijven op het bijbehorende werkblad. Daarnaast werden leerlingen gestimuleerd om goed naar elkaar te luisteren en om uitleg te vragen wanneer ze iets niet begrepen. Op het tweede deel van het werkblad stond, wederom stap voor stap, beschreven hoe de leerlingen de eindopdracht van de tweede les, het maken van een weerposter, dienden uit te voeren en wat ze hiervoor nodig hadden. Aan iedere leerling werd gevraagd een deelonderwerp voor zijn rekening te nemen, waarbij vermeld werd dat het niet het deelonderwerp mocht zijn waarover ze in de eerste les geleerd hadden. Tenslotte stond beschreven dat leerlingen hun poster dienden te controleren aan de hand van een checklist en om een poster van een andere groep te bekijken aan de hand van vragen op het werkblad.

Weerposter les 2 STIP. De eindopdracht van de tweede les bestond uit een weerposter met daarop de waterkringloop afgebeeld. Leerlingen dienden de poster zelf aan te vullen met icoontjes die betrekking hadden op de behandelde deelonderwerpen. Verder stonden onder de poster, geordend per deelonderwerp, open vragen vermeld. Bij sommige vragen waren kernwoorden weergegeven, die leerlingen moesten gebruiken in hun antwoord.

De bedoeling was dat elke leerlingen vragen beantwoordde over een ander onderwerp dan waar hij in de eerste les over geleerd had.

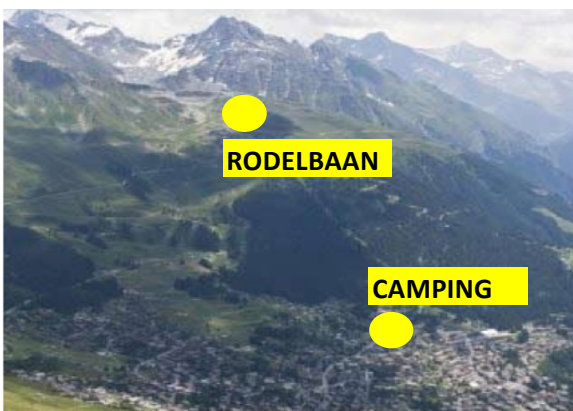
Tekstblad reguliere lessen. Het tekstblad voor de reguliere les is, wat betreft opzet en lay-out, ontwikkeld aan de hand van bestaande lesmethoden. De teksten zijn, om overeenstemming in leerdoelen te garanderen, tot stand gekomen op basis van de informatieve teksten op de werkbladen van de eerste STIP-les, aangevuld met enkele teksten uit bestaande lesmethoden die betrekking hadden op één van de deelonderwerpen. De teksten met bijbehorende afbeeldingen zijn opgedeeld in twee delen en kwamen wat betreft deelonderwerpen en belangrijkste concepten overeen met de inhoud van de STIP-lessen.

Werkblad reguliere lessen. Ook het werkblad voor de reguliere les is wat betreft type opdrachten ontwikkeld aan de hand van opdrachten uit bestaande lesmethoden. De opdrachten zijn opgesplitst in twee delen, waarbij voor elk deel zes opdrachten zijn geformuleerd. Deze deden vooral een beroep op conceptuele kennis en deels op het toepassen hiervan. Wederom kwamen de onderwerpen van de opdrachten overeen met die van de STIP-lessen.

Domeingerelateerde kennistoets. Om de kennisontwikkeling bij zowel de controlegroep als de experimentele groep te meten, zijn kennistoetsen opgesteld. De vragen op de toetsen waren gekoppeld aan de leerdoelen en deelonderwerpen. Om de leerwinst vast te stellen is zowel een voortoets als een natoets opgesteld. Beide toetsen bestonden uit vijftien parallelle, open vragen waarvan drie vragen per deelonderwerp. De parallelle vragen verschilden in formulering of hadden een andere context (zie voor voorbeelden van parallelle toetsitems Tabel 1). Van beide toetsen is door middel van een pilot gecontroleerd of de vragen niet te moeilijke woorden bevatten, of de vragen duidelijk waren, of er antwoord werd gegeven in de goede richting en ten slotte of de parallelle vragen hetzelfde maten. Op basis daarvan zijn de vragen, waar nodig, aangepast.

Voor zowel de voortoets als de natoets is de betrouwbaarheid van de toetsitems berekend (Cronbach's α). Bij de voortoets resulteerde dit in een waarde van $\alpha = .56$, bij de natoets in een waarde van $\alpha = .77$. Voor beide toetsen viel de waarde niet hoger uit als één of meerdere toetsitems zou worden verwijderd.

Tabel 1. Voorbeelden van parallelle toetsitems.

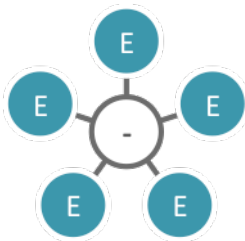
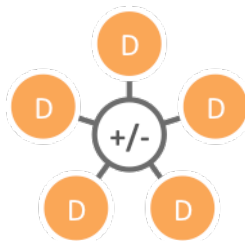
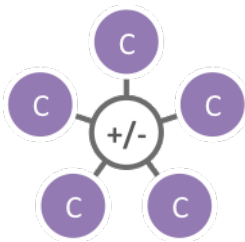
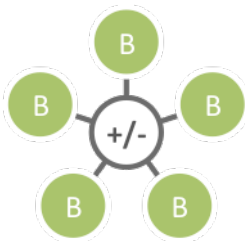
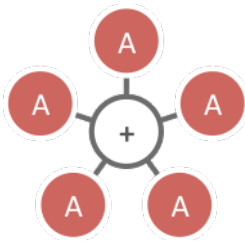
Voortoets	Natoets
	
Marieke en Peter wonen allebei in Oostenrijk. Marieke woont op de berg. Peter woont lager, in het dal. Ze hebben allebei een thermometer in de tuin hangen. Zijn de temperaturen die de thermometers aangeven dezelfde? Leg je antwoord uit.	Anne is met haar ouders op vakantie in Oostenrijk. Ze gaan een dagje naar de rodelbaan op de berg. Ze rijden met de auto van de camping naar de rodelbaan op de berg. Is de temperatuur op de camping hetzelfde als bij de rodelbaan? Leg je antwoord uit.
Iris zit op de bank. Er staat een glas water op tafel. Het licht van de lamp schijnt op het glas water. Daardoor ziet Iris een kleine regenboog. Hoe werkt dat? Leg je antwoord uit.	Wouter geeft de planten water met een plantenspuit. Er schijnt licht op de waterdruppels die uit de plantenspuit komen. Daardoor ziet Wouter een regenboog. Hoe werkt dat? Leg je antwoord uit.

Procedure

Maximaal één dag voorafgaand aan de eerste les hebben alle leerlingen, zowel in de experimentele conditie als in de controle conditie, de voortoets gemaakt. Voor het maken van de toets werd geen maximum tijd meegegeven. Nadat de kennistoetsen waren afgenomen zijn op scholen in de controle conditie de twee reguliere lessen over het weer gegeven en in de experimentele conditie de STIP-lessen gegeven. In beide condities hebben de lessen plaatsgevonden met een maximale tussenpoos van een week, waarbij voor elke afzonderlijke les één lesuur (ongeveer zestig minuten) beschikbaar was. De leerkrachten in de controle conditie hebben de lessen gegeven met behulp van het tekstblad en het werkblad met de opdrachten. In de eerste les kwamen de onderwerpen ‘Temperatuur’, ‘Regenboog’ en ‘Neerslag’ aan bod, tijdens de tweede les ‘Wind’ en ‘Onweer. Ze hebben de lessen gegeven zoals ze dat normaliter ook zouden doen. De leerkrachten in de experimentele conditie hebben de lessen gegeven volgens de STIP-methode. Zie Figuur 2 voor een representatie van deze methode.

Eerste les

Homogene expertgroepen



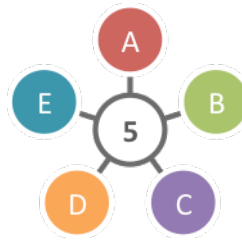
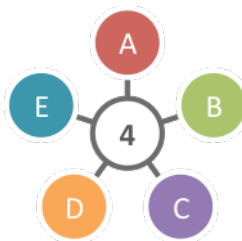
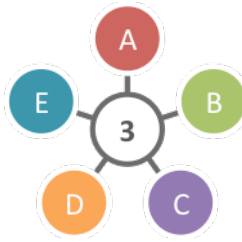
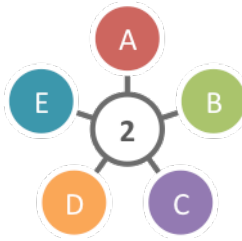
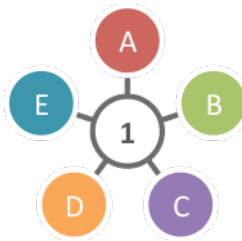
Bovengemiddelde leerling

Gemiddelde leerlingen

Ondergemiddelde leerling

Tweede les

Heterogene ontwerpgroepen



Figuur 2. Groepsindeling tijdens de STIP-lessen, volgens de benadering van de jigsaw methodiek.

De eerste STIP-les begon klassikaal met het bekijken van een introductiefilmpje en het activeren van voorkennis door middel van een woordweb. Vervolgens gingen de leerlingen in (homogene) expertgroepen uiteen om, met behulp van het werkblad, aan de slag gegaan met één van de deelonderwerpen. Bij de experimenten en opdrachten op het werkblad kreeg elk groepje gepaste procesondersteuning van de leerkracht (directe procesbegeleiding voor ondergemiddelde leerlingen versus coachende procesbegeleiding voor bovengemiddelde leerlingen). Nadat de leerlingen de eindopdracht van de eerste les hadden ingevuld, werd de eerste les ten slotte weer klassikaal afgesloten, waarbij kort werd teruggегреpen op het woordweb en op de les werd gereflecteerd. De tweede STIP-les werd opnieuw klassikaal gestart, waarbij de opdracht werd toegelicht en verteld werd in welke (heterogene) ontwerpgroepen de leerlingen waren ingedeeld. Vervolgens gingen de leerlingen uiteen ontwerpgroepen om, in drie fases, kennis uit te wisselen en te werken aan de weerposter. In de eerste fase hebben de leerlingen, met behulp van het stappenplan, elkaar verteld en uitleg gegeven over hetgeen ze in de eerste les gedaan en geleerd hebben. Daarna hebben ze, met behulp van de instructies op het werkblad, de poster voorzien van icoontjes. Vervolgens hebben de leerlingen in de laatste fase, met de informatie die ze tijdens de eerste fase hebben verkregen, de vragen beantwoord over één van de deelonderwerpen waarover ze zelf niet geleerd hadden in de eerste les. Aan het einde van de les hebben de groepen ten slotte elkaars poster beoordeeld en werd klassikaal gereflecteerd op de inhoud en het proces van de tweede les. Maximaal één dag na het afronden van de laatste les hebben de leerlingen, zowel in de experimentele conditie als in de controle conditie, de natoets gemaakt. De toetsen en lessen hebben in beide condities plaatsgevonden onder schooltijd.

Data-analyse

Domeingerelateerde kennistoetsen. Om de kennistoetsen te beoordelen, is zowel voor de voortoets als voor de natoets een antwoordmodel opgesteld. Voor de drie vragen per deelonderwerp kon in totaal een score van tien punten worden behaald, waarbij de maximale score voor de toets uitkwam op vijftig punten. Een tweede beoordelaar heeft 80 toetsen (ruim 16%) beoordeeld. De interbeoordelaarsbetrouwbaarheid (Cohen's Kappa) kwam voor de voortoets uit op een waarde van .93 en voor de natoets op een waarde van .92. Om de kennisontwikkeling van leerlingen in kaart te brengen is gebruik gemaakt van de verschilscore op de kennistoetsen (natoetsscore minus de voortoetsscore).

Opbrengst homogene samenwerking. Om de opbrengst van de homogene samenwerking te bepalen, is van elke leerling de eindopdracht van de eerste les beoordeeld. Hiervoor is een scoreformulier opgesteld, waarin de belangrijkste punten van elk deelonderwerp uit de eerste les zijn opgenomen. Voor zowel elk concept dat benoemd werd als voor elke uitleg waarin een concept juist gebruikt werd, kon één punt worden verdiend. Een tweede beoordelaar heeft 16 (ruim 17%) eindopdrachten van les 1 beoordeeld. De interbeoordelaarsbetrouwbaarheid (Cohen's Kappa) kwam uit op een waarde van .86.

Opbrengst heterogene samenwerking. Om de opbrengst van de heterogene samenwerking te bepalen, is van elke groep de eindopdracht van de tweede les beoordeeld. De eindopdracht is opgesplitst in twee onderdelen, namelijk het werkblad waarop de leerlingen belangrijke punten over elkaars deelonderwerpen hebben genoteerd en de weerposter. Hiervoor werden twee afzonderlijke scores gegeven. Om de belangrijke punten die de leerlingen over elkaars deelonderwerp hebben opgeschreven te beoordelen, is hetzelfde scoreformulier en dezelfde werkwijze gehanteerd als bij de eindopdrachten van les 1. Omdat in sommige groepen de punten door één leerling zijn opgeschreven en in andere groepen door alle leerlingen, is de groep als geheel beoordeeld. Wanneer er verschillen waren tussen hetgeen de leerlingen hadden opgeschreven, is uitgegaan van de hoogste score.

De score voor de weerposter was een combinatie van de score voor de geplaatste icoontjes en een score voor de antwoorden op de vragen. Voor elk icoontje dat op de juiste plek was geplaatst, kon één punt worden verdiend. Voor de vragen bij de poster is een antwoordmodel opgesteld. Bij sommige vragen stonden woorden vermeld die leerlingen moesten gebruiken in hun antwoord. In dat geval is voor elk woord dat op de juiste manier in het antwoord verwerkt was, één punt toegekend. Daarnaast werd één punt gegeven voor elk concept dat juist benoemd werd of dat op een goede manier in de uitleg verwerkt was. Een tweede beoordelaar heeft 15 subonderdelen (20%) beoordeeld van hetgeen de leerlingen als belangrijke punten genoteerd hadden en van vier posters (ruim 18%) met daarop de vragen over de vijf deelonderwerpen beoordeeld. Voor deze onderdelen kwam de interbeoordelaarsbetrouwbaarheid (Cohen's Kappa) respectievelijk uit op een waarde van .85. en .84.

Kennisuitwisseling. Om inzicht te krijgen in de kennisuitwisseling van de leerlingen in de heterogene groepen, is van elk groepje een audio- en video-opname gemaakt. Dit geeft de mogelijkheid om te achterhalen wat de inbreng van elke leerling binnen de groep was,

welke onderwerpen op welke manier aan bod zijn gekomen en wat de verhouding hiervan was binnen de groep.

Voor de analyse van de kennisuitwisseling zijn de audio- en video-opnames van de fase waarin de kennis moest worden gedeeld gecodeerd. Om de groepen die gebruikt zijn voor de analyse zo min mogelijk te beïnvloeden, is van elke groep het samenwerkingsproces opgenomen. Vervolgens zijn alleen de groepen die voldeden aan de ideale groepssamenstelling (dat wil zeggen, de groepen met een groepssamenstelling van vijf leerlingen, waarin elke expertise minimaal en maximaal één keer vertegenwoordigd was) in aanmerking gekomen voor analyse. Er waren in totaal 12 groepen die voldeden aan deze criteria. Van deze groepen is vervolgens een willekeurige selectie gemaakt, wat resulteerde in 6 groepen die zijn gecodeerd.

De analyse heeft in twee stappen plaatsgevonden. Ten eerste is van elke te analyseren groep de gehele dataset met behulp van ELAN (ELAN Multimedia Annotation Tool, 2013) gesegmenteerd. Een segment kon zowel een op zichzelf staande uitspraak van een leerling zijn, als een antwoord of reactie op één van de andere leerlingen en begon wanneer een leerling het woord nam en eindigde wanneer hij uitgesproken was, onderbroken werd door een andere leerling of wanneer er een stilte viel van meer dan twee seconden. Vervolgens heeft elk segment drie codes meegekregen. Ten eerste is bekeken door welke leerling de uitspraak werd gedaan. Als niet duidelijk was welke leerling het woord had, werd de code *onbekend* meegegeven. Stiltes of interrupties door leerkracht of onderzoeker werden niet gecodeerd. Vervolgens is bepaald of de gehele uitspraak als één segment geanalyseerd kon worden of dat deze opgedeeld diende te worden in subsegmenten. Dit is bijvoorbeeld het geval wanneer een leerling van onderwerp wisselt. Of een segment opgedeeld diende te worden, werd bepaald op basis van de verschillende concepten die benoemd zijn en die behoorden tot een van de expertises. Ten tweede is van elk (sub)segment het onderwerp vastgesteld. Hiervoor is eerst bepaald of de uitspraak *on-task* of *off-task* was. Alle *on-task* (sub)segmenten hebben vervolgens de code *temperatuur*, *regenboog*, *wind*, *neerslag*, *onweer* of *n.v.t.* meegekregen. Een (sub)segment kreeg de code *n.v.t.* wanneer het een taakgerelateerde, maar niet-inhoudelijke uitspraak betrof. Wanneer het onderwerp van de uitspraak niet duidelijk was, werd de code *onbekend* meegegeven. Ten derde is van de (sub)segmenten die een *on-task* code hebben gekregen de inhoud van de uitspraak bepaald (zie Tabel 2 voor een overzicht van deze codes).

Tabel 2. Codeerschema met codes die de inhoud van de uitspraak weergeven.

Code	Beschrijving	Voorbeeld
Uitleg concept	(Vraag om) uitleg over een feitelijk concept binnen de context van het weer.	“Warme lucht stijgt en koude lucht daalt, daardoor ontstaat wind.”
Uitleg proefje	(Vraag om) uitleg over een proefje, waarbij de uitkomst van het proefje duidelijk wordt, zonder dat daarbij de link gelegd wordt met het weer.	“We moesten met een ballon over ons haar wrijven en toen werd het haar statisch.”
Connectie	(Vraag om) uitleg over een proefje, waarbij een link wordt gelegd met (een feitelijke uitleg over) het weer.	“...die plusjes en minnetjes van de het haar en de ballon, zitten ook in de wolken als het onweert. Daardoor ontstaat stroom.”
Benoemen	Benoemen van een proefje of concept, zonder verdere uitleg.	“De kleuren van de regenboog.” “We gingen naar buiten om de temperatuur op te meten.”
Bevestiging	Bevestiging n.a.v. benoemen, uitleg of connectie door andere leerling.	“Ja.” “Uhuh.” “Oké.”
Stappenplan	Voorlezen van het stappenplan.	“Begin bij het eerste vakje op het werkblad. Degene die in les 1 over dat onderwerp heeft geleerd, legt uit wat hij heeft geleerd.”
Coördinatie	Taakgerelateerde, maar niet inhoudelijke uitspraken, waarbij geen kennisoverdracht plaatsvindt.	“Wie is er nu aan de beurt?” “We moeten eerst dit doen.”
Onbekend	Uitspraak waarvan de inhoud niet verstaanbaar is.	N.v.t.
N.v.t.	“Thinkink noises”, gelach of geluiden	“Als je eeh...”, “Pfff”

Het uitgangspunt hierbij was het in kaart brengen van de hoeveelheid uitleg die gegeven en gevraagd werd. Dit kon zowel feitelijke uitleg zijn over concepten binnen de context van het weer (*uitleg concept*), maar ook uitleg over de proefjes die leerlingen gedaan hebben (*uitleg proefje*) of uitleg waarbij een link werd gelegd tussen het proefje het en weer (*connectie*).

Verder zijn codes opgenomen om overige uitspraken te kunnen coderen. Wanneer leerlingen enkel dingen benoemden zonder daarbij uitleg te geven werd de code *benoemen* meegegeven. Een uitspraak kreeg de code *bevestigen* wanneer een leerling een inhoudelijke uitspraak van een andere leerling bevestigde. Daarnaast zijn twee codes opgenomen voor uitspraken die niet inhoudelijk, maar wel taakgerelateerd waren. Dit betrof de code *stappenplan*, die werd meegegeven wanneer het stappenplan voorgelezen werd, en de code *coördinatie* voor alle overige taakgerelateerde uitspraken. Een uitspraak kreeg de code *onbekend* als de inhoud van de uitspraak niet duidelijk was. Ten slotte werd de code *n.v.t.* meegegeven wanneer in de categorie “onderwerp” de code *off-task* was meegegeven of wanneer het overige uitingen betrof zoals “thinking noises”, gelach of geluiden¹.

Aangezien verwacht werd dat leerlingen kennis ontwikkelen door het geven en ontvangen van uitleg is, om de groepsinbreng tijdens de kennisuitwisseling in kaart te brengen, gebruik gemaakt van de absolute frequentie van de alle (sub)segmenten per groep waarin om uitleg werd gevraagd of uitleg werd gegeven. Dit is opgesplitst in de hoeveelheid (sub)segmenten per groep die de code *uitleg concept*, *uitleg proefje* of *connectie* hebben meegekregen. Voor de inbreng van de individuele leerling is gebruik gemaakt van de absolute frequentie van de (sub)segmenten per leerling waarin om uitleg werd gevraagd of uitleg werd gegeven. Dit is opgesplitst in de hoeveelheid (sub)segmenten per leerling die de code *uitleg concept*, *uitleg proefje* of *connectie* hebben meegekregen.

Resultaten

Kennisontwikkeling

Tabel 3 geeft een overzicht van de gemiddelde voortoetsscores, natoetsscores, verschilcores en standaarddeviaties. Variantieanalyse laat zien dat er geen significante verschillen waren in voortoetsscores tussen condities ($F(1, 242) = .11, p = .742$). Ook is geen significant interactie-effect tussen conditie en competentieniveau gevonden ($F(2, 238) = 1.55, p = .214$). In beide condities waren er significante verschillen in voortoetsscores tussen competentieniveaus (experimentele conditie: $F(2, 92) = 8.84, p < .001$; controle conditie: $F(2, 146) = 24.31, p < .001$). Post hoc analyses (Bonferroni) laten zien dat in beide condities de

¹ De intentie was dat een tweede codeur 15% van de (sub)segmenten zou coderen, om op die manier de interbeoordelaarsbetrouwbaarheid te kunnen bepalen. Dit heeft echter, vanwege persoonlijke omstandigheden van de tweede codeur, dusdanige vertraging opgeleverd dat de uiteindelijke interbeoordelaarsbetrouwbaarheid nog niet vastgesteld was ten tijde van het schrijven van deze these. Wel is het codeerschema meerdere malen samen doorgenomen en zijn fragmenten samen en afzonderlijk van elkaar gecodeerd en vervolgens besproken.

voortoetsscores van de gemiddelde leerlingen significant hoger waren dan van de ondergemiddelde leerlingen (experimentele conditie: $p = .007$; controle conditie: $p = .003$) en dat de voortoetsscores van de bovengemiddelde leerlingen zowel significant hoger waren dan van de ondergemiddelde leerlingen (experimentele conditie: $p < .001$; controle conditie: $p < .001$) als van de gemiddelde leerlingen (experimentele conditie: $p = .005$; controle conditie: $p < .001$). Gemiddeld genomen was de voorkennis dus voor elk competentieniveau gelijk tussen condities en verschillend binnen condities (oplopend met het competentieniveau). Bij alle Bonferroni analyses is eenzijdig getoetst en gebruik gemaakt van een aangepast significantieniveau van .017 per test (.05/3).

Tabel 3. Voortoetsscores, natoetsscores en verschillen per competentieniveau en conditie.

Competentieniveau	Conditie	N	Kennistoets					
			Voortoets		Natoets		Verschilscore	
			M	SD	M	SD	M	SD
Ondergemiddeld	Controle groep	30	6.97	2.24	14.22	6.03	7.25	5.27
	Experimentele groep	17	7.27	3.12	11.03	5.15	3.77	2.88
Gemiddeld	Controle groep	85	9.03	3.18	15.82	5.21	6.79	4.63
	Experimentele groep	57	9.22	2.49	13.09	5.28	3.87	4.62
Bovengemiddeld	Controle groep	34	12.69	4.54	22.67	7.65	9.97	5.76
	Experimentele groep	21	11.14	3.45	18.07	5.54	6.93	4.54
Totaal	Controle groep	149	9.45	3.88	17.06	6.73	7.61	5.17
	Experimentele groep	95	9.30	3.06	13.82	5.78	4.53	4.49

Gepaarde t-toetsen, waarbij eenzijdig is getoetst, wijzen uit dat in beide condities de natoetsscores significant hoger waren dan de voortoetsscores (experimentele conditie: $t(94) = 9.83$, $p < .001$, controle conditie: $t(148) = 17.97$, $p < .001$). In zowel de experimentele conditie als de controle conditie heeft dus kennisontwikkeling plaatsgevonden. Variantieanalyse laat echter zien dat de verschillen in de controle conditie significant hoger waren dan in de experimentele conditie ($F(1, 242) = 22.84$, $p < .001$). Hierbij is geen significant interactie-effect gevonden tussen conditie en competentieniveau ($F(2, 238) = .056$, $p = .945$). In beide condities waren er significante verschillen in verschillen tussen competentieniveaus (experimentele conditie: $F(2, 92) = 4.12$, $p = .019$; controle conditie: $F(2, 146) = 4.925$, $p = .009$). Post hoc analyses (Bonferroni) wijzen uit dat in beide condities de verschillen van de gemiddelde leerlingen en bovengemiddelde leerlingen significant van elkaar verschilden (experimentele conditie: $p = .007$; controle conditie $p = .002$). Er bleek geen significant

verschil te zijn tussen de verschillscores van de ondergemiddelde leerlingen en de gemiddelde leerlingen (experimentele conditie: $p = .931$; controle conditie $p = .671$) en tussen de verschillscores van de ondergemiddelde leerlingen de bovengemiddelde leerlingen (experimentele conditie $p = .042$; controle conditie $p = .049$). Bij alle Bonferroni analyses is gebruik gemaakt van een aangepast significantieniveau van .017 per test (.05/3).

Kennisuitwisseling

In totaal omvatte de geanalyseerde selectie 2349 (sub)segmenten in de fase ‘kennis delen’, variërend van 116 (sub)segmenten tot 983 (sub)segmenten per groep ($M = 391.50$, $SD = 328.83$). De totale tijd van de geanalyseerde selectie bedroeg 123 minuten, variërend van 7 minuten tot 43 minuten per groep ($M = 20.50$, $SD = 14.60$). Variantieanalyse wijst uit dat er tussen leerlingen van verschillende competentieniveaus geen verschillen waren in de relatieve frequentie van alle (sub)segmenten waarin om uitleg werd gevraagd of uitleg werd gegeven ($F(2, 25) = 2.50$, $p = .102$). Dus per groep hebben leerlingen van alle competentieniveaus naar verhouding evenveel uitleg gegeven of ontvangen.

Tabel 4 en Tabel 5 geven een overzicht van het totaal aantal (sub)segmenten in de fase ‘kennis delen’ per groep en de tijd die daar per groep aan besteed is. In Tabel 4 zijn daarnaast de absolute en relatieve frequenties van de (sub)segmenten per code in de categorie “onderwerp” weergegeven. In Tabel 5 zijn vervolgens de absolute en relatieve frequenties van de (sub)segmenten per code in de categorie “inhoud” weergegeven. De relatieve frequenties geven het aantal (sub)segmenten per code ten opzichte van het totaal aantal (sub)segmenten per groep weer. Wat opvalt is dat de tijd die de groepen besteed hebben aan de fase ‘kennis delen’ aanzienlijk verschilt per groep. Ook loopt het aantal uitspraken per groep waarin conceptuele uitleg werd gegeven uiteen en is in één groep (Groep E) tijdens deze fase zelfs helemaal geen uitleg gegeven over concepten. Daarnaast valt op dat in de meeste groepen de frequentie van de besproken subonderwerpen niet in balans lijkt. Verder blijkt dat over het algemeen relatief veel coördinerende uitspraken zijn gedaan. Voorbeelden hiervan zijn weergegeven in Tabel 6. Het betrof onder andere uitspraken of onduidelijkheden met betrekking tot het vervullen van de taken (nr. 6, 7, 10, 11, 15-21), discussies over de rolverdelingen (nr. 3-5) en uitspraken of discussies over hetgeen er op welk moment gedaan moest worden (nr. 1, 2, 8, 9, 11-14). Ten slotte blijkt dat het stappenplan relatief weinig werd voorgelezen.

Tabel 4. Overzicht per groep van het totaal aantal (sub)segmenten, de totale tijd besteed aan de fase ‘kennis delen’ en de absolute en relatieve frequenties van de (sub)segmenten per code in de categorie ‘onderwerp’.

	Groep					
	A	B	C	D	E	F
(Sub)segmenten	568	198	116	235	249	983
Tijd (in minuten)	34	10	7	14	12	26
On-task						
Temperatuur	49 (8.63%)	2 (1.01%)	12 (10.34%)	14 (5.96%)	26 (10.44%)	64 (6.51%)
Regenboog	66 (11.62%)	11 (5.56 %)	6 (5.17%)	36 (15.32%)	24 (9.64%)	88 (8.95%)
Wind	38 (6.69%)	12 (6.06 %)	8 (6.90%)	14 (5.96%)	30 (12.05%)	70 (7.12%)
Neerslag	34 (5.99%)	11 (5.56 %)	4 (3.45%)	10 (4.26%)	31 (12.45%)	72 (7.32%)
Onweer	34 (5.99%)	18 (9.09 %)	4 (3.45%)	43 (18.30%)	16 (6.43%)	54 (5.49%)
Off-task	18 (3.17%)	73 (36.87%)	4 (3.45%)	0 (0.00%)	11 (4.42%)	177 (18.01%)
Onbekend	19 (3.35%)	7 (3.53%)	7 (6.03%)	7 (2.98%)	6 (2.41%)	32 (3.26 %)
N.v.t. ²	310 (54.58%)	64 (32.32%)	71 (61.21%)	111 (47.23%)	105 (42.17%)	426 (43.37%)

Tabel 5. Overzicht van het totaal aantal (sub)segmenten, de totale tijd besteed aan de fase ‘kennis delen’ en de absolute en relatieve frequenties per (sub)segment per groep in de categorie ‘inhoud’.

	Groep					
	A	B	C	D	E	F
(Sub)segmenten	568	198	116	235	249	983
Tijd (in minuten)	34	10	7	14	12	26
Uitleg concept	76 (13.38%)	15 (7.58%)	12 (10.34%)	37 (15.74%)	0 (0.00%)	76 (7.73%)
Uitleg proefje	23 (4.05%)	11 (5.56%)	4 (3.45%)	7 (2.98%)	12 (4.82%)	20 (2.03%)
Connectie	5 (0.88%)	0 (0.00%)	0 (0.00%)	9 (3.83%)	2 (0.83%)	0 (0.00%)
Benoemen	79 (13.91%)	14 (7.07%)	4 (3.45%)	45 (19.15%)	96 (38.55)	168 (17.09%)
Bevestiging	28 (4.93%)	14 (7.07%)	14 (12.07%)	17 (7.23%)	12 (4.82%)	35 (3.56%)
Stappenplan	9 (1.58%)	0 (0.00%)	1 (0.86%)	0 (0.00%)	0 (0.00%)	10 (1.02%)
Coördinatie	280 (49.30%)	55 (27.78%)	61 (52.59%)	105 (44.68%)	96 (38.55%)	383 (38.96%)
Onbekend	20 (3.52%)	7 (3.54%)	7 (6.03%)	5 (2.13%)	4 (1.61%)	29 (2.95%)
N.v.t.	48 (8.45%)	82 (41.41%)	13 (11.21%)	10 (4.25%)	27 (10.84%)	262 (26.65%)

² De hoeveelheid *n.v.t.* is hier relatief groot. Dit is vanwege het feit dat alle niet-inhoudelijke, maar wel taakgerelateerde (sub)segmenten in de categorie “inhoud” (d.w.z. *stappenplan* en *coördinatie* in Tabel 5), in de categorie “onderwerp” (Tabel 4) de code *n.v.t.* hebben meegekregen.

Tabel 6. Fragmenten uit diverse protocollen met betrekking tot coördinerende uitspraken.

Groep	Nr.	Leerling	Uitspraak
A	1	B	Volgens mij moeten we alles opschrijven.
	2	D	Nee, we moeten eerst dit gaan lezen.
	3	C	Oké, wie is de leider?
	4	E	Ik ben al een keer geweest.
	5	B	Ik niet, denk ik. (...)
	6	D	Nee, [Naam leerling C] moet.
	7	B	Jij moet.
	8	D	Ik heb het nog niet eens opgeschreven. (...)
	9	C	Eerst gaan we deze invullen.
	10	B	Oké, wie is er.
	11	C	Eén. [Naam leerling A], je moet het zelf niet opschrijven.
	12	E	Zelf opschrijven?
	13	C	Nee, je eigen vakje moet je niet opschrijven.
	14	B	Jij mag jouw eigen vakje niet opschrijven, alleen die van de ander, snap je?
C	15	E	Wie in groepje één zit die mag beginnen.
	16	C	Groepje één.
	17	E	[Naam leerling A]
	18	B	Wie is groepje één dan?
	19	C	Ben jij groepje één?
	20	E	Vijf.
	21	A	Ik ben groepje één.

Met behulp van Pearson's Correlaties is bepaald of er samenhang bestond tussen de groepsinbreng tijdens de kennisuitwisseling in de heterogene groepen en de kennisontwikkeling. Er zijn significante correlaties gevonden tussen de hoeveelheid conceptuele uitleg en de verschillscore op de kennistoetsen ($r = .656, p < .001$) en tussen de hoeveelheid connecties en de verschillscore op de kennistoetsen ($r = .564, p = .002$). De hoeveelheid uitleg over een proefje en verschillscore op de kennistoetsen bleek niet significant te correleren ($r = .354, p = .065$).

Vanwege bovenstaande bevindingen en de verschillen die er tussen de groepen lijken te zijn, is de overige data nader bekeken. Ondanks dat het wel de bedoeling was, bleek hieruit dat niet alle benodigde kennis werd uitgewisseld in de geanalyseerde fase 'kennis delen', maar dat ook later in het samenwerkingsproces nog relevante uitleg werd gegeven. Een aantal voorbeelden waaruit dit blijkt zijn opgenomen in Tabel 7. In Groep E werd tijdens de fase

‘kennis delen’ vooral verteld wat leerlingen in de eerste les gedaan hadden en niet zozeer wat daar de uitkomsten van waren of wat dat betekende voor het weer (nr. 28-42). Dit is ook terug te zien in de relatieve frequenties van Groep E in Tabel 5, waaruit blijkt dat leerlingen in deze groep helemaal geen conceptuele uitleg gaven. In de fase waarin leerlingen vragen op de poster moesten beantwoorden over elkaars onderwerp, bleek dan ook dat leerlingen elkaar nog aanvullende uitleg gaven (nr. 43-47). Ook in Groep F bleek nog niet alle relevante informatie uitgewisseld tijdens de fase ‘kennis delen’ en werd er tijdens het vragen beantwoorden om extra uitleg gevraagd en extra uitleg gegeven (nr. 48-54). In Groep A werd het delen van informatie in sommige gevallen expliciet uitgesteld naar een volgende fase (nr. 1-11). In een latere fase bleek dan ook dat leerlingen nog niet alle relevante informatie hadden uitgewisseld om de vragen te kunnen beantwoorden (nr. 12-27). Daarnaast bleek dat het stappenplan, ondanks dat deze niet in elke groep gebruikt werd, vooral gelezen werd als opdracht voorafgaand aan de kennisuitwisseling en dat deze niet, zoals bedoeld, stap voor stap gelezen werd.

Tabel 7. Fragmenten uit diverse protocollen met betrekking tot het geven van uitleg in verschillende fases.

Groep	Fase	Nr.	Leerling	Uitspraak
D	Kennis delen	1	E	[Naam leerling C], oké. De neerslag mag nu.
		2	C	Ja.
		3	E	Dan mag jij straks meer informatie aan die en die geven, ja?
		4	C	En vragen stellen.
		5	E	Ja.
		6	D	Moet ik even vertellen?
		7	E	Ja, jij mag.
		8	D	Oké (...). (...)
		9	E	Eehm.. ja, oké. Zo is het wel even goed toch voor iedereen? Want ik moet... <i>[Leerling B gebaart naar leerling A]</i> Ja?
		10	B	Maar ik heb nog bijna niks opgeschreven.
		11	E	Eehm.. weet je, je hoeft maar iets op te schrijven, want jij doet het daar niet over.
	Vragen poster	12	E	Ik heb een vraag. Maar hoe komt het eigenlijk dat het zeewater, dat dat dan weer regen wordt? En hoe wordt de regen dan weer zeewater?
		13	D	Meer eehm.. kijk je hebt een potje en daar zit warm water in.
		14	E	Ja.
		15	D	Er zit dan koud water op en dan gaat de warme lucht gaat verdampen. (...)

		16	E	Maar zit dan tussen dat warme en dat koude, zit daar dan iets in, iets tussen?
		17	D	Ja, eigenlijk die lucht en dat verdampt meer...
		18	E	Maar hoe komt het dat...
		19	D	...en als dat verdampt wordt het weer water.
		20	E	Oh, dus verdamping dat komt weer naar de wolken en dat wordt dan weer regen?
		21	D	Ja. (...)
		22	C	Hoe kan met een thermometer de temperatuur worden gemeten? Wat bedoel je daarmee? Dat heb je net nog helemaal niet verteld of wel? Want je vertelde: je kan de temperatuur meten met een thermometer en over iets van het kwik, maar ik wil weten: hoe kan je met een thermometer de temperatuur dan meten?
		23	A	Er zit eigenlijk meer zo'n buisje in... wat omhoog gaat en omlaag.
		24	C	Ja, dat is het kwik?
		25	A	Ja. (...)
		26	C	Dus de thermometer kan de temperatuur meten doordat er een buisje is... een Celsiusbuisje..? Een buisje die het kwik...?
		27	A	Nee, in dat buisje.. daar binnenin zit dat kwik (...).
E	Kennis delen	28	A	We hadden zeg maar een temperatuurmeter. Toen gingen we naar buiten. Toen hadden we opgemeten hoeveel temperatuur het vandaag was (...).
		29	C	Nee, we hadden een temperatuur. Daarmee gingen we naar buiten.
		30	A	Ja, we hebben een temperatuur
		31	E	Wat was het?
		32	A	We hebben een temperatuur.
		33	E	Thermometer bedoel je?
		34	C	Temperatuur. (...)
		35	B	Bij de regenboog gingen we... hadden we een kan water en een zaklamp en een spiegel.
		36	A	Een kan water.
		37	B	En een spiegel.
		38	C	Wacht.. we hadden... ja.
		39	B	En die spiegel moest je in het water doen.
		40	A	Ja, wacht even.
		41	E	Een kan water, en dan? (...)
		42	B	En we hadden ook een wit blaadje en toen moesten we daar het licht van de zaklamp opschijnen.

	Vragen	43	E	Ja, ik weet niks. Hoe ontstaat wind? (...)
	poster	44	C	Als de warme en koude lucht bij elkaar komt dan gaat het omhoog en dan komt er wind.
		45	E	Dus warme en koude lucht komt bij elkaar en dan komt er wind?
		46	C	Ja. (...)
		47	C	Omdat warme lucht omhoog gaat en koud omlaag.
F	Vragen	48	B	Wie had wind?
	poster	49	C	Ik!
		50	B	Hoe kan het dat warme en koude lucht niet even zwaar zijn?
		51	C	Omdat de warme lucht meer ruimte nodig heeft (...).
		52	B	Koude lucht?
		53	C	Koude lucht... nee, warme lucht. Omdat de warme lucht meer ruimte nodig heeft dan de koude lucht.
				(...)
		54	C	Help me! Hoe kan de regenboog ontstaan als de zon schijnt het tegelijkertijd regent? (...) Deze vind ik moeilijk. (...)

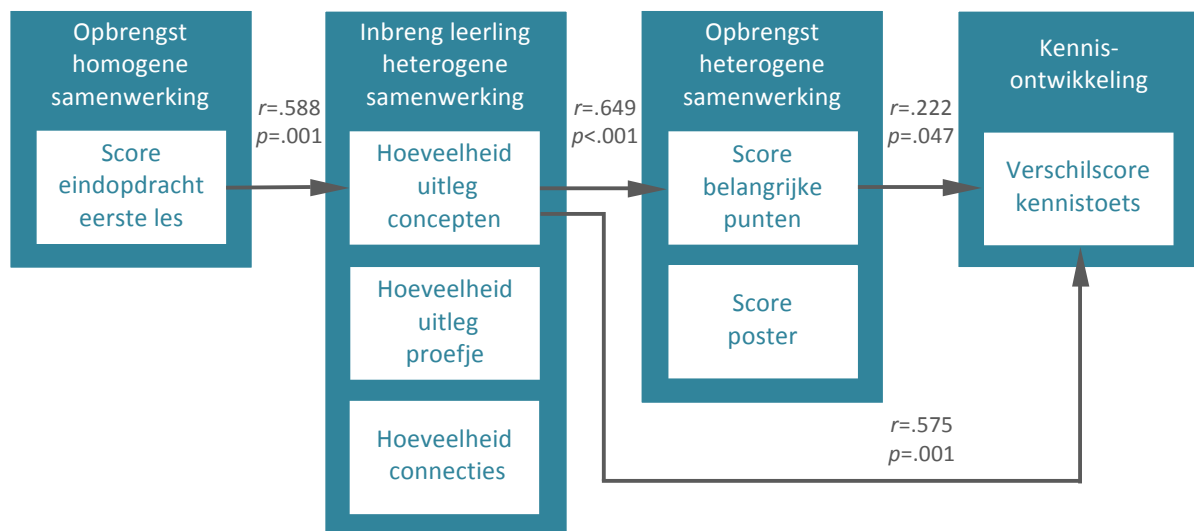
Samenhang

Tabel 8 geeft een overzicht van de gemiddelde score en standaarddeviatie van de eindopdracht van de eerste les, de gemiddelde hoeveelheid (sub)segmenten per leerling met de code *uileg proefje*, *uitleg concept* en *connectie* en de bijbehorende standaarddeviaties, de gemiddelde scores en standaarddeviaties voor de hetgeen de leerlingen als belangrijke punten hebben genoteerd en voor de poster en ten slotte de gemiddelde verschillscore met bijbehorende standaarddeviatie. Ondanks dat voor de scores op de eindopdracht van de eerste les geen maximum was vastgesteld, valt op dat deze relatief laag was.

Tabel 8. Gemiddelde scores, hoeveelheden en standaarddeviaties voor onderdelen van de opbrengsten van de samenwerking, de individuele inbreng en de kennisontwikkeling.

Opbrengst		Inbreng leerling						Opbrengst				Kennis-	
homogene samenwerking		heterogene samenwerking						heterogene samenwerking				ontwikkeling	
Score		Hoeveelheid		Hoeveelheid		Hoeveelheid		Score		Score		Verschilscore	
eindopdracht		uitleg		uitleg		connecties		belangrijke		poster		kennistoetsen	
eerste les		concept		proefje				punten					
(n = 91)		(n = 28)		(n = 28)		(n = 28)		(n = 81)		(n = 80)		(n = 95)	
M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
2.23	1.47	7.46	7.83	2.71	2.52	.57	1.35	12.85	4.23	19.34	4.27	4.53	4.49

Door middel van Pearson's Correlaties is bepaald of er een samenhang was tussen de kwaliteit van de opbrengst van de homogene samenwerking, de inbreng van de individuele leerling tijdens de kennisuitwisseling in de heterogene groepen, de kwaliteit van de opbrengst van de heterogene samenwerking en de kennisontwikkeling. In Figuur 3 zijn de relaties tussen de verschillende onderdelen gegeven die significant waren. Er zijn significante correlaties gevonden tussen de score voor de eindopdracht van de eerste les en het geven van conceptuele uitleg ($r = .001, p = .588$), het geven van conceptuele uitleg en de score voor de belangrijke punten ($r = .649, p < .001$), de score voor de belangrijke punten en de verschillscore op de kennistoetsen ($r = .222, p = .047$) en tussen de hoeveelheid conceptuele uitleg en de verschillscore op de kennistoets ($r = .575, p = .001$). Er bleek geen significante samenhang tussen de score voor de eindopdracht van de eerste les en de hoeveelheid uitleg over een proefje ($r = .057, p = .772$), de score voor de eindopdracht van de eerste les en de hoeveelheid connecties ($r = -.189, p = .334$), de hoeveelheid uitleg over een proefje en de score voor de belangrijke punten ($r = .250, p = .228$), de hoeveelheid uitleg over een proefje en de score voor de poster ($r = -.008, p = .968$), de hoeveelheid connecties en de score voor de belangrijke punten ($r = -.057, p = .787$), de hoeveelheid connecties en de score voor de poster ($r = .126, p = .524$) en tussen de score voor de poster en de verschillscore op de kennistoets ($r = .108, p = .338$).



Figuur 3. Relaties tussen de verschillende onderdelen die significant zijn.

Discussie en Conclusie

Het doel van dit onderzoek was om uit te zoeken wat het effect was van de STIP-aanpak op de kennisontwikkeling van leerlingen van verschillende competentieniveaus, hoe de mogelijke verschillen in inbreng tijdens de kennisuitwisseling in de heterogene groepen hiermee samenhangen en of er een verband bestond tussen de kwaliteit van de opbrengsten van de homogene samenwerking, de inbreng van de individuele leerling tijdens de heterogene samenwerking, de kwaliteit van de opbrengsten van de heterogene samenwerking en de kennisontwikkeling. Om dit te onderzoeken is gebruik gemaakt van twee condities: een experimentele conditie waarin lessen op basis van de STIP-aanpak werden gegeven en een controle conditie waarin reguliere lessen werden gegeven.

Voor de eerste onderzoeksvraag is onderzocht wat het effect was van de STIP-aanpak op de kennisontwikkeling van leerlingen van verschillende competentieniveaus. De verwachting was dat, afhankelijk van de kwaliteit van de kennisuitwisseling in de experimentele conditie, de kennisontwikkeling in de experimentele conditie gelijk of hoger zou zijn dan in de controle conditie. Zoals verwacht was werd geen verschil gevonden in voorttoetsscores tussen condities en liepen deze scores in beide condities op met het competentieniveau. Uit de resultaten bleek echter dat, hoewel leerlingen in beide condities kennis ontwikkelden, deze ontwikkeling bij leerlingen in de controle conditie groter was dan bij leerlingen in de experimentele conditie. Dit was voor leerlingen van alle competentieniveaus het geval. Tussen leerlingen van verschillende competentieniveaus binnen de experimentele conditie zijn, behalve tussen de gemiddelde en bovengemiddelde leerlingen, geen onderlinge verschillen gevonden in de hoeveelheid kennisontwikkeling. Een verklaring voor het gevonden verschil tussen condities kan zijn dat, ondanks dat in beide condities twee lesuren zijn besteed aan de lessen, de tijd die leerlingen hebben gehad om kennis tot zich te nemen in de experimentele groep korter was dan in de controle groep. De experimenten en de samenwerking in de eerste les vergden enige organisatie en coördinatie, zowel bij leerkrachten als leerlingen. Als gevolg daarvan kwamen leerlingen niet altijd toe aan het laatste onderdeel van de les, waarin ze geacht werden te reflecteren op hetgeen ze gedaan hadden en samen te vatten wat ze geleerd hadden. Daarnaast bleken veel leerlingen deze opdracht moeilijk te vinden. Het blijkt ook dat samenvatten een complexe taak is voor leerlingen, waarbij het onderscheiden van belangrijke en minder belangrijke informatie vaak lastig gevonden wordt (Dole, Duffy, Roehler & Pearson, 1991). Dit werd nog eens bemoeilijkt door het feit dat leerlingen, alvorens ze belangrijke informatie konden

samenvatten, deze informatie deels zelf moesten afleiden uit de gedane proefjes en hier conclusies over moesten trekken. Leerlingen waren hier vaak nog niet aan gewend. Ook het samenwerken in de heterogene groepen in de tweede les vereiste enige coördinatie. Dit bleek ook uit de relatief grote hoeveelheid coördinerende uitspraken die gedaan zijn door de leerlingen. Uit onderzoek van Weinberger, Stegmann, & Fischer (2010) blijkt dat dit leidt tot “process losses”. Ondanks dat een zekere mate van coördinatie nodig is, gaat een teveel ervan ten koste van inhoudelijke communicatie en dus van het uitwisselen van kennis. In vervolgonderzoek zou gekeken kunnen worden naar hoe leerlingen beter ondersteund kunnen worden in het afleiden van informatie en het samenvatten van bevindingen, zowel door het structureren van de eindopdracht, als door extra begeleiding van de leerkracht. Wellicht zouden leerlingen hier ook tijdens andere lessen mee kunnen oefenen. Wanneer leerlingen op deze manier meer leren, worden ze in staat gesteld om een (kwalitatief) betere bijdrage te leveren aan de kennisuitwisseling tijdens de heterogene samenwerking. Een alternatief zou zijn dat de lessen gegeven worden aan leerlingen die al meer ervaring hebben met deze vorm van werken. Daarnaast zou in vervolgonderzoek gekeken kunnen worden naar hoe het samenwerkingsproces in beide groepen beter gereguleerd kan worden en hoe leerkrachten hier meer sturing aan kunnen geven, zodat minder tijd verloren gaat aan organisatie en coördinatie, waardoor voor de leerlingen meer tijd overblijft om tot de inhoud te komen.

Een andere verklarende factor is het verschil in de mate van differentiatie door de leerkracht tijdens de homogene samenwerking. Uit het onderzoek waarbinnen deze studie heeft plaatsgevonden, bleek dat leerlingen van leerkrachten die tijdens het werken met STIP de begeleiding op meerdere manieren aanpasten (procesdifferentiatie), meer kennis ontwikkelden dan leerlingen van leerkrachten die dat niet deden (Hulsbeek & Eysink, 2015). In vervolgonderzoek zou wellicht gekeken kunnen worden naar hoe leerkrachten hier extra in begeleid en getraind kunnen worden.

Ten slotte is in dit onderzoek enkel gekeken naar de cognitieve opbrengsten, terwijl het samenwerkend leren in vergelijking met reguliere lessen ook andere voordelen kan opleveren, zoals het stimuleren van positieve relaties tussen leerlingen (Bossert, 1988), het ontwikkelen van sociale en communicatieve vaardigheden (Kreijns et al., 2003) en probleemoplossende vaardigheden (Lou et al., 1996). Veel van deze vaardigheden zijn onderdeel van de 21^e eeuwse vaardigheden, zoals communiceren, samenwerken, kritisch denken, probleemoplossende vaardigheden en sociale vaardigheden, die steeds belangrijker worden in het basisonderwijs (Onderwijsraad, 2011, 2014; Thijs, Fisser, & Van der Hoeven, 2014). Mede gezien het belang van deze vaardigheden, zou het interessant zijn om die ook in

kaart te brengen, zodat een completer beeld ontstaat van hetgeen leerlingen, die lessen volgens de STIP-aanpak hebben gevolgd, geleerd hebben.

Voor de tweede onderzoeksvraag is gekeken of er verschillen waren tussen de heterogene groepen wat betreft de kennisuitwisseling en de samenhang hiervan met de kennisontwikkeling. De verwachting was dat leerlingen die in groepen hebben gezeten waarin meer uitleg werd gegeven en ontvangen, meer kennis zouden ontwikkelen dan leerlingen die in groepen hebben gezeten waarin dit in mindere mate het geval was. Wat in de eerste plaats bleek is dat de tijd die besteed is aan de geanalyseerde fase waarin de kennis werd uitgewisseld, per groep aanzienlijk uiteen liep. De tijd die door de leerkracht beschikbaar was gesteld voor deze fase was in sommige gevallen verschillend. Ook bleek dat deze fase door sommige groepen verschillend aangepakt werd. Een voorbeeld hiervan is dat sommige leerlingen al tijdens de kennisuitwisseling afspraken wie welke vragen op de poster zou beantwoorden. In dat geval werd een deel van de relevante uitleg pas in een later stadium gegeven. Ook leken de onderwerpen die besproken zijn in de geanalyseerde fase niet altijd in balans. Wel bleek dat leerlingen van alle competentieniveaus naar verhouding evenveel uitleg gegeven of gevraagd hadden. De hoeveelheid conceptuele uitleg die per groep gegeven en ontvangen werd bleek positief samen te hangen met individuele kennisontwikkeling. Dit komt overeen met de eerder genoemde bevindingen van Webb (1982a, 1982b, 1982c). Uit de geanalyseerde protocollen bleek verder dat een deel van de uitleg enkel betrekking had op de proefjes die gedaan waren, hetgeen geen samenhang vertoonde met kennisontwikkeling. Hoewel het nauwelijks voorkwam, werd duidelijk dat wanneer de uitleg over proefjes in verband werd gebracht met theoretische concepten, het wel samen bleek te hangen met kennisontwikkeling. Wat daarnaast opviel is dat het stappenplan niet in alle groepen gebruikt werd en dat, wanneer het wel gebruikt werd, vooral gelezen werd als opdracht voorafgaand aan de kennisuitwisseling. Om een compleet beeld te krijgen van de besproken onderwerpen en van alle uitleg die gegeven is tijdens het samenwerken, zouden in vervolgonderzoek de gehele protocollen geanalyseerd moeten worden, eventueel aangevuld met de protocollen van de overige groepen. Op basis daarvan zou bepaald kunnen worden of en in welke mate het geven en ontvangen van uitleg ook een daadwerkelijke voorspeller is van kennisontwikkeling. Daarnaast zou gekeken kunnen worden in welke mate de besproken deelonderwerpen in balans dienen te zijn en of ze een voorspeller zijn van de kennisontwikkeling per deelonderwerp. Omdat de hoeveelheid uitleg per groep lijkt te verschillen zou ook onderzocht kunnen worden hoe het geven van uitleg beter ondersteund kan worden, niet alleen door het samenwerkingsproces beter te reguleren, maar ook door het gebruik van het stappenplan te

stimuleren en dit wellicht aan te passen. De opzet van het stappenplan zou op een manier aangepast kunnen worden, zodat leerlingen het meer stap voor stap gaan gebruiken en het niet enkel zien als een opdracht die ze vooraf dienen te lezen. Op die manier kunnen leerlingen op het juiste moment ondersteund worden, waarbij kan worden aangestuurd op het geven van conceptuele uitleg of het uitleg geven over de connectie tussen het experiment en conceptuele, theoretische kennis.

Voor de laatste onderzoeksvraag is gekeken of er een verband bestond tussen de opbrengst van de homogene samenwerking, de inbreng van de leerling tijdens de kennisuitwisseling in de heterogene groepen, de opbrengst van de heterogene samenwerking en de kennisontwikkeling. Er werd een positieve samenhang verwacht tussen de opbrengst van de homogene samenwerking en de inbreng van de leerling, tussen de inbreng van de leerling en de opbrengst van de heterogene samenwerking, tussen de inbreng van de leerling en kennisontwikkeling en tussen de opbrengst van de heterogene samenwerking en kennisontwikkeling. De verwachte samenhang was aanwezig, al bleek deze, wat betreft de inbreng per leerling, alleen te gelden voor de hoeveelheid conceptuele uitleg. Wat betreft de opbrengst van de heterogene samenwerking is, hoewel dit een zwakke relatie betrof, alleen een samenhang gevonden met hetgeen de leerlingen hadden opgeschreven als belangrijke punten. De score voor de poster bleek niet samen te hangen met de overige onderdelen. Een verklaring hiervoor kan zijn dat niet in alle groepen de opdracht is uitgevoerd zoals beoogd. Leerlingen zouden namelijk geen vragen beantwoorden over het onderwerp waarover ze in de homogene groepen geleerd hadden, maar over onderwerpen waarover andere leerlingen geleerd hadden. Dit bleek achteraf niet altijd het geval, waardoor de leerlingen minder afhankelijk waren van elkaars kennis, er wellicht minder uitleg is gevraagd en gegeven en daardoor minder kennis is ontwikkeld over onderwerpen waarin leerlingen niet geleerd hebben in de eerste les. In vervolgonderzoek zou gekeken kunnen worden naar een manier waarop de wederzijdse afhankelijkheid in meerdere mate gegarandeerd wordt, onder andere door dit een prominenter onderdeel te laten worden van het stappenplan. Daarnaast zou het ook hier een completer beeld geven wanneer in vervolgonderzoek alle uitspraken, ook die in latere fases zijn gedaan, met betrekking tot het geven en ontvangen van uitleg geanalyseerd zouden worden.

Concluderend kan worden gesteld dat, hoewel leerlingen die lessen op basis van de STIP-aanpak volgden minder kennis ontwikkelden dan leerlingen die reguliere lessen volgden, er potentie is voor meer kennisontwikkeling. De hoeveelheid conceptuele uitleg per groep en de hoeveelheid uitleg per groep over connecties tussen proefjes en theoretische

concepten, bleken namelijk positief samen te hangen met de individuele kennisontwikkeling. Ook werd gevonden dat hoe beter de opbrengst van de homogene samenwerking was, hoe groter de inbreng van de leerling, hoe beter de opbrengst van de heterogene samenwerking en hoe groter de kennisontwikkeling was. Daarnaast bleek dat de STIP-aanpak voorzag in gelijke inbreng van leerlingen van verschillende competentieniveaus. Verder biedt dit onderzoek een startpunt voor vervolgonderzoek naar ondersteuning van zowel leerkrachten als leerlingen en naar kennisuitwisseling in de heterogene groepssamenstelling, zodat een compleet beeld ontstaat van de verschillende onderdelen van de STIP-aanpak. Samen met de resultaten van dit onderzoek, kan dit relevante kennis en resultaten opleveren en een aanvulling vormen op het onderzoek dat al gedaan is naar de afzonderlijke onderdelen van de STIP-aanpak. Daarnaast levert het een waardevolle methodiek op die leerkrachten een handreiking biedt om te differentiëren, waardoor leerlingen uitgedaagd worden op hun eigen niveau en daarnaast werken aan waardevolle vaardigheden tijdens het samenwerken in verschillende groepssamenstellingen, hetgeen de voordelen van zowel homogene groepen als heterogene groepen combineert. Mits dit op de juiste manier ondersteund wordt, is ten slotte op cognitief vlak de potentie aanwezig om de kennisontwikkeling bij leerlingen verder te doen groeien.

Literatuur

- Adams-Byers, J., Whitsell, S. S., & Moon, S. M. (2004). Gifted students' perceptions of the academic and social/emotional effects of homogeneous and heterogeneous grouping. *Gifted Child Quarterly*, 48(1), 7-20.
- Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage.
- Bossert, S. T. (1988). Cooperative activities in the classroom. *Review of Research in Education*, 15, 225-250.
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35.
- Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research*, 61(2), 239-264.
- ELAN Multimedia Annotation Tool (2013). Verkregen via <https://tla.mpi.nl/tools/tla-tools/elan/>

- Esposito, D. (1973). Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. *Review of Educational Research*, 43(2), 163-179.
- Gillies, R. M. (2003). Structuring co-operative learning experiences in primary school. In R.M. Gillies & A. F. Ashman (Eds.), *Co-operative learning. The social and intellectual outcomes of learning in groups* (pp. 36-53). London: RoutledgeFalmer.
- Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., Jessiman, E., & Donaldson, C. (2007). Group work in elementary science: Towards organisational principles for supporting pupil learning. *Learning and Instruction*, 17(5), 549-563.
- Hulsbeek, M., & Eysink, T.H.S. (2015). Met STIP op één. *Didactief*, 45(9), 44-45.
- Inspectie van het Onderwijs (2015). *De staat van het onderwijs. Onderwijsverslag 2013/2014*. Utrecht: Inspectie van het Onderwijs.
- Janssen, J., Kirschner, F., Erkens, G., Kirschner, P. A., & Paas, F. (2010). Making the black box of collaborative learning transparent: Combining process-oriented and cognitive load approaches. *Educational Psychology Review*, 22(2), 139-154.
- Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, 19(1), 15-29.
- Karacop, A., & Doymus, K. (2013). Effects of jigsaw cooperative learning and animation techniques on students' understanding of chemical bonding and their conceptions of the particulate nature of matter. *Journal of Science Education and Technology*, 22(2), 186-203.
- Kulik, C. L. C., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal*, 19(3), 415-428.
- Kulik, J., & Kulik, C. (1991). *Research on ability grouping: Historical and contemporary perspectives*. Storrs, CT: National Research Center on the Gifted and Talented.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, 19(3), 335-353.
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66(4), 423-458.

- Onderwijsraad (2011). *Maatschappelijke achterstanden van de toekomst*. Den Haag: Onderwijsraad.
- Onderwijsraad (2014). *Een eigentijds curriculum*. Den Haag: Onderwijsraad.
- Reis, S. M., & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning and Individual Differences*, 20(4), 308-317.
- Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240.
- Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: the effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, 134(2), 223.
- Saleh, M., Lazonder, A. W., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science*, 33(2), 105-119.
- Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin*, 113, 67-81
- Thijs, A., Fisser, P., & Hoeven, M. van der (2014). *21e eeuwse vaardigheden in het curriculum van het funderend onderwijs*. Enschede: SLO.
- Tolmie, A. K., Topping, K. J., Christie, D., Donaldson, C., Howe, C., Jessiman, E., Livingston, K., & Thurston, A. (2010). Social effects of collaborative learning in primary schools. *Learning and Instruction*, 20(3), 177-191.
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L.A. & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2-3), 119-145.
- Van Bijsterveldt-Vliegenthart, M. (2011). *Naar passend onderwijs*. Verkregen via <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/kamerstukken/2011/01/31/naar-passend-onderwijs/266319.pdf>
- Van Boxtel, C., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction*, 10(4), 311-330.
- Walker, I., & Crogan, M. (1998). Academic performance, prejudice, and the jigsaw classroom: New pieces to the puzzle. *Journal of Community & Applied Social Psychology*, 8(6), 381-393.
- Webb, N. M. (1982a). Group composition, group interaction, and achievement in cooperative small groups. *Journal of Educational Psychology*, 74(4), 475-484.

- Webb, N. M. (1982b). Peer interaction and learning in cooperative small groups. *Journal of Educational Psychology*, 74(5), 642-655.
- Webb, N. M. (1982c). Student interaction and learning in small groups. *Review of Educational Research*, 52(3), 421-445.
- Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, 13(1), 21-39.
- Webb, N. M., Nemer, K. M., & Zuniga, S. (2002). Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Journal*, 39(4), 943-989.
- Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, 26(4), 506-515.