

Faciliterende effecten in onderzoekend leertaken: de rol van relaties tussen en binnen variabelen

Emiel van Lieburg

Universiteit Twente

Samenvatting

Dit onderzoek had tot doel het effect van bekendheid met relaties tussen en binnen variabelen bij onderzoekend leertaken op de leerprestatie en het onderzoeksgedrag van lerenden te bepalen. Van 57 deelnemers, verdeeld over drie condities, zijn de gemiddelde leerprestatie en het gemiddelde experimenteelgedrag (gemeten door middel van aantal geformuleerde hypothesen, aandeel geformuleerde specifieke hypothesen en het aantal uitgevoerde experimenten) middels een univariate variantieanalyse (ANOVA) vergeleken. Deze drie condities onderscheidden zich op basis van de taakinhoud van de onderzoekend leertaak die uitgevoerd diende te worden; deelnemers in de conditie 'concreet rijk' dienden een onderzoekend leertaak uit te voeren die concrete variabelen en concrete relaties tussen en binnen die variabelen bevatte, deelnemers in de conditie 'abstract' dienden een onderzoekend leertaak uit te voeren die abstracte variabelen en onbekende relaties tussen en binnen die variabelen bevatte en deelnemers in de conditie 'concreet arm' dienden een gemanipuleerde onderzoekend leertaak uit te voeren die concrete variabelen met onbekende relaties tussen en binnen die variabelen bevatte. Doel van de taakuitvoer was het bepalen van de invloed van vier onafhankelijke variabelen op een onafhankelijke variabele. Met betrekking tot de leerprestatie was de verwachting dat de gemiddelde prestatie van deelnemers in de conditie 'concreet rijk' hoger zou zijn dan die van deelnemers in de twee andere condities en dat er geen verschillen in gemiddelde leerprestatie gevonden zouden worden tussen de conditie 'abstract' en de conditie 'concreet arm'. Deze hypothese werd slechts gedeeltelijk door de resultaten ondersteund: er bleek geen statistisch significant verschil in gemiddelde leerprestatie te bestaan tussen de conditie 'concreet rijk' en de conditie 'abstract' en tussen de conditie 'abstract' en de conditie 'concreet arm', maar wel tussen de conditie 'concreet rijk' en de conditie 'concreet arm'. De verwachting ten aanzien van het experimenteelgedrag was dat deelnemers in de conditie 'concreet rijk' gemiddeld meer en specifiekere hypothesen zouden formuleren en gemiddeld minder experimenten uit zouden voeren dan deelnemers in de twee andere condities en dat er op deze vlakken geen verschillen gevonden zouden worden tussen deze beide andere condities. Hoewel de hypothesen niet op alle vlakken van de beschreven wijze van meten van experimenteelgedrag bevestigd worden, kan over het algemeen gesteld worden dat deelnemers in de conditie 'concreet rijk' beter experimenteelgedrag hebben getoond dan deelnemers in de conditie 'concreet arm' en de conditie 'abstract' en dat er geen noemenswaardige verschillen in experimenteelgedrag bestonden tussen deelnemers uit de conditie 'concreet arm' en de conditie 'abstract'. Op basis van de bevindingen is geconcludeerd dat het bij het uitvoeren van een onderzoekend leertaak met een concrete inhoud, de combinatie van concreetheid van variabelen en concreetheid van relaties tussen en binnen die variabelen is, die faciliteert. Naast conclusies worden in de discussie mogelijke verklaringen voor de gevonden resultaten gegeven, implicaties voor de praktijk bediscussieerd en aanbevelingen voor verder onderzoek gedaan.

Introductie¹

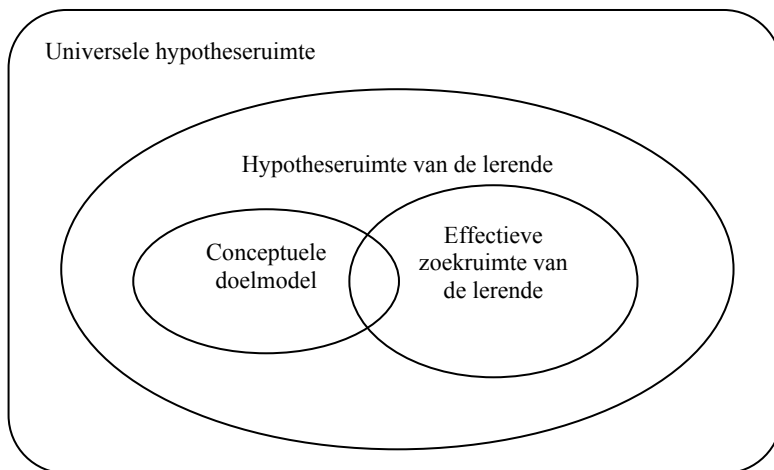
In onderzoekend leren proberen lerenden regels te ontdekken met betrekking tot de relaties tussen variabelen door middel van het uitvoeren van experimenten (Wilhelm & Beishuizen, 2003). Op deze wijze nemen ze de leerstof op een actieve en zelfgestuurde wijze op; een belangrijk aspect van deze constructivistische vorm van leren (De Jong & Van Joolingen, 1998). Het gebruik van onderzoekend leren in het onderwijs lijkt veelbelovend, maar lerenden hebben moeite met het uitvoeren van onderzoekend leertaken (De Jong & Van Joolingen, 1998). Eén van de factoren die invloed heeft op onderzoekend leren is de aard van het domein waarin de taak is ingebed; in isomorfe taken (taken met dezelfde onderliggende structuur en daarmee identiek met betrekking tot de relaties tussen de onafhankelijke variabelen en de afhankelijke variabele) leiden verschillende domeinen tot variaties in leerprestatie en leerprocessen (Kuhn, Garcia-Mila, Zohar, & Anderson, 1995; Wilhelm & Beishuizen, 2003). Echter, om de effectiviteit van onderzoekend leren te kunnen optimaliseren is er meer inzicht nodig in de invloed van de specifieke karakteristieken van het domein op de leerprestatie.

Volgens Wilhelm en Beishuizen (2003) bestaat er een redelijke mate van consensus over de processen waaruit onderzoekend leren is opgebouwd: hypotheseformatie, experimentontwerp en data-interpretatie. De Jong en Njoo (1992) noemen deze processen transformatief; processen die direct kennis opleveren. Naast de transformatieve processen hebben zij het concept van leerregulatie, met de daaruit voortkomende regulatieve processen, toegevoegd. Deze regulatieve processen, zoals plannen en monitoren, zijn nodig om het onderzoekend leerproces te managen (De Jong & Njoo, 1992). Klahr en Dunbar (1988) hebben, om het onderzoekend leerproces te beschrijven, het 'Scientific Discovery as Dual-space Search' (SDDS) model geïntroduceerd. Volgens De Jong en Van Joolingen (1998) is het een belangrijke bijdrage aan het veld van onderzoekend leren. Het basisidee van het model is dat wetenschappelijk redeneren plaatsvindt in twee gerelateerde probleemruimtes; de hypotheseruimte en de experimenteeruimte. De hypotheseruimte is een zoekruimte die alle mogelijke regels betreffende de fenomenen die in het domein geobserveerd kunnen worden, beschrijft. De experimenteeruimte bestaat uit de experimenten die uitgevoerd kunnen worden in het domein, evenals de uitkomsten van deze experimenten (De Jong & Van Joolingen, 1998). Het model geeft aan dat onderzoekend leren uit drie hoofdcomponenten bestaat: het doorzoeken van de hypotheseruimte, het testen van hypothesen en het evalueren van bewijs, deze worden verder uitgewerkt in sequentiële subprocessen. Volgens Klahr en Dunbar verloopt dit proces van onderzoekend leren als volgt: het resultaat van het doorzoeken van de hypotheseruimte is een gespecificeerde hypothese, die als input dient voor een experiment. Deze hypothese kan getest worden door een experiment uit te voeren, een voorspelling over de uitkomst te doen en deze uitkomst vervolgens te observeren. Het resultaat van hiervan is een beschrijving van bewijs die gebruikt wordt als input voor het evalueren van dit bewijs. Dit zorgt voor de beslissing of het bewijs en andere overwegingen de hypothese verwerpen of bevestigen en of de experimenten eventueel voortgezet dienen te worden. Voorkennis en experimenteerresultaten leiden het doorzoeken van de hypotheseruimte en de actuele hypothese leidt de zoektocht in de experimenteeruimte, die gebruikt kan worden voor het verzamelen van informatie om hypothesen te formuleren.

Om in complexe domeinen een meer gedetailleerde beschrijving van het gedrag van de lerende mogelijk te maken, hebben Van Joolingen en De Jong (1997) het SDDS model uitgebreid. Zij hebben de structuur van de hypotheseruimte en de experimenteeruimte in meer detail uitgewerkt en zijn met mechanismen die de zoektocht van de lerende in deze ruimtes beschrijft, gekomen: de zoektocht van de lerende in de hypotheseruimte, evenals een representatie van de actuele staat van kennis van de lerende. Hun uitbreiding heeft geleid tot de volgende subruimtes: de universele hypotheseruimte, de hypotheseruimte van de lerende, de effectieve zoekruimte van de lerende, de ruimte met juiste hypothesen en het conceptuele doelmodel (zie Figuur 1). Volgens hen bevat de universele hypotheseruimte alle mogelijke hypothesen over een bepaald domein, onafhankelijk van het feit of ze juist zijn of niet. De hypotheseruimte van de lerende wordt gezien als een subruimte van de

¹ De begeleiders van dit onderzoek waren Dr. P. Wilhelm en Dr. A. W. Lazonder. Beiden zijn werkzaam op de afdeling Instructietechnologie aan de Faculteit Gedragwetenschappen van de Universiteit Twente.

universele hypotheseruimte en bevat de kennis van de lerende over de variabelen en de relaties in het domein, tevens onafhankelijk van het feit of ze juist zijn of niet. Dit gedeelte van de universele hypotheseruimte kan de lerende direct, zonder additionele informatie, doorzoeken. De effectieve zoekruimte van de lerende, een subruimte van de hypotheseruimte van de lerende, bevat de hypothesen die de lerende zinvol acht om te testen. Deze ruimte representeert de gedachten en ideeën en actuele staat van kennis van de lerende (en is daarom niet vaststaand maar dynamisch). De ruimte met juiste hypothesen (die in Figuur 1 niet is weergegeven vanwege duidelijkheidsredenen) bevat alle juiste hypothesen die het domein beschrijven. Het conceptuele doelmodel is een subset van de ruimte met juiste hypothesen en bestaat uit de kennis die ontdekt moet worden. Aan het eind van het ontdekkingsproces wordt van de lerende verwacht dat hij of zij een verzameling relaties heeft gevonden die equivalent zijn aan dit doelmodel.



Figuur 1. Het uitgebreide SDDS model (Van Joolingen & De Jong, 1997).

Naast het ontwerp van het uitgebreide SDDS model, verklaren Van Joolingen en De Jong (1997) dat de hypotheseruimte in feite bestaat uit twee ruimtes, een variabelenruimte en een relatieruimte (omdat een hypothese over een relatie tussen twee of meer variabelen gaat). Volgens hen moet een lerende, wanneer de hypotheseruimte doorzocht wordt, zowel de variabelenruimte (om variabelen om een hypothese over te formuleren, te identificeren) als de relatieruimte (om een relatie die tussen twee of meer variabelen bestaat, te selecteren) doorzoeken.

Volgens Gijlers (2005), die het model heeft bewerkt en de naam van de 'effectieve zoekruimte van de lerende' heeft gewijzigd in de 'domeinruimte van de lerende', overlappen de domeinruimte van de lerende en het conceptuele doelmodel niet altijd. Bovendien zijn ze niet altijd van gelijke grootte en valt het conceptuele doelmodel niet noodzakelijkerwijs (volledig) binnen de hypotheseruimte van de lerende. Dit betekent dat de lerende niet altijd over (voldoende) relevante domeinkennis beschikt en hij of zij soms zijn of haar hypotheseruimte moet vergroten of alle relevante variabelen en relaties moet kennen, voordat hij of zij relevante hypothesen kan formuleren.

Zoals eerder vermeld, gaat het bij onderzoekend leren over het ontdekken van regels (Wilhelm & Beishuizen, 2003). Het meeste redeneren in het onderzoekend leren proces is daarom inductief. Echter, deductief redeneren speelt ook een rol bij onderzoekend leren omdat geformuleerde hypothesen getest kunnen worden op validiteit, wat een deductief redeneerproces is (Wilhelm & Beishuizen, 2003). De meeste studies over de invloed van domeineigenschappen op redeneerprestatie gebruiken deductieve redeneertaken. Daarom zal, om meer duidelijkheid te verschaffen over de specifieke domeineigenschappen die van invloed zijn op de uitkomsten van onderzoekend leertaken, eerst nader gekeken worden naar deze deductieve redeneerstudies en hun bevindingen.

Initieel psychologisch onderzoek met deductieve redeneertaken begon met het bestuderen van syllogismen (Benjafield, 1997). Een syllogisme bestaat uit twee premissen en een conclusie en binnen een syllogisme specificeert ieder van de twee premissen een relatie tussen twee categorieën, bijvoorbeeld: 'alle koeien zijn dieren' en 'alle koeien geven melk' en daarom, 'sommige dieren geven

melk'. Volgens Henle (1962) kunnen premissen vaak verschillend geïnterpreteerd worden, wat kan leiden tot fouten in het redeneerproces, zoals gevonden door Ceraso en Provitera (1971). Zij gaven aan dat de manier waarop mensen redeneren, afhangt van de wijze waarop de premissen worden geïnterpreteerd of gecodeerd. Verder ontdekten ze dat mensen soms onlogische conclusies accepteren in abstracte syllogismen, omdat ze het moeilijk vinden om de premissen correct te interpreteren. Revlis (1975) ontdekte dat het gebruik van betekenisvol (concreet) materiaal in de premissen deze minder ambigu kan maken, waardoor misinterpretaties voorkomen worden. Echter, inhoudelijk realistische syllogismen kunnen ook leiden tot een overtuigingsbeoordeling. Zo vonden Evans, Barston en Pollard (1983) dat mensen de correctheid van een conclusie van syllogismen soms beoordelen op basis van overtuiging in plaats van logische validiteit.

Overige deductieve redeneertaken die op grote schaal onderwerp van onderzoek zijn geweest zijn Wason's selectietaak (Wason, 1966) en Wason's THOG-taak (Wason, 1977). Wason's selectietaak lijkt op het eerste gezicht eenvoudig, maar blijkt voor de meesten moeilijk op te lossen. De standaard (abstracte) versie van de taak is als volgt: vier kaarten worden getoond, één met een klinker, één met een medeklinker, één met een even cijfer en één met een oneven cijfer. De kaarten worden vergezeld met een 'als ... dan' voorwaardelijke zin, zoals: 'als er een klinker op de ene kant van de kaart staat, staat er een even cijfer op de andere kant'. De taak is om de kaarten te selecteren die omgedraaid moeten worden om te bepalen of de voorwaardelijke zin goed of fout is. In een experiment met deze taak, draaiden 59 van de 128 deelnemers de 'E' en de '4' om en 42 deelnemers enkel de 'E' (Johnson-Laird & Wason, 1970). De meeste deelnemers concludeerden dus dat ze de 'E' en de '4', of enkel de 'E' om moesten draaien, terwijl het enige juiste antwoord is te concluderen dat het omdraaien van de 'E' en de '5' nodig is om te bepalen of de voorwaardelijke zin goed of fout is. Dit initiële experiment met de selectietaak heeft aangetoond dat mensen moeite hebben met het op een juiste wijze oplossen ervan. Maar wat gebeurt er wanneer de inhoud meer realistisch wordt gemaakt en de abstracte letters en cijfers worden vervangen door concrete alledaagse woorden?

In een experiment uitgevoerd door Wason en Shapiro (1971) werd bij de ene helft van de deelnemers de standaard abstracte versie van de taak afgenomen, bij de andere helft een nieuwe versie met een concrete inhoud. In deze concrete versie waren de letters vervangen door de woorden 'Manchester' en 'Leeds' en de cijfers door de woorden 'trein' en 'auto'. De voorwaardelijke zin in deze taak was: 'elke keer als ik naar Manchester ga, reis ik per trein'. Een aanzienlijk faciliterend effect werd waargenomen; 10 van de 16 deelnemers maakten de juiste selectie bij de taak met de concrete inhoud, terwijl slechts 2 van de 16 deelnemers de juiste selectie maakten bij de taak met de abstracte inhoud. Een zelfs groter faciliterend effect werd gerapporteerd door Johnson-Laird, Legrenzi en Legrenzi (1972). Zij gebruikten ook een abstracte letter-cijfer versie van de selectietaak maar vergeleken de prestatie op deze versie met die van een versie waarin deelnemers de rol van postsorteerders moesten spelen. Aan deze deelnemers werd de volgende voorwaardelijke zin meegedeeld: 'wanneer een envelop dichtgeplakt is, zit er een postzegel op van 50 cent'. De taak was om die enveloppen te selecteren die nader onderzocht moesten worden om er zodoende achter te komen of de regel overschreden werd. Dit gerapporteerde faciliterende effect werd in twijfel getrokken door Manktelow en Evans (1979). Volgens hen bestond het effect doordat het de herinneringen van mensen over het correcte tegengestelde voorbeeld opriep, in plaats van dat het logisch redeneren promootte. Griggs en Cox (1982) rapporteerden een studie waarin ze dit idee van het oproepen van herinneringen testten. Zij namen de postversie van de selectietaak over en testten deze op deelnemers uit Florida, VS. Deze deelnemers waren niet bekend met een regeling voor verschillende posttarieven voor open en dichtgeplakte enveloppen (de deelnemers uit het onderzoek van Johnson-Laird e.a. waren dit wel). Zoals Griggs en Cox hadden voorspeld, leidde de prestatie van de deelnemers uit Florida niet tot het faciliterende effect, zoals gerapporteerd door Johnson-Laird e.a. Cheng en Holyoak (1985) herhaalden het onderzoek ook met deelnemers die niet bekend waren met een dergelijke regeling voor posttarieven. Zij ontdekten dat prestatiefacilitering mogelijk was wanneer een reden voor de regel werd gegeven. Vanwege de bevindingen in dit onderzoek, geeft Manktelow (1999) aan dat culturele ervaring met een regel of een duidelijke verklaring van een reden voor een regel, die ervoor zorgt dat bij mensen de kennis die ze hebben over regels opgeroepen wordt, ervoor zorgt dat mensen correct redeneren. Cheng en Holyoak (1985) gebruikten een versie van de selectietaak waarin de alledaagse inhoud achterwege werd gelaten, maar die ingebed was in een regelings- en overtredingscontext. Ze gebruikten de volgende regel: 'wanneer iemand actie 'A'

uitvoert, moet men eerst aan preconditionie 'B' voldoen' en vonden een significant faciliterend effect. Deze bevinding suggereert dat afgezien van de concreetheid van de afbeeldingen op de kaarten, ook de context waarin een probleem geplaatst wordt een faciliterende rol speelt bij redeneren met de selectietaak. Een mogelijke reden voor deze bevinding wordt aangedragen door Cosmides (1989). Volgens haar zijn mensen uitgerust met aangeboren structuren voor regels, zogenaamde 'Darwinistische algoritmen', nodig om te overleven. Zij geeft aan dat mensen vanwege deze structuren voor regels in staat zijn om mensen te detecteren die 'het voordeel pakken, zonder daarvoor de kosten te betalen' of, in andere woorden, 'de regel overtreden'.

Resultaten van studies met Wason's THOG-taak pleitten ook voor een door concrete inhoud gecreëerd faciliterend effect met betrekking tot redeneerprestatie. In een experiment waarbij deelnemers geconfronteerd werden met de originele geometrische versie van de THOG-taak, gaf slechts 35 % van de deelnemers het juiste antwoord (Wason & Brooks, 1979). Girotto en Legrenzi (1989) bedden deze originele geometrische versie in een realistische inhoud in en vonden een aanzienlijke toename in het aantal deelnemers dat het juiste antwoord gaf (89 %).

Over het algemeen kan gesteld worden dat het realistischer maken van deductieve redeneertaken, door het vervangen van de abstracte door een concrete inhoud, een faciliterend effect heeft. Het zou kunnen dat een meer realistische inhoud, visualisering van de taak mogelijk maakt en mensen zodoende kunnen vertrouwen op voorkennis, gedachtes en opvattingen. Echter, dit kan, vooral wanneer de taak niet overeenkomt met iemands voorkennis en overtuigingen, ook leiden tot tegenstrijdigheden, het logisch redeneren belemmeren en daardoor leiden tot een verminderde prestatie (Evans, Barston & Pollard, 1983). Verder heeft onderzoek aangetoond dat wanneer een regel gepresenteerd wordt, ervaring met die regel of het begrijpen van die regel (eventueel door het geven van een reden voor de regel) bijdraagt aan de redeneerprestatie en dat, afgezien van een concrete inhoud, ook de context van het probleem een rol speelt bij de redeneerprestatie (Cheng & Holyoak, 1985). Geconcludeerd kan worden dat het bij faciliterende effecten bij inhoudelijk concrete redeneerproblemen dus niet alleen gaat om bekendheid met de objecten of afbeeldingen, maar dat bekendheid met de regel of de context van een probleem ook een rol speelt. Deze regel of context van het probleem zorgt mogelijk voor connecties en relaties tussen de objecten of afbeeldingen van het probleem en wordt gecreëerd wanneer de inhoud concreet gemaakt wordt. Zo bestond er in de originele abstracte versie van de selectietaak (Johnson-Laird & Wason, 1970) geen connectie of relatie tussen de afgebeelde letters en cijfers op de kaarten, maar werd deze connectie of relatie wel gecreëerd toen Wason en Shapiro (1971) de taak concreet maakten door de letters en cijfers te vervangen door de concrete objecten 'Manchester', 'Leeds', 'trein' en 'auto' en het onderwerp van het probleem (reizen) aangaven in de voorwaardelijke zin.

Zoals vermeld gebruiken de meeste studies naar de invloed van domeineigenschappen op de redeneerprestatie, deductieve redeneertaken. Een uitzondering hierop is een onderzoek van Bruner Goodnow en Austin (1956), waarin de rol van inhoudseffecten in een inductieve redeneertaak onderzocht werd. Zij gebruikten een concrete en een abstracte versie van een leertaak over concepten, waarin deelnemers een regel, die bepaalde kenmerken van het concept beschreef, moesten ontdekken. Om deze regel te ontdekken konden deelnemers concepten selecteren en de experimenteleider vragen of het geselecteerde concept aan de regel voldeed. Bruner e.a. vergeleken de taakprestatie van de twee groepen en ontdekten dat hoewel lerenden in beide groepen de regel ontdekten, de lerenden in de concrete conditie meer concepten hadden geselecteerd en meer incorrecte hypothesen hadden geformuleerd dan lerenden in de abstracte conditie.

Domeinverschillen bij onderzoekend leertaken leiden, zoals eerder vermeld, tot variaties in leerprestatie en leerprocessen. Eén van de onderzoeken waarin dit effect van het domein, waarin een onderzoekend leren taak is ingebed, op het leergedrag en -uitkomst werd onderzocht, is uitgevoerd door Kuhn, Garcia-Mila, Zohar en Anderson (1995). In hun studie voerden deelnemers vier onderzoekend leertaken uit; twee in een fysiek domein (hoe de snelheid van een modelboot en de snelheid van een raceauto wordt beïnvloed door kenmerken van de boot en de raceauto) en twee in een sociaal domein (hoe schoolprestatie wordt beïnvloed door kenmerken van de leersituatie en hoe de populariteit van televisieprogramma's voor kinderen wordt beïnvloed door kenmerken van het programma). De taakstructuur was als volgt: iedere taak bestond uit vijf onafhankelijke variabelen met twee of drie niveaus en een afhankelijke variabele met vier niveaus. In ieder van de vier taken bestond er interactie tussen twee onafhankelijke variabelen, had een variabele met drie niveaus een kromlijnig

effect (één niveau had een hoofdeffect, de twee anderen geen effect op de uitkomst) en twee variabelen waren irrelevant. Taak van de deelnemers was om te variëren in de verschillende niveaus van de onafhankelijke variabelen, om zodoende te ontdekken of en hoe ze de afhankelijke variabele beïnvloedden. De vier taken waren isomorf (de relaties tussen de onafhankelijke variabelen en de afhankelijke variabele waren in iedere taak gelijk). Voor en na de taakuitvoering werden deelnemers gevraagd naar hun ideeën over de effecten van de onafhankelijke variabelen. Kuhn e.a. ontdekten dat deelnemers het lastiger vonden om hun causale theorieën te weerleggen in het sociale domein dan in het fysieke domein. Daarnaast ontdekten ze dat deelnemers meer moeilijkheden hadden bij het ontdekken van het interactie-effect in de autotaak en meer moeilijkheden hadden bij het ontdekken van het kromlijnige effect in de boottaak, terwijl er geen verschillen in moeilijkheid tot ontdekken van deze effecten werd gevonden in de twee taken uit het sociale domein.

Wilhelm en Beishuizen (2003) gebruikten de taakstructuur van Kuhn e.a. (1995) om het effect van concrete en abstracte taakinhoud op de leerprestatie en het leerproces te onderzoeken. Zij legden vijftig universiteitsstudenten een onderzoekend leertaak voor. Aan één groep werd een abstracte versie van de taak voorgelegd. In deze abstracte versie bestonden de vijf onafhankelijke variabelen uit kaarten met verschillende geometrische vormen. De niveaus van iedere onafhankelijke variabele bestonden uit verschillende kleuren van deze vormen. Deelnemers werden geïnstrueerd om rijen te maken met de verschillende niveaus van de vijf vormen. Voor iedere rij ontvingen ze een bepaald aantal punten. De taak was om verschillende rijen te maken en er zodoende achter te komen hoe de vormen in verschillende kleuren het aantal punten, verkregen voor een bepaalde rij, beïnvloedde. Aan de andere groep werd een isomorfe versie van de taak, ingebed in een concrete (bekende) context, voorgelegd. In deze concrete versie werd een alledaags verhaal gepresenteerd, waarin een fictief persoon voor een probleem stond. Deelnemers moesten de effecten van vijf onafhankelijke variabelen (zoals type fiets) op het aantal minuten dat de persoon te laat op school aankwam, ontdekken. In deze versie van de taak, bestonden de niveaus uit subcategorieën van de onafhankelijke variabelen (bijvoorbeeld: bij de onafhankelijke variabele 'fiets', waren de niveaus 'normale fiets' en 'racefiets'). De bedoeling was ook hier om verschillende rijen te maken (met één niveau van iedere onafhankelijke variabele), om er zodoende achter te komen hoe de onafhankelijke variabelen het aantal minuten dat de persoon te laat op school arriveerde, beïnvloedde. Wilhelm en Beishuizen waren geïnteresseerd in het effect van de taakinhoud op de leerprestatie en het leerproces. Om de leerprestatie te bepalen berekenden ze een begripsscore voor iedere deelnemer. Deze score was gebaseerd op de kennis die een deelnemer zou kunnen genereren over effecten van iedere onafhankelijke variabele. Ze ontdekten dat deelnemers die de abstracte taak hadden uitgevoerd, significant lager scoorden dan deelnemers die de concrete taak hadden uitgevoerd. Daarnaast berekenden ze subscores voor de verschillende effecten van de variabelen (zoals het interactie-effect), om na te gaan of de verschillen in begripsscore toe te schrijven waren aan verschillen in het ontdekken van een bepaald effect. Ze ontdekten dat deelnemers die de concrete taak hadden uitgevoerd significant hoger scoorden op de subscores voor het interactie-effect en de irrelevante effecten, maar dat er geen verschillen tussen de groepen aanwezig waren met betrekking tot subscore voor het kromlijnige effect. Verder ontdekten ze dat deelnemers op de concrete taak, significant meer hypothesen hadden gegenereerd, meer gerichte experimentplannen gemaakt hadden en meer inductieve experimentplannen gemaakt hadden, dan deelnemers op de abstracte taak. Door het inspecteren van de verzameling experimenten, uitgevoerd door iedere deelnemer, op bewijs voor het bestaan van het interactie-effect, ontdekten ze dat 50 % van de deelnemers die de concrete taak hadden uitgevoerd en 58 % van de deelnemers die de abstracte taak hadden uitgevoerd, experimenten hadden gedaan waarin dit bewijs aan het licht kwam, terwijl 33 % van de studenten in de concrete groep en slechts 15 % van de studenten in de abstracte groep melding hadden gemaakt van het bestaan van dit type effect.

Uit de resultaten van het onderzoek van Wilhelm en Beishuizen (2003) kan afgeleid worden dat het inbedden van een onderzoekend leertaak in een concrete inhoud, de prestatie faciliteert. Dit komt mogelijk door het feit dat de hypotheseruimte van de lerende in een concrete taak initieel groter is dan die in een abstracte taak. Door het bestaan van voorkennis bij deelnemers in een taak met een bekend domein kunnen relevante hypothesen geformuleerd worden en met deze geformuleerde hypothesen kan direct worden begonnen met het uitvoeren van gerichte experimenten. Omdat deelnemers in een abstracte taak niet kunnen beschikken over voorkennis met betrekking tot de taakinhoud, dienen zij te beginnen met het uitvoeren van (ongerichte) experimenten, totdat ze, op basis

van de resultaten van die experimenten, voldoende domeinkennis hebben vergaard om relevante hypothesen te kunnen formuleren. Verder kan op basis van dit onderzoek geconcludeerd worden dat concreetheid in onderzoekend leertaken leidt tot beter experimenteel gedrag. In hun studie ontdekten Wilhelm en Beishuizen ook verschillen tussen de groepen op basis van de subscores voor de verschillende effecten. Zij concludeerden dat dit mogelijk verklaard kan worden door het feit dat deelnemers in de concrete conditie konden vertrouwen op mechanismen die de gevonden effecten in de taak konden verklaren. Dit effect van verklarende mechanismen (zie Koslowski, 1996) kan geïllustreerd worden aan de hand van een voorbeeld: 'schoenen' (normale schoenen of sneakers) was één van de irrelevante variabelen in de concrete conditie. Wanneer deze irrelevantie ontdekt wordt is dat aannemelijk, aangezien de persoon in de taak niet naar school loopt, maar met de fiets gaat. Deelnemers in de abstracte conditie konden niet vertrouwen op verklarende mechanismen (die bijvoorbeeld irrelevantie van de variabele 'cirkel' aannemelijk maken) en hadden daarom mogelijk meer moeite de irrelevante effecten te ontdekken. Verder verklaren Wilhelm en Beishuizen dat, omdat het genereren van een alternatieve hypothese of verklarend mechanisme eenvoudiger is in een concreet dan in een abstract domein, deelnemers in de concrete conditie minder moeilijkheden hadden met het ontdekken van het interactie-effect. Echter, zoals geïllustreerd door het onderzoek van Kuhn e.a. (1995), kunnen verschillen in leerprestatie tussen concrete domeinen ook voorkomen. Zij ontdekten dat deelnemers het lastiger vonden om hun causale theorieën te weerleggen in het sociale domein, dan in het fysieke domein. Een mogelijke verklaring hiervoor is dat sociale domeinen, in tegenstelling tot fysieke domeinen, meer ruimte laten voor persoonlijke opvattingen, gedachten en aannames van mensen en dat deelnemers, wanneer ze hun causale theorieën weerlegden, dus ook hun persoonlijke opvattingen, gedachten en aannames aan de kant zetten.

Geconcludeerd kan worden dat, analoog aan bevindingen uit onderzoek met deductieve redeneertaken zoals de selectietaak, concreetheid in onderzoekend leertaken de prestatie faciliteert. Echter, dit effect wordt niet vanzelfsprekend veroorzaakt door concreetheid van de variabelen of objecten in het taakdomein. Andere factoren, zoals bekendheid met een regel of de context van een probleem (in redeneertaken) en domeinkenmerken die helpen bij het genereren van (alternatieve) hypothesen en het interpreteren van bevindingen (in onderzoekend leertaken), spelen ook een rol bij de prestatiefacilitering. Over het algemeen kan gesteld worden dat lerenden geholpen kunnen worden bij het uitvoeren van onderzoekend leertaken door het taakdomein concreet te maken of door bij lerenden (voor)kennis betreffende het domein te activeren of aan te bieden. Maar wat is het dat er voor zorgt dat een taakdomein (meer) concreet wordt? En wat voor een soort (voor)kennis betreffende het domein is verstandig om te activeren of aan te bieden.

Aangezien Van Joolingen en De Jong (1997) verklaarden dat de hypotheseruimte bestaat uit twee subruimtes: een variabelenruimte en een relatieruimte, lijkt het erop dat concreetheid in onderzoekend leertaken zowel gaat over de beschikking over voorkennis over variabelen (die zorgt voor een initieel gevulde variabelenruimte) als over de beschikking over voorkennis over relaties (die zorgt voor een initieel gevulde relatieruimte). Dit roept de vraag op welke soort concreetheid (en daarmee de beschikking over voorkennis) de prestatie het meest faciliteert: concreetheid van de variabelen, concreetheid van de relaties of wellicht is het de combinatie van concreetheid van variabelen en concreetheid van relaties die zorgt voor het faciliterende effect? Aangezien Wilhelm en Beishuizen (2003) verklaarden dat onderzoekend leren gaat over het ontdekken van regels (en daarmee over het ontdekken van relaties tussen de variabelen) ligt het voor de hand om aan te nemen dat het niet alleen de concreetheid van de variabelen is die faciliteert, maar ook de concreetheid van de relaties tussen en binnen die variabelen. Wanneer dieper ingegaan wordt op de functie die een initieel gevulde hypotheseruimte speelt met betrekking tot de facilitering van prestatie, kan verondersteld worden dat het waarschijnlijk de combinatie van beide types concreetheid is die hierbij een sleutelrol speelt. Immers, een initieel gevulde hypotheseruimte faciliteert, omdat het de lerende toegang verschaft tot het genereren van relevante hypothesen betreffende het taakdomein en hij of zij zodoende direct kan beginnen met het uitvoeren van gerichte experimenten. Aangezien een hypothese in onderzoekend leren een bewering is over de relatie tussen twee of meer onafhankelijke variabelen en de afhankelijke variabele (Van Joolingen en De Jong, 1997), is zowel kennis over de variabelen als kennis over de relaties in het domein nodig om een hypothese te formuleren.

Om meer inzicht te verkrijgen over de mogelijk grote rol van de combinatie van concreetheid van variabelen en concreetheid van relaties tussen en binnen deze variabelen, als zijnde faciliterende factor

in onderzoekend leertaken, zal de volgende onderzoeksvraag onderzocht worden: ‘wat is het effect van concreetheid van variabelen (het bestaan van een initieel gevulde variabelenruimte), concreetheid van relaties tussen en binnen variabelen (het bestaan van een initieel gevulde relatieruimte) en concreetheid van zowel variabelen als de relaties ertussen en erbinnen (het bestaan van een initieel gevulde variabelenruimte én een initieel gevulde relatieruimte) op de leerprestatie en het experimenteelgedrag van lerenden op onderzoekend leertaken?’ (waarbij experimenteelgedrag gebaseerd is op de onderzoekend leren processen ‘hypotheseformatie’ en ‘experimentontwerp’ en tot uitdrukking komt in het aantal geformuleerde hypothesen, het aandeel geformuleerde specifieke hypothesen en het aantal uitgevoerde experimenten). Dit zal onderzocht worden door te manipuleren in de aanwezigheid van de twee subruimtes van de hypotheseruimte (de variabelen- en de relatieruimte). Om dit mogelijk te maken zijn drie isomorfe onderzoekend leertaken ontworpen.

De eerste taak (conditie ‘concreet rijk’) bevat zowel concrete variabelen als concrete relaties tussen en binnen deze variabelen en zal daarmee, bij de lerende, ruimte laten voor de aanwezigheid van zowel voorkennis over de variabelen als voorkennis over de relaties tussen en binnen die variabelen (aanwezigheid van zowel een initieel gevulde variabelenruimte als een initieel gevulde relatieruimte).

De tweede taak (conditie ‘concreet arm’) bevat concrete variabelen, maar onbekende relaties tussen en binnen deze variabelen, zodat bij de lerende ruimte is voor de aanwezigheid van voorkennis over de variabelen (aanwezigheid van een initieel gevulde variabelenruimte), maar niet voor de aanwezigheid van voorkennis over de relaties tussen en binnen die variabelen (afwezigheid van een initieel gevulde relatieruimte). Een voorbeeld van twee concrete variabelen met onbekende relaties ertussen en erbinnen is ‘bloemen’ (met de niveaus ‘tulp’ en ‘narcis’) en kleding (met de niveaus ‘broek’ en ‘trui’). Hoewel lerenden hoogst waarschijnlijk over voorkennis beschikken over deze variabelen en hun niveaus, ontbreekt voorkennis over de relaties tussen bijvoorbeeld ‘bloemen’ en ‘kleding’, ‘tulp’ en ‘narcis’ of ‘tulp’ en ‘broek’.

De derde taak (conditie ‘abstract’) bevat abstracte variabelen met onbekende relaties tussen en binnen deze variabelen, zodat bij de lerende geen ruimte is voor aanwezigheid van voorkennis over de variabelen en de relaties ertussen en erbinnen (afwezigheid van zowel een initieel gevulde variabelenruimte als een initieel gevulde relatieruimte).

Als een replica van het onderzoek van Wilhelm en Beishuizen (2003) en om het faciliterende effect van zowel concreetheid van variabelen als concreetheid van relaties ertussen en erbinnen (aanwezigheid van zowel een initieel gevulde variabelenruimte als een initieel gevulde relatieruimte) te onderzoeken, zal de gemiddelde leerprestatie en het experimenteelgedrag van lerenden in de conditie ‘concreet rijk’ en de conditie ‘abstract’ met elkaar vergeleken worden. Om het faciliterende effect van concreetheid van relaties tussen en binnen variabelen (aanwezigheid van een initieel gevulde relatieruimte) te onderzoeken, zal de gemiddelde leerprestatie en het experimenteelgedrag van lerenden in de conditie ‘concreet rijk’ en de conditie ‘concreet arm’ met elkaar vergeleken worden. Daarnaast zal, om het faciliterende effect van concreetheid van variabelen (aanwezigheid van een initieel gevulde variabelenruimte) te onderzoeken, de gemiddelde leerprestatie en het experimenteelgedrag van lerenden in de conditie ‘concreet arm’ en de conditie ‘abstract’ met elkaar vergeleken worden.

Met betrekking tot de eerste vergelijking (conditie ‘concreet rijk’ met conditie ‘abstract’) wordt verwacht dat, gelijk aan de bevindingen in het onderzoek van Wilhelm en Beishuizen (2003), lerenden in de conditie ‘concreet rijk’ beter zullen presteren dan lerenden in de conditie ‘abstract’. Verder wordt verwacht dat lerenden in de conditie ‘concreet rijk’ meer efficiënt experimenteelgedrag zullen vertonen (meer hypothesen zullen formuleren, een hoger aandeel specifieke hypothesen zullen formuleren en minder experimenten zullen uitvoeren), aangezien lerenden in deze conditie kunnen profiteren van een initieel gevulde hypotheseruimte door de beschikbaarheid van voorkennis over de (concrete) variabelen en hun relaties en ze daardoor in staat zijn om relevante hypothesen te formuleren voordat ze met experimenteren beginnen. Ze zullen daardoor waarschijnlijk meer en specifiekere hypothesen formuleren en zullen waarschijnlijk minder experimenten nodig hebben om de taak af te maken. De hypotheseruimte van de lerenden in de ‘abstracte’ conditie zal initieel leeg zijn, aangezien zij niet over voorkennis beschikken over de (abstracte) variabelen en hun relaties. Lerenden in deze conditie zullen bij het aanvangen van de taak niet kunnen beschikken over relevante hypothesen en zodoende moeten beginnen met het uitvoeren van (ongerichte) experimenten. Verder

zullen lerenden in de ‘concreet rijk’ conditie mogelijk profiteren van het feit dat ze (op basis van hun voorkennis) mogelijk kunnen vertrouwen op verklarende mechanismen (die ze helpt bij het interpreteren van de bevindingen en het formuleren van alternatieve hypothesen).

Met betrekking tot de tweede vergelijking (conditie ‘concreet rijk’ met conditie ‘concreet arm’), wordt verwacht dat lerenden in de conditie ‘concreet rijk’ beter zullen presteren dan lerenden in de conditie ‘concreet arm’ en dat ze efficiënter experimenteergedrag zullen vertonen, aangezien lerenden in deze conditie kunnen profiteren van een initieel gevulde hypotheseruimte door de beschikbaarheid van voorkennis over de (concrete) variabelen en hun relaties. Lerenden in de conditie ‘concreet arm’ beschikken weliswaar over voorkennis over de (concrete) variabelen (en daarmee over een initieel gevulde variabelenruimte), ze ontberen voorkennis over de relaties tussen en binnen deze variabelen en zijn daardoor niet in staat om relevante hypothesen te formuleren voordat ze beginnen met experimenteren.

Met betrekking tot de derde vergelijking (conditie ‘concreet arm’ met conditie ‘abstract’), wordt verwacht dat er geen verschillen tussen deze condities gevonden zullen worden met betrekking tot leerprestatie en experimenteergedrag; lerenden in de ‘concreet arm’ conditie beschikken weliswaar over een initieel gevulde variabelenruimte, ze kunnen er niet van profiteren, aangezien hun relatieruimte initieel leeg is en ze daardoor niet in staat zijn om relevante hypothesen te formuleren voordat ze met experimenteren beginnen.

Method

Deelnemers

Zevenenvijftig studenten (31 vrouwen, 26 mannen, gemiddelde leeftijd: 19.6 jaar, SD.: 1.8) van de Faculteit Gedragwetenschappen van de Universiteit van Twente hebben deelgenomen aan het onderzoek. Met deelname aan het onderzoek verdienden ze twee proefpersoonpunten. Deelnemers zijn aselect toegewezen aan één van de in totaal drie experimentele condities, zodat er uiteindelijk 19 deelnemers aan elke conditie zijn toegewezen.

Materialen

Onderzoekend leertaken

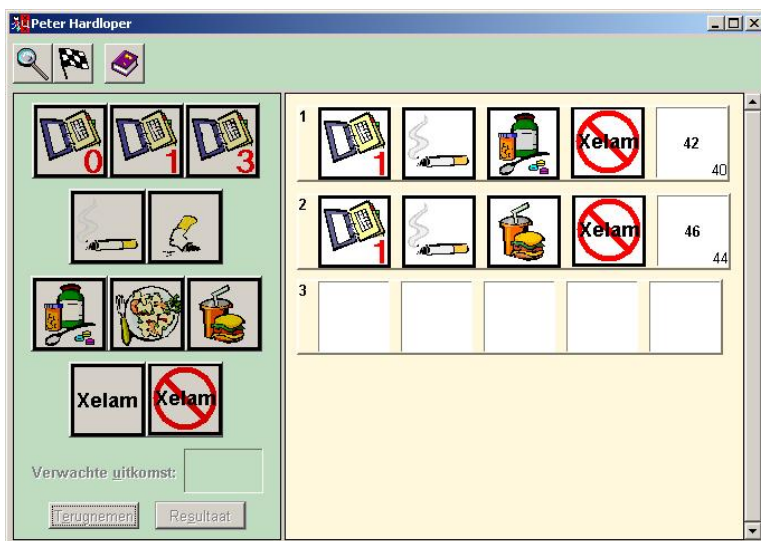
Drie isomorfe onderzoekend leertaken zijn ontworpen: een concrete taak met een rijke interne relatiestructuur (conditie ‘concreet rijk’), een concrete taak met een arme interne relatiestructuur (conditie ‘concreet arm’) en een abstracte taak (conditie ‘abstract’). De taken bestaan uit vier onafhankelijke variabelen die de deelnemers konden manipuleren en uit één afhankelijke variabele, de uitkomst van het geheel van gemanipuleerde onafhankelijke variabelen. Twee van de onafhankelijke variabelen hebben twee niveaus en de overige twee onafhankelijke variabelen hebben drie niveaus.

De concrete taak met een rijke interne relatiestructuur is opgebouwd rond een verhaal waarin een fictief persoon zich voorbereidt op een hardlooptwedstrijd. In deze taak bestaan de onafhankelijke variabelen uit verschillende factoren (trainen, roken, voeding en een bepaald middel genaamd ‘xelam’) waarvan een combinatie de afhankelijke variabele (de tijd die de persoon doet over het lopen van de tien kilometer) beïnvloed. De niveaus van de onafhankelijke variabelen bestaan bij de factor ‘trainen’ uit: ‘niet trainen’, ‘één keer per week trainen’ en ‘drie keer per week trainen’, bij de factor ‘voeding’ uit: ‘eten van koolhydraatrijke sportvoeding’, ‘eten wat de pot schaft’ en ‘eten van junkfood’, bij de factor ‘roken’ uit: ‘wel stoppen met roken’ en ‘niet stoppen met roken’ en bij de factor ‘xelam’ uit: ‘wel gebruik van xelam’ en ‘geen gebruik van xelam’.

De onafhankelijke variabelen van de concrete taak met een arme interne relatiestructuur bestaan uit verschillende concrete factoren (muziekinstrumenten, bestek, bloemen en kleding). De niveaus van de onafhankelijke variabelen bestaan uit bepaalde soorten van deze factoren (de twee niveaus van de factor bloemen zijn bijvoorbeeld: ‘tulpe’ en ‘narcis’). De afhankelijke variabele bestaat uit een puntentotaal, dat verkregen wordt voor een bepaalde combinatie van de niveaus van de onafhankelijke variabelen.

De onafhankelijke variabelen van de abstracte taak bestaan uit abstracte vormen (een driehoek, een vierkant, een cirkel en een vraagteken). De niveaus van de onafhankelijke variabelen bestaan uit verschillende kleuren (voor de vorm ‘driehoek’ kon bijvoorbeeld gekozen worden voor de kleur ‘bruin’ of ‘blauw’). De afhankelijke variabele bestaat uit een puntentotaal, dat verkregen wordt voor een bepaalde combinatie van de niveaus van de onafhankelijke variabelen.

De onderzoekend leertaken zijn uitgevoerd in FILE, een gecomputeriseerde leeromgeving waarin het mogelijk is verschillende leeromgevingen te programmeren (zie Hulshof, Wilhelm, Beishuizen & Van Rijn, 2005). In deze simulatieomgeving is het mogelijk om experimenten uit te voeren door aanpassing van de onafhankelijke factoren en het effect van deze aanpassing waar te nemen in de afhankelijke variabele (de uitkomst). De werking van het programma zal uitgelegd worden aan de hand van de concrete taak met een rijke interne relatiestructuur (de werking van de overige twee taken is identiek), de zogenaamde ‘hardloop taak’. De interface van deze taak in het programma is te zien in Figuur 2.



Figuur 2. Interface van de ‘hardloop taak’ in FILE.

Om de opdrachtomschrijving van de taak te lezen kan in de omgeving op het pictogram van het ‘boekje’ geklikt worden. Vervolgens kan gestart worden met het uitvoeren van experimenten door, aan de linkerkant van het scherm, van iedere variabele een gewenst niveau te selecteren, een ‘verwachte uitkomst’ te selecteren en op de button ‘resultaat’ te klikken. Uitgevoerde experimenten verschijnen vervolgens aan de rechterkant van het scherm, waarbij bij ieder specifiek experiment het gekozen niveau van iedere variabele, de ingegeven verwachte uitkomst en de werkelijke uitkomst (de waarde van de afhankelijke variabele, in dit geval de tijd die de fictieve persoon uit de opdrachtomschrijving erover doet om tien kilometer hard te lopen) weergegeven worden. In Figuur 2 zijn twee experimenten uitgevoerd. Het eerste experiment bestaat hierbij uit de niveaus: ‘één keer per week trainen’, ‘blijven roken tijdens de voorbereiding’, ‘eten van koolhydraatrijke sportvoeding’ en ‘geen gebruik van xelam’. De ingegeven verwachte uitkomst in dit experiment was 40 minuten en de werkelijke uitkomst 42 minuten. Eenmaal gedane experimenten blijven staan aan de rechterkant van het scherm, zodat vergelijkingen tussen experimenten constant mogelijk blijft. Er kunnen oneindig veel experimenten uitgevoerd worden. Wanneer het aantal reeds uitgevoerde experimenten te groot wordt om in het scherm te tonen, verschijnt een scrollbar waarmee gescrolled kan worden tussen de uitgevoerde experimenten. Vergelijken van bepaalde gewenste experimenten is tevens mogelijk door deze gewenste experimenten te selecteren (door erop te klikken) en vervolgens op het pictogram van het ‘loepje’ te klikken. Geselecteerde experimenten worden dan onder elkaar in een apart scherm getoond. Verder slaat het programma alle handelingen van de gebruiker met bijbehorende tijdsvermelding op in een logfile.

Kleurenblindheidtest

Aangezien de niveaus van de variabelen in de abstracte taak onderscheiden werden door middel van kleuren, kan eventueel aanwezige kleurenblindheid bij deelnemers zorgen voor een ongewenste negatieve beïnvloeding van de prestatie. Om dit te ondervangen is voorafgaand aan het experiment een digitale versie van de Ishihara kleurenblindheidtest afgenomen (Ishihara, 1982). Deze test bestaat uit acht afbeeldingen, opgebouwd uit gekleurde bolletjes. Taak is om aan te geven welk cijfer er in het bolletjespatroon van elke afbeelding verborgen zit. De verborgen cijfers differentiëren voor kleurenblindheid voor de kleuren rood of groen. Van de acht voorgelegde afbeeldingen moesten er minstens zeven juist geïnterpreteerd worden om te slagen voor de test.

Prestatieformulier

Om de prestatie van de deelnemers te meten is een prestatieformulier ontwikkeld. Door middel van dit prestatieformulier kan nagegaan worden in hoeverre de individuele deelnemer het onderliggende model van de taak heeft ontdekt. Op het prestatieformulier staan de verschillende onafhankelijke variabelen uit een bepaalde taak, met daaronder voldoende ruimte om aantekeningen en ontdekkingen te vermelden. De bijbehorende instructie vraagt de deelnemer op te schrijven wat hij of zij heeft ontdekt over de invloed van de verschillende onafhankelijke variabelen op de afhankelijke variabele (de uitkomst).

Procedure

Het onderzoek is uitgevoerd in een tijdsbestek van vier weken waarbij elke deelnemer in het bijzijn van een proefleider werd getest. De digitale Ishihara kleurenblindheidtest werd als eerste afgenomen. Vervolgens werd samen met de betreffende deelnemer, aan de hand van een introductietaak, de simulatieomgeving en het doel van de taak doorgenomen. Vervolgens begonnen deelnemers met de daadwerkelijke taakuitvoer. Hiervoor kregen ze maximaal 45 minuten de tijd, wanneer ze echter het gevoel hadden dat ze alles hadden ontdekt over de relaties binnen de taak konden ze ook eerder stoppen. Ze noteerden al hun ontdekkingen over de invloed van de onafhankelijke variabelen op de afhankelijke variabele op het prestatieformulier. Deelnemers kregen 40 minuten de tijd om experimenten uit te voeren, vervolgens kregen ze nog maximaal vijf minuten de tijd om het prestatieformulier eventueel verder aan te vullen. Tijdens deze vijf minuten waren de deelnemers nog wel in de gelegenheid om gedane experimenten in de simulatieomgeving te bekijken, ze mochten echter geen nieuwe experimenten meer uitvoeren. Tijdens het werken van de deelnemer aan de taak is door de onderzoeker een formulier ingevuld waarop genoteerd kon worden wat de deelnemer precies onderzocht en wat de (eventuele) bijbehorende hypothese was. Om dit te achterhalen zijn door de onderzoeker vragen gesteld aan de desbetreffende deelnemer. Deze vragen waren: 'wat ga je onderzoeken?' en 'wat denk je dat er uit komt/het resultaat is?' en werden gesteld op het moment dat een deelnemer een verwachte uitkomst invulde en op het moment dat een deelnemer reeds uitgevoerde experimenten met elkaar vergeleek (door te scrollen of door de 'vergelijkfunctie' te selecteren met behulp van het pictogram van het vergrootglas).

Variabelen

Leerprestatie

Leerprestatie op de taak is bepaald door middel van scoring van de prestatieformulieren. Voor de eerste drie onafhankelijke variabelen van iedere taak (driehoek, vierkant en cirkel bij de abstracte taak, training, roken en voeding bij de concrete taak met een rijke interne relatiestructuur en muziekinstrumenten, bestek en bloemen bij de concrete taak met een arme interne relatiestructuur) kon een maximum van drie punten per variabele behaald worden, waarbij voor een vermelding van een effect, een juiste formulering van de richting van het effect en een juiste formulering van de omvang van het effect steeds één punt toegekend werd. De vierde onafhankelijke variabele van iedere taak (het vraagteken bij de abstracte taak, het middel 'xelam' bij de concrete taak met een rijke interne relatiestructuur en kleding bij de concrete taak met een arme interne relatiestructuur) interacteerde met

één van de andere onafhankelijke variabelen. Voor (gedeeltelijke) ontdekking van dit effect werden ook punten toegekend; één punt voor melding van het effect, één punt voor een juiste formulering van de richting van het effect en één punt voor de juiste formulering voor de juiste omvang van het effect, zodat ook hier een maximum van drie punten bepaald kon worden. Zodoende kon op elke taak een maximum van twaalf punten behaald worden.

Twee beoordelaars (beiden direct betrokken bij het onderzoek) hebben, op basis van een gezamenlijk ontwikkeld scoringsprotocol en afzonderlijk van elkaar, van acht deelnemers uit elke conditie het prestatieformulier gescoord om zodoende de inter-beoordelaars betrouwbaarheid te bepalen. Overeenstemming (Cohen's Kappa) bleek .86 voor de concrete taak met een rijke interne relatiestructuur, .92 voor de concrete taak met een arme interne relatiestructuur en 1.00 voor de abstracte taak. Op basis van deze resultaten is besloten het scoringsprotocol ongewijzigd te hanteren en de scoring van de overige prestatieformulieren te verdelen onder beide beoordelaars.

Hypotheses

Voor het registreren van de hypothesen van de deelnemers is een hypotheseformulier ontwikkeld. Op dit formulier kan door de onderzoeker, bij ieder door de individuele deelnemer uitgevoerd experiment, aangegeven worden wat de deelnemer beoogt te onderzoeken en wat zijn of haar hypothese omtrent de uitkomst van het experiment is. Daarnaast kan het nummer van het specifieke experiment waar de intentie en hypothese van de deelnemer betrekking op had, aangegeven worden. Wanneer een deelnemer bepaalde uitgevoerde experimenten met elkaar gaat vergelijken kan ook hiervan aangegeven worden wat beoogd wordt te onderzoeken met de vergelijking en wat eventuele bijbehorende hypothesen zijn. Daarnaast kunnen de nummers van de experimenten die de deelnemer met elkaar vergelijkt, aangegeven worden. Invulling van dit formulier door twee onderzoekers leidde bij vijf gezamenlijk geobserveerde deelnemers tot 90 % overeenstemming. Op basis van dit resultaat is besloten de observering van de overige deelnemers (en daarbij de invulling van het hypotheseformulier) onder de twee onderzoekers te verdelen.

Voor het bepalen van het totaal aantal geformuleerde hypothesen en de verdeling van de typen hypothesen (met betrekking tot mate van specificiteit) zijn de hypotheseformulieren gescoord. Hierbij is het totale aantal geformuleerde hypothesen per deelnemer bepaald, alsmede een procentuele verdeling met betrekking tot typen hypothesen. Deze typen hypothesen bestonden uit: hypothesen waarin melding gemaakt wordt van een effect (type 'effect'), hypothesen waarin tevens melding gemaakt wordt van een richting van een effect (type 'richting') en hypothesen waarin tevens melding gemaakt wordt van een omvang van een effect (type 'omvang'). Een vierde type hypothese is die van type hypothese waarin geen melding gemaakt wordt van een concrete hypothese (type 'blanco'). Kortom: geen vermelding van effect, richting of omvang.

Twee beoordelaars hebben van acht deelnemers uit elke conditie het bijbehorende hypotheseformulier beoordeeld om zodoende de inter-beoordelaars betrouwbaarheid te bepalen. Deze (Cohen's Kappa) bleek .97 voor de concrete taak met een rijke interne relatiestructuur, .96 voor de concrete taak met een arme interne relatiestructuur en .97 voor de abstracte taak. Op basis van deze resultaten is besloten de scoring van de overige hypotheseformulieren te verdelen onder beide beoordelaars.

Experimenten

Voor het registreren van het tijdens de taak uitgevoerde totaal aantal experimenten, aantal unieke experimenten en aantal dubbele experimenten is gebruik gemaakt van de logfiles. Deze individuele logfiles zijn geanalyseerd en op basis hiervan zijn de gemiddelde aantallen bepaald.

Design en data analyse

Een univariate variantieanalyse (ANOVA) is gebruikt om de drie verschillende condities (concreet rijk, concreet arm en abstract), op basis van gemiddelde prestatie van de deelnemers, gemiddeld totaal aantal uitgevoerde experimenten door de deelnemers en gemiddeld totaal aantal geformuleerde hypothesen door de deelnemers, met elkaar te vergelijken. Daarnaast is deze toets gebruikt om het verschil in gemiddelde hypotheseverdeling van de deelnemers tussen de condities te analyseren. *Post*

hoc vergelijkingen tussen gemiddelden zijn uitgevoerd met de LSD methode (α was hierbij ingesteld op .05).

Resultaten

Kleurenblindheid

Van de zeventenvijftig deelnemers bleken twee deelnemers meer dan één afbeelding van de Ishihara kleurenblindheidtest foutief te hebben geïnterpreteerd (Ishihara, 1982). Van deze deelnemers is aangenomen dat ze kleurenblind waren en daarom problemen konden ondervinden bij het interpreteren van de afbeeldingen in de abstracte taak. Aangezien deze twee deelnemers zich beiden bevonden in de groep voor de concrete taak met een arme interne relatiestructuur is aangenomen dat deze geconstateerde kleurenblindheid geen probleem vormde met betrekking tot beïnvloeding van de prestatie.

Leerprestaties

Uit de uitgevoerde variantieanalyse bleek dat er statistisch significante verschillen bestonden in gemiddelde leerprestaties van de deelnemers (zie Tabel 1) tussen de drie condities, $F(2, 54) = 3.42$, $p < .05$. *Post hoc* vergelijkingen lieten zien dat er geen statistisch significant verschil aanwezig was in gemiddelde leerprestatie tussen de conditie ‘concreet rijk’ en de conditie ‘abstract’. De gemiddelde leerprestatie van de deelnemers in de conditie ‘concreet rijk’ bleek echter statistisch significant hoger dan die van deelnemers in de conditie ‘concreet arm’. Verder is er geen statistisch significant verschil gevonden in gemiddelde leerprestatie tussen de conditie ‘concreet arm’ en de conditie ‘abstract’, maar bleek dat er in dit verband wel sprake was van een trend ($p = .062$), in de richting van een hogere gemiddelde leerprestatie van deelnemers in de conditie ‘abstract’.

Tabel 1. Gemiddelde scores en gemiddelde percentages typen hypothesen (en standaarddeviaties), per conditie

	Conditie		
	Concreet rijk	Concreet arm	Abstract
Leerprestatie	9.58 (2.32)	7.37 (3.08)	9.05 (2.72)
Totaal aantal experimenten	21.16 (12.92)	19.16 (11.00)	24.11 (8.81)
Aantal unieke experimenten	15.79 (5.85)	13.84 (5.65)	18.16 (5.53)
Aantal dubbele experimenten	5.37 (8.55)	5.32 (6.48)	5.95 (5.68)
Aantal hypothesen	15.63 (6.45)	16.32 (7.83)	17.37 (6.01)
Aantal hypothesen (zonder type ‘blanco’)	13.68 (6.50)	7.32 (5.45)	8.37 (5.59)
<i>Percentages type hypothesen</i>			
Type ‘blanco’	14.23 (16.20)	50.10 (27.23)	54.54 (21.76)
Type ‘effect’	4.97 (6.03)	2.26 (4.36)	5.22 (9.37)
Type ‘richting’	67.32 (16.83)	32.50 (25.64)	15.50 (19.72)
Type ‘omvang’	13.48 (15.44)	15.26 (20.89)	24.74 (17.20)

Noot. $N = 57$.

Hypothesen

De drie condities zijn met elkaar vergeleken op basis van het gemiddeld totaal aantal geformuleerde hypothesen (zie Tabel 1). Uit deze analyse bleek dat er op dit vlak geen statistisch significante verschillen tussen de condities aanwezig waren, $F(2, 54) = .31$, $p = .73$. Echter, wanneer de hypothesen van het type ‘blanco’ buiten beschouwing werden gelaten (dit type hypothese geeft immers aan dat er eigenlijk geen hypothese geformuleerd is), blijkt dat er wel degelijk statistisch significante verschillen tussen de condities aanwezig waren, $F(2, 54) = 6.44$, $p < .01$. *Post hoc* vergelijkingen lieten zien dat het gemiddelde aantal geformuleerde hypothesen (zonder type ‘blanco’) statistisch significant hoger was in de conditie ‘concreet rijk’ dan in de conditie ‘abstract’ en de conditie ‘concreet arm’. Verder bleek dat er op dit punt geen verschil aanwezig was tussen de conditie ‘concreet arm’ en de conditie ‘abstract’.

De gemiddelde verdeling van de typen hypothesen (type 'blanco', type 'effect', type 'richting' en type 'omvang') in elke conditie is gebruikt om inzicht te krijgen in het onderzoekend leerproces. Deze gemiddelde verdeling van hypothesen is uitgedrukt in percentages ten opzichte van het gemiddeld totaal aantal geformuleerde hypothesen (zie Tabel 1).

Uit de analyse van deze gemiddelde verdeling van typen hypothesen tussen de drie condities bleek dat er geen verschillen aanwezig waren in de gemiddelde percentages geformuleerde type 'effect' hypothesen ($F(2, 54) = 1.03, p = .37$) en dat er geen verschillen aanwezig waren in de gemiddelde percentages geformuleerde type 'omvang' hypothesen ($F(2, 54) = 2.15, p = .13$). Dit wil zeggen dat het gemiddelde percentage geformuleerde hypothesen, waarin deelnemers enkel en alleen aangeven dat er een effect is en het gemiddelde percentage geformuleerde hypothesen waarin deelnemers de omvang (inclusief het effect en de richting ervan) aangeven binnen de drie condities niet van elkaar verschillen. Echter, er bleken statistisch significante verschillen tussen de condities te bestaan in de gemiddelde percentages geformuleerde type 'richting' hypothesen, $F(2, 54) = 29.92, p < .01$. *Post hoc* vergelijkingen toonden aan dat het gemiddelde percentage geformuleerde type 'richting' hypothesen statistisch significant hoger was in de conditie 'concreet rijk' dan in de conditie 'abstract' en de conditie 'concreet arm'. Daarnaast bleek uit deze vergelijkingen dat het gemiddelde percentage geformuleerde type 'richting' hypothesen statistisch significant hoger was in de conditie 'concreet arm' dan in de conditie 'abstract'. Dit wil zeggen, dat het gemiddelde percentage geformuleerde hypothesen, waarin deelnemers een effect en tevens de richting van dit effect aangeven, statistisch significant het hoogst was in de conditie concreet 'rijk' en statistisch significant het laagst was in de conditie 'abstract'.

Daarnaast bleken er statistisch significante verschillen tussen de condities te bestaan in de gemiddelde percentages geformuleerde type 'blanco' hypothesen, $F(2, 54) = 18.85, p < .01$. *Post hoc* vergelijkingen lieten zien dat het gemiddelde percentage geformuleerde type 'blanco' hypothesen statistisch significant hoger was in de conditie 'abstract' dan in de conditie 'concreet rijk'. Daarnaast bleek dit gemiddelde percentage statistisch significant hoger in de conditie 'concreet arm' dan in de conditie 'concreet rijk'. Dit wil zeggen, dat het gemiddelde percentage hypothesen waarin deelnemers aangeven dat ze geen hypothese hebben, statistisch significant hoger is in de conditie 'abstract' en in de conditie 'concreet arm' dan in de conditie 'concreet rijk' en dat er op dit vlak geen verschil was tussen de conditie 'abstract' en de conditie 'concreet arm'.

Experimenten

Uit de analyse van de gemiddelde aantallen uitgevoerde experimenten (zie Tabel 1) bleek dat er tussen de condities geen statistisch significante verschillen aanwezig waren in het gemiddeld totaal aantal uitgevoerde experimenten ($F(2, 54) = .97, p = .39$), het gemiddelde aantal uitgevoerde unieke experimenten ($F(2, 54) = 2.75, p = .07$) en het gemiddelde aantal uitgevoerde dubbele experimenten ($F(2, 54) = .05, p = .95$).

Discussie

Dit onderzoek had tot doel het effect van bekendheid met relaties tussen en binnen variabelen bij onderzoekend leertaken op de leerprestatie en het onderzoeksgedrag te bepalen. Dit is gedaan door middel van het vergelijken van de leerprestatie en het experimenteelgedrag op drie onderzoekend leertaken die varieerden in aanwezigheid van de subruimtes van de hypotheseruimte: de variabelenruimte en de relatieruimte.

Verwacht werd dat de gemiddelde leerprestatie van de deelnemers in de conditie 'concreet rijk' hoger zou zijn dan die van de deelnemers in de twee andere condities ('abstract' en 'concreet arm'). Verder werd verwacht dat er geen verschil in gemiddelde leerprestatie aanwezig zou zijn tussen de conditie 'concreet arm' en de conditie 'abstract'.

Met betrekking tot het experimenteelgedrag was de verwachting dat het experimenteelgedrag van deelnemers in de conditie 'concreet rijk' beter zou zijn (een lager gemiddeld aantal uitgevoerde experimenten, een hoger gemiddeld aantal geformuleerde hypothesen en een gemiddeld hoger aandeel

specifieke hypothesen) dan die van deelnemers in de twee andere condities en dat er geen verschil in experimenteelgedrag aanwezig zou zijn tussen de conditie 'concreet arm' en de conditie 'abstract'.

De verwachtingen met betrekking tot de gemiddelde leerprestatie werden gedeeltelijk door de gevonden resultaten ondersteund. Zo bleek er geen verschil te bestaan in gemiddelde leerprestatie tussen de conditie 'concreet rijk' en de conditie 'abstract', terwijl verwacht werd dat deelnemers in de conditie 'concreet rijk' een gemiddeld hogere leerprestatie zouden behalen dan deelnemers in de conditie 'abstract' (zoals eerder aangetoond in de studie van Wilhelm en Beishuizen (2003)). Mogelijke verklaring hiervoor is dat deelnemers in de conditie 'abstract' een mogelijk meer gestructureerde zoektocht naar de effecten gehanteerd hebben, doordat deelnemers in deze conditie zich niet konden beroepen op voorkennis betreffende de variabelen en de relaties ertussen en erbinnen (en daarmee op relevante hypothesen). Omdat de taken in het onderzoek van Wilhelm en Beishuizen van grotere omvang waren dan de taken in dit onderzoek (vijf in plaats van vier onafhankelijke variabelen) heeft een mogelijk meer gestructureerde werkwijze van lerenden op een abstracte onderzoekend leertaak in hun onderzoek mogelijk niet geleid tot een (voldoende) verbeterde leerprestatie. De gemiddelde leerprestatie van deelnemers in de conditie 'concreet rijk' bleek echter statistisch significant hoger dan die van deelnemers in de conditie 'concreet arm'. Een van tevoren gestelde en voor de hand liggende verklaring is dat deelnemers in de conditie 'concreet rijk' konden beschikken over voorkennis over de variabelen en hun relaties en dus konden profiteren van een initieel gevulde hypotheseruimte die hen in staat stelde om relevante hypothesen op te stellen voordat ze met experimenteren begonnen. Verder konden deelnemers in deze conditie profiteren van het feit dat ze konden vertrouwen op verklarende mechanismen, waardoor ze ontdekkingen beter konden interpreteren en/of beter in staat waren om alternatieve hypothesen te formuleren. Deelnemers in de conditie 'concreet arm' beschikten weliswaar over voorkennis over de variabelen (beschikking over een initieel gevulde variabelenruimte), ze beschikten niet over voorkennis over de relaties tussen deze variabelen (geen beschikking over een initieel gevulde relatieruimte) en waren daarom niet in staat om relevante hypothesen op te stellen voordat ze met experimenteren begonnen. Verder konden ze, doordat ze geen voorkennis hadden over de relaties tussen de variabelen, niet profiteren van verklarende mechanismen. In lijn met de verwachtingen is er geen statistisch significant verschil in gemiddelde leerprestatie gevonden tussen deelnemers in de conditie 'concreet arm' en de conditie 'abstract'. Deelnemers in de conditie 'concreet arm' beschikten weliswaar over voorkennis over de variabelen, ze beschikten niet over voorkennis over de relaties tussen en binnen deze variabelen en waarom daarom, netzo als de deelnemers in de conditie 'abstract', niet in staat om relevante hypothesen op te stellen voordat ze met experimenteren begonnen. Kennis van de relaties tussen en binnen variabelen lijkt voor de prestatiefacilitering bij onderzoekend leertaken dan ook van cruciaal belang.

De verwachtingen met betrekking tot experimenteelgedrag werden, evenals die van de leerprestatie, gedeeltelijk door de resultaten ondersteund; zo bleek tussen de drie condities geen statistisch significant verschil te bestaan met betrekking tot het gemiddelde aantal totaal, unieke en dubbel uitgevoerde experimenten. Mogelijke verklaring hiervoor is de grootte van de standaarddeviaties (die in sommige gevallen zelfs de omvang van de gemiddeldes overschreden).

Met betrekking tot het gemiddelde aantal geformuleerde hypothesen leek er op het eerste gezicht geen statistische verschillen tussen de drie condities. Wanneer echter het aandeel type 'blanco' hypothesen buiten beschouwing werd gelaten (en het dus om uitspraken over een gerichte verwachting gaat) bleken er wel degelijk verschillen tussen de condities te bestaan; deelnemers in de conditie 'concreet rijk' bleken (conform de verwachting) statistisch significant meer hypothesen geformuleerd te hebben dan deelnemers in de twee andere condities. Tussen de beide andere condities bleken geen verschillen te bestaan.

Ten aanzien van de specificiteit van de geformuleerde hypothesen bleken er geen verschillen tussen de condities te bestaan tussen de gemiddelde percentages geformuleerde type 'effect' hypothesen. Verder bleek dat er ook geen statistisch significante verschillen tussen de condities waren met betrekking tot de gemiddelde percentages van het meest specifieke type hypothese (type 'omvang'). Hoewel deze bevinding niet overeenkomstig is met vooraf opgestelde verwachting, lijkt deze uitkomst toch te verklaren. Wanneer je een uitspraak wilt doen (hypothese stellen) over de omvang van een effect is de kans dat je hypothese vervolgens verworpen wordt groot. Immers, de werkelijke uitkomst van het effect hoeft er maar één punt naast te zitten of je opgestelde hypothese

over de omvang van de uitkomst blijkt foutief. Van Joolingen en De Jong (1993) spreken in dit verband over 'angst voor verwerping'. Volgens hen hebben studenten de tendens om hypothesen die een grote kans op verwerping hebben te vermijden, zoals hypothesen waarin de relatie een grote mate van precisie heeft. Het ligt daarom meer voor de hand om een hypothese te formuleren over de richting van een effect. Over deze richting van een effect konden deelnemers in de conditie 'concreet rijk' wel uitspraken doen die een geringe kans tot verwerping hadden; ze beschikten immers over voorkennis over de variabelen en hun relaties. Uit de analyse van de hypotheseverdeling bleek deze meer voor de hand liggende type specifieke hypothese (type 'richting') inderdaad voor een statistisch significant gemiddeld groter aandeel geformuleerd te zijn bij de deelnemers in de conditie 'concreet rijk' dan bij de deelnemers in de twee andere condities. Daarnaast is, doordat lerenden in de conditie 'concreet rijk' konden beschikken over verklarende mechanismen, het formuleren en onderzoeken van een omvanghypothese vaak overbodig, aangezien de bevindingen van een eerder geformuleerde richtinghypothese betekenisvol geïnterpreteerd konden worden. Verder bleek het gemiddelde percentage geformuleerde type 'richting' hypothese hoger te zijn in de conditie 'concreet arm' dan in de conditie 'abstract'.

Naast de gemiddelde percentages geformuleerde type 'richting hypothesen, bleken ook de gemiddelde percentages geformuleerde hypothesen van het type 'blanco' (geen hypothese geformuleerd) conform de verwachting: het gemiddelde percentage hypothesen van dit type bleek hoger te zijn in de conditie 'concreet arm' en de conditie 'abstract' dan in de conditie 'concreet rijk' (voorkennis over de relaties tussen variabelen ontbreekt (geen beschikking over een initieel gevulde relatieruimte), waardoor het vormen van een hypothese over een effect zeer lastig is). Kanttekening die bij de gemiddelde verdeling van typen hypothesen geplaatst moet worden is dat de standaarddeviaties erg groot bleken en dat er dus, met betrekking tot deze verdeling, een aanzienlijke spreiding tussen individuele deelnemers aanwezig was.

Over het algemeen kan gesteld worden dat deelnemers in de conditie 'concreet rijk' beter experimenteelgedrag hebben getoond dan deelnemers in de conditie 'concreet arm' en de conditie 'abstract' en dat er geen noemenswaardige verschillen in experimenteelgedrag bestonden tussen deelnemers uit de conditie 'concreet arm' en de conditie 'abstract'. Samen met de bevindingen betreffende leerprestaties is het aannemelijk dat wanneer men bij het uitvoeren van een onderzoekend leertaak kan beschikken over een initieel (gedeeltelijk) gevulde hypotheseruimte (en dus over voorkennis beschikt over zowel de variabelen als hun relaties) dit een faciliterend effect heeft met betrekking tot de taakuitvoer en -prestatie. Het is bij het uitvoeren van onderzoekend leertaken met een concrete taakinhoud dus de combinatie van concreetheid van variabelen en de concreetheid van relaties die faciliteert. Concreetheid van de variabelen an sich lijkt niet tot facilitatie te leiden, aangezien er geen significante verschillen betreffende gemiddelde leerprestatie zijn gevonden tussen deelnemers uit de conditie 'concreet arm' en de conditie 'abstract'. Het is in dit verband interessant om te vermelden dat er zelfs sprake was van een trend die pleit voor een gemiddeld hogere leerprestatie van deelnemers in de conditie 'abstract'. Hierbij lijkt het dus alsof de beschikking van voorkennis over de variabelen, zonder daarbij te kunnen beschikken over voorkennis over hun relaties, de leerprestatie belemmert. Kortom, er lijkt sprake van een soort interactie; wanneer de variabelenruimte (initieel) gevuld is en de relatieruimte niet, presteert men slechter dan wanneer beiden (initieel) leeg zijn. Een mogelijke verklaring hiervoor is dat hoewel beide taken fictief waren, dit in de conditie 'abstract' duidelijker naar voren kwam (de variabelen bestonden hier immers uit abstracte figuren). In de conditie 'concreet arm' bestonden de variabelen uit herkenbare objecten. Logische relaties tussen deze variabelen (en tussen de niveaus binnen een variabele) ontbraken bewust om er voor te zorgen dat deelnemers niet konden beschikken over voorkennis over deze relaties (en het dus moesten stellen zonder een initieel gevulde relatieruimte). Een mogelijkheid is dat deelnemers, door de herkenbaarheid van de variabelen, uit alle macht op zoek zijn gegaan naar enige vorm van (volgens hen) logische relaties ertussen en erbinnen. Deze mogelijke zoektocht naar betekenis en verklaringen (die ook blijkt uit het gemiddeld hogere percentage geformuleerde richtinghypothesen bij deelnemers in de conditie 'concreet arm') zou bij constatering van het ontbreken van enige logische vorm van relaties tussen en binnen de variabelen geleid kunnen hebben tot onbegrip, verwarring en (wellicht) het negeren van bewijs, volgens De Jong en Van Joolingen (1998), een van de problemen waar studenten bij onderzoekend leren mogelijk mee geconfronteerd worden), iets wat de ontdekking van effecten (en daarbij de leerprestatie) natuurlijk niet ten goede komt. In de conditie 'abstract' ligt

een zoektocht naar betekenis en verklaringen daarentegen niet voor de hand; abstracte figuren zijn per definitie niet betekenisvol en een logische relatie ertussen is daarom niet te verwachten, iets wat mogelijk een meer gestructureerde zoektocht naar de effecten tot gevolg heeft gehad.

Met de in dit onderzoek gedane bevindingen dient zowel bij het ontwerpen als bij het aanbieden van onderzoekend leertaken rekening gehouden te worden; onderzoekend leertaken kunnen eenvoudiger gemaakt worden door ze concreter te maken en/of voorkennis met betrekking tot de taakinhoud (over de variabelen in de taak en de relaties ertussen en erbinnen) bij lerenden te activeren, dan wel aan te bieden.

Hoewel in dit onderzoek aangenomen is dat deelnemers bij de concrete taken over reeds (gedeeltelijk) gevulde (sub)hypotheseruimtes van de lerende konden beschikken (en deze aanname gezien de geringe moeilijkheidsgraad van de domeinen ook verantwoord lijkt), is het niet met zekerheid te stellen dat deze kennis ook daadwerkelijk aanwezig was. Het daadwerkelijk meten van deze kennis (en dus de omvang van de (sub)hypotheseruimtes) in verder onderzoek zal leiden tot een hogere validiteit van de in dit onderzoek beweerde bevindingen.

Zoals vermeld bleek er geen verschil in gemiddelde leerprestatie tussen deelnemers in de conditie 'concreet rijk' en de conditie 'abstract' en was er sprake van een trend die pleit voor een gemiddeld hogere leerprestatie van deelnemers in de conditie 'abstract' dan van deelnemers in de conditie 'concreet arm'. Vervolgonderzoek zou meer duidelijkheid kunnen verschaffen over de (geformuleerde) verklaringen en redenen hiervoor. Zo zou onderzocht kunnen worden of confrontatie met een abstracte onderzoekend leertaak inderdaad leidt tot een meer gestructureerde werkwijze van lerenden, dan wanneer geconfronteerd met een taak in een bekend domein. Daarnaast zou vervolgonderzoek zich kunnen richten op identificatie van de gesuggereerde zoektocht naar betekenis en verklaringen met betrekking tot de relaties tussen en binnen variabelen in een taak waarbij enige logische vorm van relaties hiertussen en hierbinnen ontbreekt (zoals in de taak uit de conditie 'concreet arm' het geval was). Daarnaast zou verder onderzoek zich kunnen richten op een verdere detaillering omtrent de omvang van de relatieruimte, door onderscheid tussen taken te maken met betrekking tot de aanwezigheid van relaties tussen en/of binnen variabelen. Zou zo de gemiddelde leerprestatie en/of experimenteergedrag van deelnemers in een conditie waarbij enkel relaties binnen variabelen bestaan (relaties tussen de niveaus van een variabele) vergeleken kunnen worden met die van deelnemers in condities waarbij daarnaast relaties tussen de variabelen bestaan, dan wel geen relaties bestaan, dan wel enkel relaties tussen variabelen bestaan. Dergelijk onderzoek zal leiden tot meer duidelijkheid over de (faciliterende) rol van concreetheid in onderzoekend leertaken.

Referenties

- Benjafield, J. G. (1997). *Cognition*. Upper Saddle River: Prentice-Hall.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. Huntington, NY: Wiley and Sons.
- Ceraso, J., & Provitera, A. (1971). Sources of error in syllogistic reasoning. *Cognitive Psychology*, 2, 400-410.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187-316.
- De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68(2), 179-201.
- De Jong, T., & Njoo, M. (1992). Learning and instruction with computer simulations: Learning processes involved. In E. de Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer based learning environments and problem solving*. Berlin: Springer.
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, 11, 295-306.
- Gijlers, A. H. (2005). *Confrontation and co-construction: Exploring and supporting collaborative scientific discovery learning with computer simulations*. (Doctoral dissertation, University of Twente, Enschede, The Netherlands).
- Giroto, V., & Legrenzi, P. (1989). Mental representation and hypothetico-deductive reasoning: The case of the THOG problem. *Psychological Research*, 51, 129-135.
- Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73, 407-420.
- Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69, 366-378.
- Hulshof, C. D., Wilhelm, P., Beishuizen, J. J., & Van Rijn, H. (2005). FILE: a tool for the study of inquiry learning. *Computers in Human Behavior*, 21, 945-956.
- Ishihara, S. (1982). *Ishihara's test for colour deficiency*. Tokyo: Kanehara.
- Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, 63, 395-400.
- Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, 1, 134-148.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-48.
- Koslowski, B. (1996). *Theory and evidence: the development of scientific reasoning*. Londen: MIT.
- Kuhn, D., Garcia-Mila, M., Zohar, A., & Anderson, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, 60(4), (Serial no. 245).
- Manktelow, K. I. (1999). *Reasoning and thinking*. Hove: Psychology press.
- Manktelow, K. I., and Evans, J. St. B. T. (1979). Facilitation of reasoning by realism: effect or non effect? *British Journal of Psychology*, 70, 477-488.
- Revlis, R. (1975). Syllogistic reasoning: logical decisions from a complex data base. In R. J. Falmagne (Eds.), *Reasoning: Representation and process*. New York: Wiley.
- Van Joolingen, W. R., & De Jong, T. (1993). Exploring a domain through a computer simulation: traversing variable and relation space with the help of a hypothesis scratchpad. In D. Towner, T. de Jong & H. Spada (Eds.), *Simulationbased experiential learning*. Berlin, Germany: Springer-Verlag.
- Van Joolingen, W. R., & De Jong, T. (1997). An extended dual search space model of learning with computer simulations. *Instructional Science*, 25, 307-346.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Eds), *New horizons in psychology*. Harmondsworth: Penquin.
- Wason, P. C. (1977). Self-contradictions. In P. N. Johnson-Laird, & P. C. Wason (Eds.), *Thinking: Readings in cognitive science*. Cambridge: Cambridge University Press.

- Wason, P. C., & Brooks, P. J. (1979). THOG: The anatomy of a problem. *Psychological Research, 41*, 79-90.
- Wason, P. C., & Shapiro D. A. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology, 23*, 63-71.
- Wilhelm, P., & Beishuizen, J. J. (2003). Content effects in self-directed inductive learning. *Learning and Instruction, 13*, 381-402.