

De rol van voorkennis bij onderzoekend leren met computersimulaties

Mieke Hagemans

Universiteit Twente



## Samenvatting

In dit onderzoek is de rol van voorkennis tijdens onderzoekend leren met computersimulaties onderzocht. Van 36 deelnemers zijn de leerprestatie en de onderzoekend leren processen 'hypotheseformatie' en 'experimenteren' met elkaar vergeleken op een concrete onderzoekend leertaak (in een bekend domein) en een abstracte onderzoekend leertaak (in een onbekend domein). Op beide taken dienden deelnemers de invloed van een aantal onafhankelijke variabelen op een afhankelijke variabele te bepalen. Voorafgaand aan het experiment is voorkennis over het onderwerp van de concrete taak gemeten. Bij de abstracte taak is aangenomen dat voorkennis geen rol speelde. Met betrekking tot de leerprestatie was de verwachting dat deelnemers beter zouden presteren op de concrete taak dan op de abstracte taak. Deze hypothese wordt door de resultaten ondersteund. De verwachting ten aanzien van het onderzoekend leren proces 'hypotheseformatie' was dat deelnemers meer en specifiekere hypothesen zouden formuleren op de concrete taak dan op de abstracte taak. Uit de resultaten bleek dat deelnemers, conform de verwachting, meer hypothesen hebben geformuleerd op de concrete taak. Ten aanzien van de specificiteit van de geformuleerde hypothesen is een verdeling in typen hypothesen gemaakt, uitgedrukt in percentages ten opzichte van het totaal aantal geformuleerde hypothesen. Uit de resultaten bleek dat het percentage hypothesen waarin deelnemers aangaven dat er eigenlijk geen hypothese bij een uitgevoerd experiment geformuleerd kon worden ('blanco hypothesen'), zoals verwacht, groter is op de abstracte taak dan op de concrete taak. Het percentage hypothesen waarin deelnemers aangaven dat een variabele een effect had ('effecthypothesen'), bleek op beide taken niet van elkaar te verschillen. Het percentage hypothesen waarin deelnemers zowel het bestaan als de richting van het effect van een variabele vermeldden ('richtinghypothesen'), bleek, naar verwachting, groter te zijn op de concrete taak. Het percentage hypothesen waarin deelnemers zowel het bestaan, de richting als de omvang van het effect van een variabele vermeldden ('omvanghypothesen'), bleek, tegen de verwachting in, groter te zijn op de abstracte taak. Ten aanzien van het onderzoekend leren proces 'experimenteren' was de verwachting dat deelnemers meer experimenten uit zouden voeren op de abstracte taak dan op de concrete taak. Uit de resultaten is gebleken dat er geen statistisch significant verschil bestaat ten aanzien van het totaal aantal uitgevoerde experimenten op beide taken, maar dat er wel sprake lijkt te zijn van een trend die in lijn is met de verwachting. In de discussie worden mogelijke verklaringen gegeven voor de gevonden resultaten. Tevens worden implicaties voor de praktijk bediscussieerd en worden aanbevelingen voor verder onderzoek gedaan.

## Introductie

In het dagelijkse leven bestaan veel momenten waarop we nieuwe dingen willen leren. In essentie bestaan er twee manieren om in een dergelijke leerbehoefte te voorzien; aan de ene kant kun je kennis vergaren die aangeboden wordt door iets of iemand anders dan jezelf, bijvoorbeeld wanneer een docent je iets vertelt over de Tweede Wereldoorlog of wanneer je een studieboek leest over natuurkundige principes. Aan de andere kant kun je kennis opdoen door middel van onderzoekend leren. Mensen maken in het dagelijkse leven vaak gebruik van onderzoekend leren, veelal zonder zich ervan bewust te zijn. Dit is bijvoorbeeld het geval wanneer verschillende routes naar het werk met elkaar vergeleken worden of wanneer geprobeerd wordt de werking van een softwareprogramma onder de knie te krijgen door het systematisch uitproberen van alle functies. Dergelijke exploraties dragen bij aan het vergaren van nieuwe kennis.

Op onderwijsgebied kent onderzoekend leren een lange geschiedenis. Het vindt zijn oorsprong in de Gestaltpsychologie van Bruner (De Jong & Van Joolingen, 1998). Onderzoekend leren heeft het afgelopen decennium aan populariteit gewonnen toen een paradigmashift ervoor zorgde dat onderwijsonderzoekers meer nadruk gingen leggen op de rol van de lerende (Veermans & Van Joolingen, 2004). Onderzoekend leren past beter binnen deze filosofie dan expositorisch onderwijs, omdat het de lerende ziet als actieve deelnemer aan zijn of haar eigen leerprocessen (Veenman, 1993). De lerende construeert zijn of haar eigen kennis door experimenten uit te voeren en uiteindelijk de onderliggende principes van een bepaald domein te ontdekken. Omdat computersimulaties de capaciteit hebben lerenden te voorzien van leeromgevingen waarin exploratie centraal staat, worden ze gezien als krachtige tools voor onderzoekend leren (Rheid, Zhang, & Chen, 2003).

Ondanks de veelbelovende verwachtingen blijkt dat de effectiviteit van onderzoekend leren laag is wanneer lerenden over onvoldoende voorkennis beschikken. Tuovinen en Sweller (1999) deden onderzoek naar het werken met databases, waarbij gekeken werd naar het oefenen met bestaande voorbeelden ten opzichte van het oefenen door middel van exploratie. De resultaten lieten zien dat lerenden zonder voorkennis over databases substantieel meer profiteerden van het oefenen met bestaande voorbeelden in vergelijking met lerenden die oefenden door middel van exploratie. Echter, wanneer lerenden bekend waren met het domein werd er ten aanzien van de leerresultaten geen significant verschil gevonden tussen de typen instructie. Deze resultaten suggereren dat bekendheid met een bepaald domein een faciliterend effect heeft op de leerprocessen en dat dit er mogelijk voor zorgt dat lerenden met veel voorkennis meer ontdekken en daardoor meer leren dan lerenden die over weinig voorkennis beschikken. Wat precies dit faciliterende effect is, wordt uit de studie van Tuovinen en Sweller echter niet duidelijk. Dit onderzoek<sup>1</sup> probeert dan ook een antwoord te vinden op de vraag wat de rol van voorkennis is tijdens onderzoekend leren.

Onderzoekend leren is het proces van het ontdekken van regels over relaties tussen onafhankelijke en afhankelijke variabelen in een bepaald domein. Dit gebeurt op basis van het uitvoeren van experimenten waarin de onafhankelijke variabelen worden gemanipuleerd (Wilhelm, 2001; Wilhelm & Beishuizen, 2003). Voordat experimenten worden uitgevoerd, worden hypothesen gegenereerd over het te onderzoeken domein (De Jong, 2006). Vervolgens worden de experimenten door de lerende zelf ontworpen en uitgevoerd, waarna de resultaten gebruikt worden om kennis af te leiden over het te bestuderen domein. Onderzoekend leren moedigt de lerende aan actief deel te nemen aan zijn of haar eigen leerproces om zodoende zelf een kennisbasis op te bouwen (Van Joolingen, 1999; Swaak, De Jong, & Van Joolingen, 2004; Gijlers & De Jong, 2005). Met geschikte ondersteuning leidt onderzoekend leren, in vergelijking met traditionele onderwijsvormen, tot een dieper en betekenisvoller begrip van een domein (De Jong, 2006). Echter, de mate waarin lerenden profiteren van onderzoekend leren lijkt af te hangen van de voorkennis over het te bestuderen domein (Tuovinen & Sweller, 1999).

Volgens Hulshof (2001) zijn de belangrijkste activiteiten tijdens onderzoekend leren het verzamelen van data door middel van het uitvoeren van experimenten en het opstellen van hypothesen op basis van de gevonden data. De formatie van hypothesen en het uitvoeren van experimenten kunnen beide gezien worden als actieve zoekprocessen. De probleemoplossingtheorie van Newell en

---

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

Simon (1972) kan volgens Hulshof (2001) gerelateerd worden aan deze activiteiten. Volgens deze theorie kan probleemoplossen gedefinieerd worden als het bewegen in een probleemruimte, waarin een verschuiving plaats vindt van de ene probleemtoestand naar de andere. Het resultaat van het bewegen in een probleemruimte is een actief zoekproces.

Simon en Lea (1974) introduceerden een systeem gebaseerd op het induceren van gegeneraliseerde regels, de 'Generalized Rule Inducer'(GRI) (in: Klahr & Dunbar, 1988). Dit systeem is een uitbreiding van de probleemoplossingstheorie. De GRI is een informatieverwerkingsstelsel dat prestaties op twee verschillende taken verklaart; probleemoplossingstaken en taken waarbij regels geïnduceerd dienen te worden. De probleemruimten die gebruikt worden voor het oplossen van problemen en het induceren van regels zijn verschillend; probleemoplossingstaken vinden alleen plaats in een op zichzelf staande ruimte met regels. Het doorzoeken van twee probleemruimten is vereist bij taken waarbij regels geïnduceerd dienen te worden. Hierbij is een probleemruimte nodig waarin zich regels bevinden en daarnaast is er een ruimte nodig waarin zich bewijzen bevinden.

Klahr en Dunbar (1988) hebben een uitbreiding op de GRI gemaakt. Zij introduceerden het 'Scientific Discovery as Dual-space Search' model (SDDS). Het SDDS model beschrijft een aantal processen van wetenschappelijk redeneren en kan op elke context toegepast worden waarin hypothesen worden gegenereerd en data wordt verzameld. Het basisidee van het model is dat wetenschappelijk redeneren plaats vindt in twee gerelateerde probleemruimten: de hypotheseruimte en de experimentruimte. De hypotheseruimte bevat alle hypothesen die de lerende tijdens het onderzoeksproces kan formuleren (Gijlers & De Jong, 2005). De experimentruimte bestaat uit alle experimenten die de lerende kan uitvoeren (Klahr & Dunbar, 1988; Gijlers & De Jong, 2005).

In principe stelt het SDDS model dat onderzoekend leren uit drie kernactiviteiten bestaat; het doorzoeken van de hypotheseruimte, het testen van hypothesen en het evalueren van bewijs voor deze hypothesen (Klahr & Dunbar, 1988; Veenman, 1993; Van Joolingen & De Jong, 1997). Het resultaat van het doorzoeken van de hypotheseruimte is een gespecificeerde hypothese. Door middel van het uitvoeren van een experiment wordt deze hypothese getest, wordt er een voorspelling gedaan en wordt het resultaat geobserveerd. Het resultaat van het testen van de hypothese is een beschrijving van het bewijs dat voor of tegen de hypothese pleit. Dit resultaat wordt als input gebruikt om het bewijs te evalueren. Deze evaluatie heeft als doel te bepalen of het bewijs de hypothese bevestigt of dat verder onderzoek naar de hypothese nodig is.

De drie kernactiviteiten zijn onderverdeeld in verschillende subprocessen (Klahr & Dunbar, 1988). Twee van deze subprocessen, het oproepen van een "frame" en het gebruiken van voorkennis, doen een beroep op de kennis van de lerende. Beide subprocessen vinden plaats in de hypotheseruimte. Bij het oproepen van een "frame" zoekt de lerende in zijn of haar geheugen naar situaties of ervaringen waaruit hypothesen afgeleid kunnen worden. Om deze hypothese volledig te kunnen specificeren gebruikt de lerende zijn of haar voorkennis. De aanwezigheid of afwezigheid van voorkennis impliceert dus een verschillende configuratie van het doorzoeken van de hypotheseruimte.

Klahr en Dunbar (1988) hebben hun ideeën getest door middel van een programmeerbare robot, genaamd BigTrak. Deelnemers aan het onderzoek dienden de werking van de herhalingsbutton (RPT button) van BigTrak te ontdekken. De mogelijke regels die de functie van deze button beschrijven, werden gezien als de hypotheseruimte. De experimentruimte bestond uit alle mogelijke programma's en gedragingen van BigTrak. Tijdens een training werden deelnemers eerst geïnstrueerd over de manier waarop commando's in het programma van BigTrak gecombineerd konden worden om zo een bepaald doel te bereiken. Na de training participeerden de deelnemers in een instructieloze fase. Aan de deelnemers werd verteld dat er slechts één herhalingsbutton in het programma aanwezig is en dat deze een numerieke parameter hanteert. Deelnemers werden gevraagd de werking van deze herhalingsbutton te onderzoeken door middel van het formuleren en testen van hypothesen. Negentien van de twintig deelnemers ontdekten de werking van de herhalingsbutton binnen de toegestane tijd.

Uit de analyse van de hypothesen die de deelnemers formuleerden, bleek dat zij verschillende benaderingen gebruikten om de hypotheseruimte te doorzoeken. Vanuit de analyses redeneerden Klahr en Dunbar (1988) dat het doorzoeken van de hypotheseruimte gebeurde op basis van het doorzoeken van twee geheugenruimten; de ruimte waarin zich de mogelijke experimenten bevinden en de ruimte waarin zich de mogelijke hypothesen bevinden. Ze identificeerden op basis van die redenering twee hoofdbenaderingen die deelnemers gebruikten voor het genereren van nieuwe hypothesen. De eerste benadering, waarin het geheugen werd doorzocht, noemden Klahr en Dunbar de theoretische

benadering. De tweede benadering, waarbij de resultaten van eerder gedane experimenten werden gegeneraliseerd om hypothesen te kunnen genereren, noemden Klahr en Dunbar de experimenteerbenadering.

Dertien deelnemers hanteerden de experimenteerbenadering. Zij doorliepen twee fasen. In de eerste fase formuleerden ze mogelijke hypothesen en evalueerden deze door middel van het uitvoeren van experimenten. Wanneer bleek dat weinig vooruitgang werd geboekt ten aanzien van het ontdekken van de werking van de herhalingsbutton, switchten experimenteerders naar het uitvoeren van experimenten zonder hierbij een expliciete hypothese te stellen. Deze tweede fase kan gezien worden als een exploratie van de experimentruimte. Experimenteerders gebruikten hierbij verschillende strategieën: sommigen stelden nieuwe hypothesen op door deze af te leiden uit eerdere experimenten en anderen gebruikten dezelfde hypothese om deze pas na veelvuldig experimenteren te verwerpen. Deelnemers die de experimenteerbenadering hanteerden, waren in staat de werking van de herhalingsbutton te ontdekken.

Zeven deelnemers hanteerden de theoretische benadering. Deze benadering bestond uit het genereren van hypothesen en het uitvoeren van experimenten om de waarde van de gegenereerde hypothesen te testen. De hypotheseruimte werd doorzocht waarna de hypothesen in de experimentruimte werden getest. Wanneer genoeg bewijs was verzameld om de hypothese te verwerpen, werd gewicht naar een andere hypothese. Bewijs voor een mogelijke hypothese werd gezocht in het geheugen, zodat een hypothese geformuleerd kon worden, in plaats van het veelvuldig uitvoeren van experimenten. Theoretici kozen over het algemeen de correcte hypothese. Wanneer de hypothese niet werd bevestigd, ontdekten ze de correcte hypothese door het veranderen van een element in hun zoekproces.

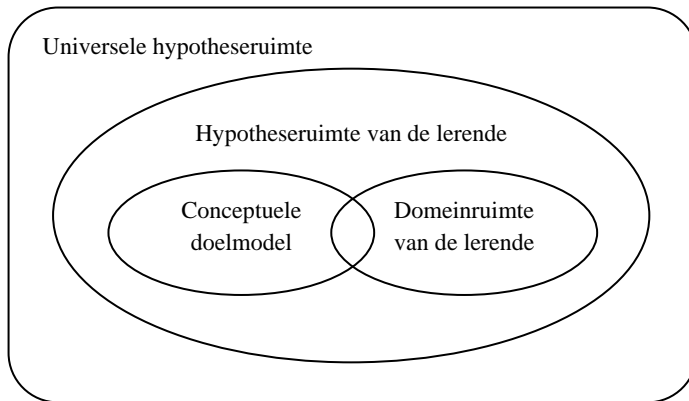
Er kan geconcludeerd worden dat beide benaderingen verschillen in de manier waarop nieuwe hypothesen worden gegenereerd. Binnen de theoretische benadering wordt de hypotheseruimte doorzocht en binnen de experimenteerbenadering de experimentruimte. Theoretici hadden minder tijd nodig om de werking van de herhalingsbutton te ontdekken. Experimenteerders hadden meer experimenten nodig waarin een expliciete hypothese ontbrak. Gezien de resultaten van het onderzoek van Klahr en Dunbar (1988) lijkt het waarschijnlijk dat de keuze voor een theoretische benadering dan wel experimenteerbenadering afhankelijk is van het voorkennisniveau van een lerende.

Klahr en Dunbar (1988) noemden drie voorkennis categorieën die mogelijk de hypothesen van deelnemers beïnvloedden: a) linguïstische kennis over de betekenis van het woord 'herhaling', b) programmeerkennis over herhalingen en c) specifieke kennis over BigTrak. Deelnemers weten dat het woord 'herhaling' betekent dat iets opnieuw gedaan wordt. Deelnemers met programmeerervaring leidden kennis ten aanzien van een herhalingsbutton af uit andere programma's. Ten aanzien van BigTrak wisten deelnemers dat er verschillende typen buttons bestonden. Klahr en Dunbar veronderstelden dat het effect van de verschillende soorten voorkennis merkbaar is tijdens de formatie van hypothesen en dat dit verschillen in experimenteren teweegbrengt. Deze veronderstelling blijft echter een aanname gezien het feit dat Klahr en Dunbar de invloed van de drie voorkennis categorieën niet hebben gemeten.

Van Joolingen en De Jong (1997) en Gijlers en De Jong (2005) hebben een extensie van het SDDS model gemaakt. De reden voor deze extensie is tweeledig; allereerst was het gebruikte domein in het onderzoek van Klahr en Dunbar (1988) eenvoudig doordat slechts één regel ontdekt hoefde te worden. Bij onderzoekend leren met computersimulaties is het domein vaak complex en spelen meerdere variabelen en relaties een rol (Van Joolingen & De Jong, 1997). Ten tweede is het SDDS model niet in staat om een gedetailleerde beschrijving van de zoekoperaties door de hypotheseruimte weer te geven. Dergelijke beschrijvingen kunnen nodig zijn voor het begrijpen van de manier waarop voorkennis het doorzoeken van de hypotheseruimte en de experimentruimte beïnvloedt.

De hoofdelementen in het uitgebreide SDDS model zijn een gedetailleerde bewerking van de hypotheseruimte en van mechanismen die het zoekproces door de hypotheseruimte beschrijven. Van Joolingen en De Jong (1997) en Gijlers en De Jong (2005) representeren de kennis van de lerende over een domein in termen van subruimten van de hypotheseruimte (zie Figuur 1). De complete hypotheseruimte wordt de universele hypotheseruimte genoemd en bestaat uit alle mogelijke hypothesen over een bepaald domein. De hypotheseruimte van de lerende is de ruimte waarin zich alle veronderstellingen, variabelen en relaties bevinden waarover de lerende kennis bezit. De lerende kan deze ruimte gebruiken om te doorzoeken en om het te ontdekken domein te beschrijven. De

hypotheseruimte van de lerende is verder onderverdeeld in twee verschillende ruimten; de domeinruimte van de lerende en de ruimte van het conceptuele doelmodel. De domeinruimte van de lerende representeert de kennis van de lerende en bestaat uit de hypothesen over relaties in het domein die de lerende als “waar” beschouwd. De lerende test deze hypothesen en bevestigt of verworpt ze of markeert ze als nader te onderzoeken hypothesen. In de ruimte van het conceptuele doelmodel bevinden zich alle valide veronderstellingen over het domein.



Figuur 1. Het uitgebreide SDDS model. (Bron: Gijlers & De Jong, 2005).

Het uitgebreide SDDS model is dynamisch. Wanneer de lerende kennis vergaart over het onderliggende model, beweegt de domeinruimte van de lerende richting de ruimte van het conceptuele doelmodel. Het uitgebreide SDDS model differentieert voor veronderstellingen die al getest zijn en voor veronderstellingen waarvan de waarheid bij de lerende nog niet bekend is. Deze veronderstellingen zijn gesitueerd in verschillende componenten van het model.

Uit onderzoek van Klahr en Dunbar (1988) en Van Joolingen en De Jong (1997) is gebleken dat voorkennis vooral van invloed is op de onderzoekend leren processen ‘hypotheseformatie’ en ‘experimenteren’. Deze bevindingen zouden kunnen verklaren waarom de effectiviteit van onderzoekend leren laag is wanneer lerenden over weinig voorkennis beschikken (Tuovinen & Sweller, 1999). Lerenden met veel voorkennis hebben een grote hypotheseruimte waardoor veel hypothesen geformuleerd kunnen worden. Deze lerenden gebruiken de experimentruimte om de gegenereerde hypothesen te toetsen en al dan niet te bevestigen. De hypotheseruimte van lerenden die over weinig voorkennis beschikken is beperkter. Deze lerenden kunnen weinig of geen geschikte hypothesen over het gestelde probleem formuleren. De experimentruimte moet geëxploreerd worden om een geschikte hypothese te kunnen formuleren of om een hypothese af te kunnen leiden uit uitgevoerde experimenten.

In dit onderzoek wordt de rol van voorkennis onderzocht tijdens onderzoekend leren met computersimulaties. Het onderzoek van Klahr en Dunbar (1988) is hierbij als uitgangspunt genomen. De door hun gebruikte experimentele onderzoeksopzet is op twee punten aangepast, zodat een meer gecontroleerde vergelijking uitgevoerd kon worden. Ten eerste wordt in dit onderzoek de voorkennis over het domein gemeten. Het tweede verschil betreft de ‘within-subject’ vergelijking van de leerprestatie en de onderzoekend leren processen ‘hypotheseformatie’ en ‘experimenteren’ op een concrete en een abstracte onderzoekend leertaak. De concrete taak is gesitueerd in een bekend domein; voorkennis over het onderwerp van deze taak is voorafgaand aan het experiment gemeten. De abstracte taak is gesitueerd in een onbekend domein. Deze taak is fictief, zodat kon worden aangenomen dat de deelnemers aan het onderzoek geen voorkennis over de taakinhoud of het domein zouden hebben.

Het vergelijken van de onderzoekend leren processen en de leerprestatie op beide taken zal inzicht bieden in de veronderstelde relatie tussen domeinkennis en onderzoekend leren vaardigheden. Ten aanzien van de leerprestatie wordt verwacht dat lerenden beter zullen presteren op de concrete taak dan op de abstracte taak. Lerenden zullen waarschijnlijk hun voorkennis gebruiken op de concrete taak, maar kunnen dat niet op de abstracte taak. De hypothese ten aanzien van het onderzoekend leren proces ‘hypotheseformatie’ is dat lerenden meer en specifiekere hypothesen zullen formuleren op de concrete taak dan op de abstracte taak. Wat betreft het onderzoekend leren proces ‘experimenteren’ is

de verwachting dat er meer experimenten uitgevoerd zullen worden op de abstracte taak dan op de concrete taak. De achterliggende gedachte hierbij is dat deelnemers aan het onderzoek een op toeval en ontdekkingen gebaseerde benadering zullen hanteren tijdens het uitvoeren van de abstracte taak en een op voorkennis gebaseerde benadering op de concrete taak. Doordat er over de abstracte taak geen voorkennis aanwezig is, gezien het feit dat deze taak fictief is, is er een gebrek aan kennis over de variabelen en relaties in het domein. In vergelijking met de concrete taak is de hypotheseruimte van deelnemers ten aanzien van de abstracte taak kleiner, zodat er minder hypothesen geformuleerd kunnen worden. Voor het formuleren van relevante hypothesen over het domein zullen binnen een op toeval en ontdekkingen gebaseerde aanpak veel experimenten nodig zijn.

## **Methode**

### *Deelnemers*

Aan het onderzoek hebben 38 studenten (16 mannen, 22 vrouwen, gemiddelde leeftijd 19,3 jaar,  $SD = 1.8$ ) van de Faculteit Gedragswetenschappen van de Universiteit Twente deelgenomen. Dertig deelnemers hadden de Nederlandse nationaliteit en acht deelnemers waren van Duitse afkomst. De Duitse deelnemers hebben het staatsexamen Nederlands als tweede taal (NT2) succesvol afgelegd en waren daardoor goed in staat de gesproken en geschreven teksten in het onderzoek te begrijpen. Door deelname aan het onderzoek konden deelnemers twee proefpersoonpunten behalen.

### *Materialen*

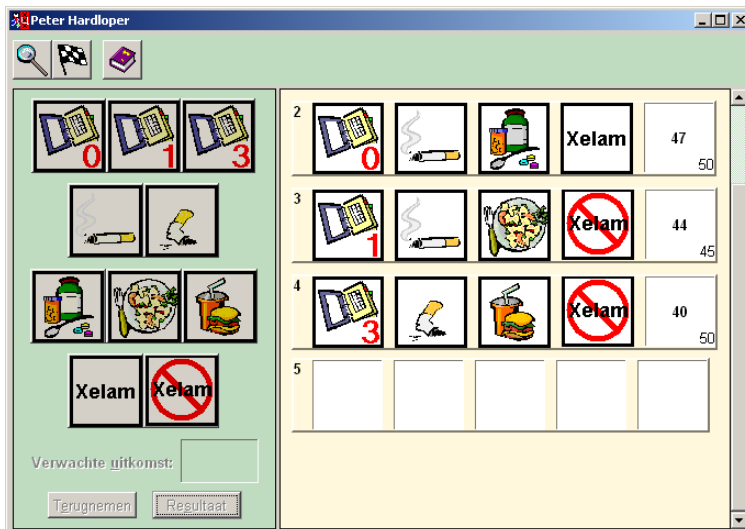
#### *Onderzoekend leren taken.*

Alle deelnemers hebben twee onderzoekend leren taken uitgevoerd, te weten een concrete en een abstracte taak. De taken zijn ontwikkeld in een auteursomgeving genaamd FILE (zie Hulshof, Wilhelm, Beishuizen, & Van Rijn, 2005).

Figuur 2 geeft de interface van het hoofdscherm van de concrete taak weer. In deze taak werd deelnemers gevraagd te onderzoeken welke invloed vier onafhankelijke variabelen hadden op de tijd waarin een jongen, genaamd Peter, de 10 kilometer loopt. De onafhankelijke variabelen met bijbehorende niveaus waren: a) het aantal trainingen per week (nul, één of drie keer), b) roken (wel roken, niet roken), c) voeding (koolhydraatrijke sportvoeding, eten wat de pot schaft, junkfood) en d) Xelam (wel, niet). De invloed van de eerste drie variabelen is bij een breed publiek bekend. De vierde variabele, 'Xelam', was fictief. Kennis over deze factor was dan ook niet aanwezig. Deelnemers wisten niet hoe Xelam de prestaties op de 10 kilometer beïnvloedde.

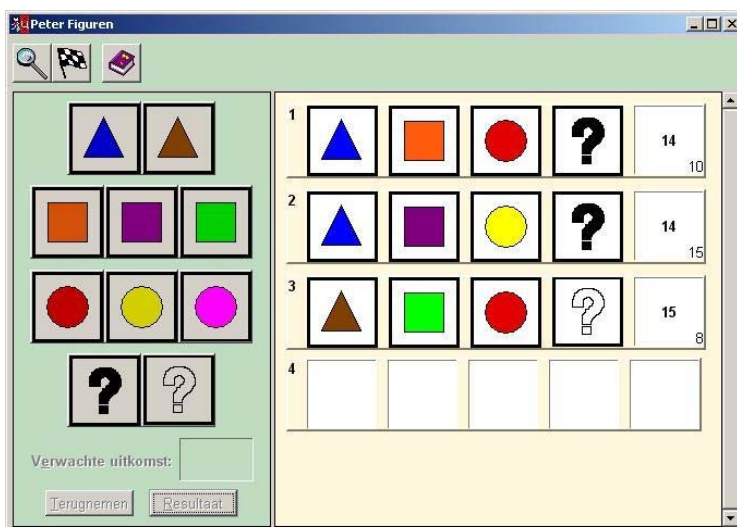
Door het variëren van de niveaus van de vier variabelen kon de tijd waarin Peter de 10 kilometer loopt, bepaald worden. De taak bevatte een interactie-effect tussen de variabelen 'training' en 'Xelam'; de werking van Xelam kon bepaald worden door deze te combineren met één van de niveaus van de variabele 'training'. Door het uitvoeren van verschillende experimenten kon uiteindelijk de invloed van de verschillende variabelen en niveaus op de tijd waarin Peter de 10 kilometer loopt, bepaald worden.





Figuur 2. Interface van de concrete taak.

In de abstracte taak werden deelnemers geconfronteerd met vier soorten abstracte figuren waarvan ze de invloed moesten bepalen op een puntentotaal (zie Figuur 3). De onafhankelijke variabelen met bijbehorende niveaus in deze taak waren: a) een driehoek (blauw, bruin), b) een vierkant (oranje, paars, groen), c) een cirkel (rood, geel, roze) en d) een vraagteken (zwart, wit). De abstracte taak was qua onderliggend model isomorf aan de concrete taak. Echter, de omvang van de effecten verschilde, evenals de volgorde van de onafhankelijke variabelen met hun bijbehorende niveaus in de interface. De abstracte taak bevatte, evenals de concrete taak, een interactie-effect; de invloed van het vraagteken was mede afhankelijk van het niveau van de cirkel.



Figuur 3. Interface van de abstracte taak.

Deelnemers konden in de simulatieomgeving experimenten uitvoeren door de niveaus van de vier onafhankelijke variabelen te variëren en het effect ervan af te leiden uit de afhankelijke variabele. De simulatie registreerde alle door de deelnemers uitgevoerde handelingen, met bij behorende tijdsvermelding, in een logfile.

Om de opdrachtomschrijving van de desbetreffende taak te kunnen lezen, dienden deelnemers in het hoofdscherm van de simulatieomgeving de button met het pictogram van het boek aan te klikken. De opdrachtomschrijving opende vervolgens in een nieuw scherm. Vervolgens kon gestart worden met het uitvoeren van experimenten. Elke onafhankelijke variabele werd aan de linkerkant van het scherm weergegeven als een reeks afbeeldingen, met één afbeelding voor elk niveau. Voor elke variabele dienden deelnemers een gewenst niveau te selecteren. Wanneer een niveau was gekozen,

was het onmogelijk tegelijkertijd een tweede niveau van dezelfde variabele te selecteren. De simulatieomgeving gaf dit weer door de desbetreffende variabele grijs te kleuren. Indien deelnemers een ongewenst niveau hadden geselecteerd, kon de button 'terugnemen' aangeklikt worden. Het laatst geselecteerde niveau van de variabele werd dan verwijderd en het gewenste niveau kon dan alsnog geselecteerd worden. Na het selecteren van de gewenste niveaus voor de variabelen dienden deelnemers een verwachte uitkomst in te geven. Om het experiment te voltooien, diende de button 'resultaat' aangeklikt te worden. Uitgevoerde experimenten verschenen aan de rechterkant op het scherm. Bij elk experiment werden de geselecteerde niveaus van de onafhankelijke variabelen, de verwachte uitkomst en de werkelijke uitkomst weergegeven. Een compleet experiment bleef staan en kon niet meer veranderd worden. Eerder gedane experimenten verdwenen van het scherm wanneer er meer experimenten uitgevoerd werden dan op het scherm weergegeven konden worden. Door middel van de scroll-balk konden deelnemers de eerder gedane experimenten bekijken en vergelijken. Tevens konden deelnemers experimenten vergelijken door de te vergelijken experimenten te selecteren en vervolgens de button met het pictogram van de loep aan te klikken. De geselecteerde experimenten werden dan onder elkaar in een apart scherm getoond. Ook hier kon eventueel de scroll-balk gebruikt worden om de geselecteerde experimenten te bekijken. Tot slot kon de taak afgesloten worden door de button met het pictogram van de finishvlag aan te klikken.

#### *Voortoets.*

Door middel van een voortoets is de kennis over de variabelen in de concrete taak gemeten. De abstracte taak was "inhoudsloos", zodat kon worden aangenomen dat deelnemers hierover geen voorkennis bezaten. De voortoets bestond uit tien meerkeuzevragen en één open vraag. Alle meerkeuze vragen kenden een 'als-dan'-vorm.

Drie meerkeuzevragen gingen over het algemene effect van de onafhankelijke variabelen 'training', 'roken' en 'voeding'. Een voorbeeld van een vraag over het algemene effect van de variabele 'training' was: "Je hebt je twee maanden voorbereid op de wedstrijd over 10 kilometer. Naarmate je vaker traint, zonder overtraining, zul je de tien kilometer waarschijnlijk a) sneller lopen, b) net zo snel lopen of c) langzamer lopen".

De overige zeven meerkeuzevragen bevatten een vergelijking tussen twee niveaus van de variabelen. Een voorbeeld van een vraag met een niveauvergelijking over de variabele 'voeding' was: "Je hebt je twee maanden voorbereid op de wedstrijd over 10 kilometer. In die periode heb je koolhydraatrijke sportvoeding gegeten. Stel dat je in die periode junkfood had gegeten. Dan zou je de 10 kilometer waarschijnlijk a) sneller lopen, b) net zo snel lopen of c) langzamer lopen".

De open vraag ging over de variabele 'Xelam'. Deze vraag was bedoeld om de preconcepties van de deelnemers over deze fictieve variabele te meten. Deelnemers werd gevraagd wat ze dachten dat Xelam was en wat de invloed ervan zou kunnen zijn op de prestatie op de tien kilometer.

De betrouwbaarheid van de meerkeuzevragen is berekend door middel van de K-R 20 formule (Dousma, Horsten, & Brants, 1997). Voor de voortoets geldt  $K-R 20 = .61$ .

#### *Ishihara kleurenblindheidtest.*

In de abstracte taak werden de verschillende niveaus van de onafhankelijke variabelen door middel van kleuren van elkaar onderscheiden. Eventuele kleurenblindheid bij deelnemers zou kunnen zorgen voor een ongewenste beïnvloeding van de prestatie. Om voor dit eventuele ongewenste effect te controleren, is een digitale versie van de Ishihara kleurenblindheidtest afgenomen (Ishihara, 1982). Deze test bestond uit acht afbeeldingen, opgebouwd uit gekleurde bolletjes. De deelnemers moesten aangeven welke cijfers er in het bolletjespatroon van elke afbeelding verborgen zaten. Het allereerste cijfer was voor iedereen, ongeacht kleurenblindheid, zichtbaar. De overige zeven cijfers differentieerden voor kleurenblindheid voor de kleuren rood of groen. Drie van de zeven cijfers onderscheidden kleurenblindheid voor de kleur rood en de overige vier cijfers onderscheidden kleurenblindheid voor de kleur groen.

#### *Antwoordformulier.*

De prestatie van de deelnemers op de concrete en de abstracte taak is gemeten door het scoren van een speciaal ontwikkeld antwoordformulier dat deelnemers invulden. Op deze manier kon nagegaan

worden in hoeverre de individuele deelnemer het onderliggende model van de desbetreffende taak had ontdekt.

Het antwoordformulier bevatte een instructie waarin de deelnemer gevraagd werd op te schrijven wat hij of zij had ontdekt over de invloed van de onafhankelijke variabelen en hun niveaus op de afhankelijke variabelen in de concrete en abstracte taak, respectievelijk de tijd in minuten om de 10 kilometer te lopen en het puntentotaal. Op het antwoordformulier stonden de verschillende onafhankelijke variabelen uit de concrete en abstracte taak vermeld met daaronder ruimte om aantekeningen te maken en ontdekkingen ten aanzien van de afhankelijke variabele te vermelden.

### *Procedure*

Het onderzoek is uitgevoerd in een tijdsbestek van een maand. Voor afname van het onderzoek stond maximaal twee uur per deelnemer. Het onderzoek is voor elke deelnemer individueel afgenomen.

Allereerst werden deelnemers welkom geheten en dienden de deelnemers een formulier te ondertekenen waarin de rechten en plichten van de deelnemers en de onderzoeker vastgelegd waren. Vervolgens hebben de deelnemers de voortoets gemaakt. Na het inleveren van de voortoets bij de onderzoeker werd de digitale Ishihara kleurenblindheidstest afgenomen. Vervolgens werd de werking van de simulatie uitgelegd aan de hand van een introductietaak die geen verband hield met de beide taken (het bepalen van de kosten van een wintersportvakantie) en konden de deelnemers vragen stellen alvorens ze de concrete en abstracte taken zouden gaan uitvoeren.

Voor het uitvoeren van de taken werd aan de deelnemers verteld dat ze tijdens het uitvoeren ervan het antwoordformulier in moesten vullen en dat er tijdens het uitvoeren van de taken geen vragen gesteld mochten worden. Tot slot werd de deelnemers verteld dat de onderzoeker tijdens het uitvoeren van de taken vragen zou stellen op de volgende momenten: voordat de deelnemers een verwachte uitkomst selecteerden, wanneer de deelnemers gingen scrollen en wanneer er gebruik gemaakt werd van het vergelijken van een set experimenten door middel van de button met het pictogram van de loep. Op die momenten is gevraagd naar wat de deelnemers precies gingen onderzoeken en wat ze dachten dat de uitkomst zou zijn en waarom, om zodoende de hypothesen van de deelnemers te kunnen achterhalen.

Elke deelnemer heeft vervolgens de concrete en abstracte taak uitgevoerd. Om volgorde-effecten te voorkomen heeft de helft van de deelnemers eerst de concrete taak uitgevoerd en vervolgens de abstracte taak. De andere helft van de deelnemers kreeg deze taken in omgekeerde volgorde aangeboden. Voor het uitvoeren van de taken kregen de deelnemers per taak maximaal 45 minuten de tijd. Er mocht 40 minuten geëxperimenteerd worden in de simulatieomgeving, daarna kregen de deelnemers nog vijf minuten de tijd om het antwoordformulier eventueel verder aan te vullen. Tijdens deze vijf minuten mochten de deelnemers nog wel de gedane experimenten bekijken, maar geen nieuwe experimenten meer uitvoeren. Als deelnemers het gevoel hadden dat ze alle relaties in de taak ontdekt hadden, konden ze eerder stoppen en dit aangeven bij de onderzoeker. Om onduidelijkheden te voorkomen, kregen deelnemers bij de taken een formulier met daarop de betekenis van de afbeeldingen van de onafhankelijke variabelen en hun niveaus.

Tijdens het uitvoeren van de taken is door de onderzoeker een formulier ingevuld waarop genoteerd werd wat er precies onderzocht werd en wat de daarbij behorende verwachte uitkomst, ofwel hypothese, was. Om dit voor de desbetreffende deelnemer te achterhalen, zijn per uitgevoerd experiment op van te voren bepaalde momenten vragen gesteld (zie hierboven). Na 40 minuten werd aangegeven dat de tijd voor de eerste taak verstreken was en dat het antwoordformulier, indien nodig, nog verder aangevuld mocht worden. Vervolgens werd de taak afgesloten en werd overgegaan naar de tweede taak. Voor deze taak is dezelfde procedure gehanteerd als voor de eerste taak.

Na het voltooien van de tweede taak werden de deelnemers bedankt voor hun deelname en werd verteld dat de proefpersoonpunten toegekend zouden worden. Tot slot werd de deelnemers gevraagd geen informatie over het onderzoek aan anderen mede te delen.

## *Codering*

### *Prestatie.*

Om de domeinruimte van de deelnemers te meten, zijn de scores van de deelnemers op de concrete en de abstracte taak bepaald met behulp van een scoringsprotocol voor het antwoordformulier. Voor de beschrijving van de ontdekte invloed van de vier onafhankelijke variabelen konden per onafhankelijke variabele op beide taken maximaal drie punten behaald worden. Indien de deelnemer vermeldde dat een variabele een effect had, werd één punt toegekend. Dit was bijvoorbeeld het geval wanneer een deelnemer bij de variabele 'voeding' aangaf dat gezondere voeding van invloed is op de prestatie. Twee punten werden toegekend wanneer ook de correcte richting van het effect werd vermeld. Deelnemers kregen bijvoorbeeld twee punten voor de variabele 'voeding' wanneer vermeld werd dat je met het eten van wat de pot schaft een betere prestatie op de tien kilometer neerzet dan met het eten van junkfood en dat je met het eten van koolhydraatrijke sportvoeding een betere prestatie neerzet dan met het eten van wat de pot schaft of met het eten van junkfood. Indien de deelnemer het effect, de correcte richting van het effect en de correcte omvang van het effect vermeldde, werden drie punten toegekend. Dit was bijvoorbeeld het geval wanneer vermeld werd dat de prestatie op de tien kilometer met het eten van wat de pot schaft met twee minuten verbeterde in vergelijking tot het eten van junkfood en dat het eten van koolhydraatrijke sportvoeding de prestaties met twee minuten verbeterde ten opzichte van het eten van wat de pot schaft en met vier minuten verbeterde in vergelijking met het eten van junkfood.

De variabelen 'Xelam' en 'vraagteken' interacteerden met één van de andere onafhankelijke variabelen. In de concrete taak interacteerde de variabele 'Xelam' met de variabele 'training' en in de abstracte taak interacteerde de variabele 'vraagteken' met de variabele 'cirkel'. Indien de deelnemer vermeldde met welke variabele Xelam of het vraagteken interacteerde, zodat aangenomen kon worden dat het interactie-effect ontdekt was, werden de punten toegekend volgens het scoringsstelsel. Indien de deelnemer dit niet vermeldde, werden nul punten toegekend. De vermelding van het interactie-effect mocht ook aangegeven worden bij de variabele waarmee Xelam of het vraagteken interacteerde, dus bij de variabele 'training' of bij de variabele 'cirkel'.

Twee beoordelaars, beiden direct betrokken bij het onderzoek, hebben het scoringsprotocol ontwikkeld om de antwoordformulieren te scoren. Vervolgens hebben beide beoordelaars van acht willekeurig geselecteerde deelnemers zowel de concrete als de abstracte taak afzonderlijk van elkaar gescoord om de inter-beoordelaars overeenstemming te bepalen. Cohen's Kappa = .86 voor de concrete taak en voor de abstracte taak geldt Cohen's Kappa = 1.00. Op basis van deze resultaten is besloten het protocol ongewijzigd te hanteren voor het scoren van de overige antwoordformulieren. De overige te scoren antwoordformulieren zijn onder beide beoordelaars verdeeld en vervolgens gescoord.

### *Hypothesen.*

Om de intentie van de deelnemers ten aanzien van de uitgevoerde experimenten en de daarbij behorende hypothesen te kunnen registreren, is een hypotheseformulier ontwikkeld. Op dit formulier kon, bij elk door de individuele deelnemer uitgevoerd experiment, door de onderzoeker genoteerd worden wat de desbetreffende deelnemer beoogde te onderzoeken en wat zijn of haar verwachting was. Hetgeen de deelnemer had onderzocht kon op het formulier aangegeven worden door het aankruisen van één of meerdere van de op het formulier vermelde variabelen en niveaus uit de desbetreffende taak. Voor het noteren van de hypothese bij het uitgevoerde experiment was een schrijfruimte beschikbaar. Het nummer van het desbetreffende experiment kon aangegeven worden door het nummer aan te kruisen in een op het formulier bestaande ruimte. Daarnaast konden in dezelfde ruimte de nummers van de door de deelnemers vergeleken experimenten genoteerd worden.

Invulling van het hypotheseformulier door twee onderzoekers leidde bij vijf gezamenlijk onderzochte deelnemers tot 90% overeenstemming. Op basis van dit resultaat is besloten de overige te onderzoeken deelnemers, met de bijbehorende invulling van de hypotheseformulieren, te verdelen onder deze twee onderzoekers.

Voor het meten van de hypotheseruimte is het totaal aantal geformuleerde hypothesen op de concrete en de abstracte taak bepaald en zijn de geformuleerde hypothesen geclassificeerd. Hieruit is een procentuele verdeling van de typen hypothesen ontstaan. De verdeling in typen hypothesen is als

volgt bepaald: wanneer een deelnemer een hypothese formuleerde waarin melding gemaakt werd van een effect, was er sprake van een 'effecthypothese'. Wanneer eveneens melding gemaakt werd van de richting van het effect, was er sprake van een 'richtinghypothese'. Indien melding gemaakt werd van zowel het bestaan van het effect, de richting van het effect als de omvang van het effect, was er sprake van een 'omvanghypothese'. Wanneer deelnemers een uitspraak deden waarbij niets gezegd werd over het bestaan van een effect, de richting van het effect of de omvang van het effect, was er sprake van een 'blanco hypothese'.

Twee direct bij het onderzoek betrokken beoordelaars hebben de hypotheseformulieren van acht willekeurig geselecteerde deelnemers voor zowel de concrete als de abstracte taak afzonderlijk van elkaar gescoord om de inter-beoordelaars overeenstemming te bepalen. Cohen's Kappa = .97 voor zowel de concrete als de abstracte taak. Op basis van dit resultaat is besloten de scoringsmethode ten aanzien van de verdeling van de typen hypothesen ongewijzigd te hanteren en de overige te scoren hypotheseformulieren te verdelen onder beide beoordelaars.

### *Experimenten.*

De experimentruimte van de deelnemers is gemeten door het totaal aantal door de deelnemers uitgevoerde experimenten op beide taken te bepalen aan de hand van een analyse van de logfiles uit de simulatieomgeving. Daarnaast is, eveneens door middel van een analyse van de logfiles, het aantal unieke experimenten en het aantal dubbel uitgevoerde experimenten op beide taken bepaald.

### *Design en data analyse*

Een 'within-subject' design is gebruikt om de prestatie, het totaal aantal uitgevoerde experimenten, het aantal uniek en dubbel uitgevoerde experimenten en het totaal aantal geformuleerde hypothesen van de deelnemers op de concrete en de abstracte taak met elkaar te vergelijken. Daarnaast is dit design gebruikt om de hypotheseverdeling van de deelnemers tussen beide taken te analyseren en te vergelijken. Hierbij is steeds gebruik gemaakt van een gepaarde *t*-toets.

## **Resultaten**

### *Kleurenblindheid*

Van de 38 deelnemers bleek één deelnemer vier van de zeven afbeeldingen in de digitale Ishihara kleurenblindheidtest (Ishihara, 1982) niet correct te hebben geïnterpreteerd. Van deze deelnemer is aangenomen dat hij problemen zou kunnen ondervinden bij het waarnemen van de afbeeldingen in de abstracte taak en dat er daardoor mogelijk ongewenste effecten ten aanzien van de prestatie van deze deelnemer zouden kunnen ontstaan. Er is besloten de deelnemer te verplaatsen naar een onderzoek dat in dezelfde periode als dit onderzoek is uitgevoerd. Uit dat onderzoek is een niet-kleurenblinde deelnemer random geselecteerd om aan dit onderzoek deel te nemen.

### *Prestatie*

De scores op de voortoets zijn gebruikt om de voorkennis over de concrete taak te bepalen ( $M = 9.36$ ,  $SD = 0.76$ ). In de analyse van de resultaten zijn de scores van de deelnemers die de drempelwaarde van acht van de tien punten niet behaalden niet meegenomen. Het ging hierbij om twee deelnemers met een score van respectievelijk vier en zeven punten.

De leerprestatiescores op beide taken (zie Tabel 1) zijn gebruikt om verschillen in leerprestatie te bepalen. De leerprestatiescore op de concrete taak bleek statistisch significant hoger dan de leerprestatiescore op de abstracte taak,  $t(35) = 2.74$ ,  $p < .05$ , wat impliceert dat deelnemers beter hebben gepresteerd op de concrete taak dan op de abstracte taak.

## Experimenten

Beide taken zijn met elkaar vergeleken op basis van het totaal aantal door de deelnemers uitgevoerde experimenten (zie Tabel 1). Uit de analyse bleek dat er geen statistisch significant verschil bestaat tussen het totaal aantal uitgevoerde experimenten op de concrete en de abstracte taak, wel lijkt er sprake van een trend,  $t(35) = -1.89, p = .07$ .

Naast een vergelijking van het totaal aantal door de deelnemers uitgevoerde experimenten, zijn ook het aantal unieke en dubbele experimenten op de concrete en de abstracte taak met elkaar vergeleken (zie Tabel 1). Uit de analyse bleek dat er geen statistisch significant verschil bestaat ten aanzien van het aantal uniek uitgevoerde experimenten op beide taken,  $t(35) = -1.21, p = .23$ . Tevens bleek dat er ook geen statistisch significant verschil bestaat wat betreft het aantal dubbel uitgevoerde experimenten op beide taken,  $t(35) = -1.80, p = .08$ .

Tabel 1. Gemiddelde scores en gemiddelde percentages typen hypothesen (en standaarddeviaties) op de concrete en de abstracte taak

	Concrete taak		Abstracte taak	
Leerprestatie	10.08	(1.99)	9.17	(2.40)
Totaal aantal uitgevoerde experimenten	21.00	(10.37)	24.03	(11.36)
Aantal unieke experimenten	16.64	(5.56)	17.69	(6.48)
Aantal dubbele experimenten	4.36	(6.53)	6.33	(7.34)
Aantal hypothesen	15.11	(5.15)	16.97	(6.82)
Aantal hypothesen zonder 'blanco hypothesen'	12.67	(5.37)	7.89	(5.87)
<i>Percentages typen hypothesen</i>				
Blanco hypothesen	16.92	(17.04)	56.95	(23.68)
Effecthypothesen	4.18	(5.76)	4.83	(7.90)
Richtinghypothesen	62.61	(20.24)	12.98	(20.49)
Omvanghypothesen	16.51	(18.09)	25.24	(20.48)

Noot.  $N = 36$

## Hypothesen

Het totaal aantal door de deelnemers geformuleerde hypothesen op de concrete en de abstracte taak is met elkaar vergeleken (zie Tabel 1). Uit de analyse bleek dat deelnemers statistisch significant meer hypothesen hebben geformuleerd op de abstracte taak dan op de concrete taak,  $t(35) = -2.26, p < .05$ . Wanneer de 'blanco hypothesen' in de analyse buiten beschouwing werden gelaten – omdat dit type hypothesen aangeeft dat er eigenlijk geen hypothese bij een uitgevoerd experiment is geformuleerd – bleek uit de analyse dat er een omgekeerd effect bestaat: deelnemers hebben significant meer hypothesen geformuleerd op de concrete taak dan op de abstracte taak,  $t(35) = 5.89, p < .01$ .

De verdeling van de typen hypothesen ('blanco hypothesen', 'effecthypothesen', 'richtinghypothesen' en 'omvanghypothesen') is geanalyseerd om inzicht te krijgen in het onderzoekend leren proces 'hypotheseformatie'. De verdeling van de typen hypothesen is uitgedrukt in percentages ten opzichte van het totaal aantal door de deelnemers geformuleerde hypothesen (zie Tabel 1). Analyse liet zien dat er geen statistisch significant verschil bestaat ten aanzien van het percentage geformuleerde 'effecthypothesen' op de concrete taak en de abstracte taak,  $t(35) = -.39, p = .70$ . Dit wil zeggen dat het percentage hypothesen, waarin deelnemers enkel en alleen aangaven dat een variabele een effect had, op beide taken niet van elkaar verschilt.

Uit de analyse bleek dat deelnemers een statistisch significant groter percentage 'richtinghypothesen' hebben geformuleerd op de concrete taak dan op abstracte taak,  $t(35) = 11.74, p < .01$ . Dit betekent dat op de concrete taak het percentage hypothesen waarin deelnemers het bestaan van het effect en tevens de richting van dit effect aangaven, groter is dan op de abstracte taak. Verder bleek dat deelnemers op de abstracte taak een statistisch significant groter percentage 'omvanghypothesen' hebben geformuleerd dan op de concrete taak,  $t(35) = -2.45, p < .05$ . Tevens bleek dat deelnemers op de abstracte taak een statistisch significant groter percentage 'blanco hypothesen' hebben geformuleerd dan op de concrete taak,  $t(35) = -10.28, p < .05$ . Dit wil zeggen dat op de abstracte taak zowel het percentage hypothesen waarin deelnemers de omvang van het effect,

inclusief het effect en de richting van het effect, aangaven als het percentage hypothesen waarin deelnemers aangaven dat ze geen hypothese hadden, groter is dan op de concrete taak.

## Discussie

Het doel van dit onderzoek was het bepalen van de rol van voorkennis tijdens onderzoekend leren met computersimulaties. Hierbij is gekeken naar de onderzoekend leren processen ‘hypotheseformatie’ en ‘experimenteren’ en naar de leerprestatie. De verwachting ten aanzien van de leerprestatie was dat deelnemers beter zouden presteren op de concrete taak dan op de abstracte taak. Ten aanzien van het onderzoekend leren proces ‘hypotheseformatie’, waarbij gekeken is naar het totaal aantal geformuleerde hypothesen en de specificiteit van deze hypothesen, werd verwacht dat deelnemers op de concrete taak meer en specifiekere hypothesen zouden formuleren dan op de abstracte taak. De hypothese met betrekking tot het onderzoekend leren proces ‘experimenteren’, waarbij gekeken is naar het totaal aantal uitgevoerde experimenten, was dat deelnemers meer experimenten uit zouden voeren op de abstracte taak dan op de concrete taak.

De hypothese ten aanzien van de leerprestatie wordt door de resultaten ondersteund. Gezien het feit dat er binnen proefpersonen gemeten is, is het zeer waarschijnlijk dat een verschil in voorkennis de verklarende factor voor het verschil in leerprestatie is; deelnemers hebben op de concrete taak mogelijk hun voorkennis gebruikt. Op de abstracte taak was dit voor de deelnemers niet mogelijk gezien het feit dat deze taak fictief was en er geen voorkennis over bestaat.

Ten aanzien van het onderzoekend leren proces ‘experimenteren’ lieten de resultaten zien dat het totaal aantal uitgevoerde experimenten op beide taken niet van elkaar verschilt. Wel lijkt er sprake te zijn van een trend die conform de verwachting is. Een mogelijke verklaring voor het resultaat zou kunnen liggen in de grootte van de standaarddeviaties (zie Tabel 1). Deze geven aan dat er sprake is van zeer veel spreiding ten aanzien van het totaal aantal uitgevoerde experimenten op beide taken.

Wat betreft het onderzoekend leren proces ‘hypotheseformatie’ kan op basis van de resultaten geconcludeerd worden dat deelnemers op de concrete taak meer experimenten hebben uitgevoerd waarbij een hypothese geformuleerd is dan op de abstracte taak. De resultaten lieten zien dat, wanneer de ‘blanco hypothesen’ in de berekeningen buiten beschouwing worden gelaten, er statistisch significant meer hypothesen zijn geformuleerd op de concrete taak dan op de abstracte taak, wat in overeenstemming is met de verwachting. Ten aanzien van de specificiteit van de geformuleerde hypothesen bleek uit de resultaten dat er, zoals verwacht, meer ‘blanco hypothesen’ geformuleerd zijn op de abstracte taak dan op de concrete taak.

Wat betreft de specificiteit van de hypothesen lieten de resultaten ten aanzien van de overige typen hypothesen zien dat wat betreft de ‘effecthypothesen’ het percentage hypothesen op beide taken niet van elkaar verschilt. Het resultaat ten aanzien van de ‘richtinghypothesen’ is conform de verwachting. Een mogelijke verklaring voor het resultaat ten aanzien van de ‘omvanghypothesen’, gezien vanuit een op voorkennis gebaseerde benadering, zou kunnen zijn dat op de concrete taak bij de deelnemers een verklaringsmechanisme een rol heeft gespeeld. Dit wil zeggen dat de aanwezigheid van kennis over een bepaald mechanisme dat het effect van een variabele in de taak verklaart, net zo belangrijk is als het bewijs over een bepaald effect van een variabele dat uit de data van de taak naar voren komt (Koslowski, 1996). De bij de deelnemers aanwezige kennis over mechanismen verklaart de gevonden effecten in de taak. Een voorbeeld ter illustratie: op de concrete taak dienden deelnemers het effect van roken op het aantal minuten waarin Peter de 10 kilometer loopt te onderzoeken. Wanneer de deelnemers door middel van het uitvoeren van experimenten beide niveaus van de variabele ‘roken’ onder verschillende condities hebben onderzocht en verschillende uitkomsten het resultaat zijn, is het plausibel dit effect toe te schrijven aan het al dan niet roken, omdat er een mechanisme aanwezig is dat de verschillen in uitkomsten verklaart; namelijk dat roken een negatieve invloed heeft op de tijd waarin Peter de 10 kilometer loopt, omdat roken slecht is voor de gezondheid. De richting van het effect van de variabelen en hun niveaus in de concrete taak strookt met de voorkennis van de deelnemers. Bij het formuleren van hypothesen is het daarom voor deelnemers, die een op voorkennis gebaseerde benadering hanteren, minder noodzakelijk om reeds ontdekte resultaten uit eerdere experimenten over de omvang van het effect te bekijken en

op te nemen in de te formuleren hypothese om uiteindelijk het onderliggende model te kunnen ontdekken.

Bij de abstracte taak is er geen mechanisme aanwezig dat het effect van de variabelen verklaart. Deelnemers beschikken niet over kennis op basis waarvan ze uitspraken kunnen doen over de verwachte richting van een effect tussen niveaus van variabelen. Gezien vanuit een op toeval en ontdekkingen gebaseerde benadering, welke verondersteld werd gehanteerd te worden op de abstracte taak, zijn experimenten nodig om tot relevante hypothesen over de variabelen en relaties in het domein te komen om uiteindelijk het onderliggende model te kunnen ontdekken. Op basis van het bekijken van de uitkomsten van eerder gedane experimenten en de daardoor vergaarde domeinkennis kunnen deelnemers tot zeer specifieke hypothesen komen. Het ligt daarom meer voor de hand dat deelnemers bij het formuleren van een hypothese over de variabelen en hun niveaus in de abstracte taak niet alleen de richting van het effect, maar ook de zojuist bekeken omvang van het effect vermelden.

Uit de resultaten van dit onderzoek blijkt dat de aanwezigheid van voorkennis de leerprestatie op onderzoekend leren taken ten goede komt. Onderzoekend leren lijkt daarom vooral effectief te zijn wanneer leerlingen al over enige kennis van het domein beschikken. Met deze bevinding dient rekening gehouden te worden bij zowel het ontwerpen als het aanbieden van onderzoekend leren taken in het onderwijs; de aanwezigheid van voorkennis dient gewaarborgd te worden. Wanneer voorkennis aanwezig is, zal deze geactiveerd moeten worden. Indien voorkennis niet aanwezig blijkt te zijn, dient domeininformatie aangeboden te worden.

In dit onderzoek is gekeken naar de rol van voorkennis tijdens onderzoekend leren gebaseerd op computersimulaties. Verder onderzoek zou zich niet alleen kunnen richten op het onderzoeken van de rol van voorkennis, maar ook op de ontwikkeling van domeinkennis gedurende onderzoekend leren met computersimulaties. Een andere aanbeveling voor verder onderzoek is het onderzoeken van het meest geschikte moment voor het aanbieden van domeininformatie, wanneer niet over voorkennis wordt beschikt. Uit de resultaten van dit onderzoek is namelijk gebleken dat de aanwezigheid van domeinkennis de leerprestatie op onderzoekend leren taken ten goede komt. Domeininformatie zou bijvoorbeeld voorafgaand aan een taak of tijdens een taak aangeboden kunnen worden. Er kan gekeken worden naar de invloed van het moment van aanbieden op de leerprestatie en de onderzoekend leren vaardigheden. Verwacht wordt dat het aanbieden van domeininformatie tijdens de taak zal leiden tot een betere leerprestatie, omdat er direct een koppeling tussen de taak en de aangeboden informatie gemaakt kan worden, waardoor de informatie op een betekenisvolle manier opgenomen en verwerkt kan worden.



## Referenties

- De Jong, T. (2006). Scaffolds for scientific discovery learning. In J. Elen & D. Clark (Eds.), *Handling complexity in learning environments: research and theory* (pp. 107-128). London: Elsevier Science Publishers.
- 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.
- Dousma, T., Horsten, A., & Brants, J. (1997). *Tentamineren*. Groningen: Wolters-Noordhoff.
- Gijlers, H., & De Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of Research in Science Teaching*, 42(3), 264-282.
- Hulshof, C.D. (2001). *Discovery about ideas and ideas about discovery*. PhD. Thesis, University of Twente.
- 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(6), 945-956.
- Ishihara, S. (1982). *Ishihara's test for colour deficiency*. Tokyo: Kanehara.
- 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*. London: MIT.
- Reid D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning*, 19(1), 9-20.
- Swaak, J., De Jong, T., & Van Joolingen, W. R. (2004). The effects of discovery learning and expository instruction on the acquisition of definitional and intuitive knowledge. *Journal of Computer Assisted Learning* 20, 225-234.
- Touvinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91, 334-341.
- Van Joolingen, W.R. (1999). Cognitive tools for discovery learning. *International Journal of Artificial Intelligence in Education*, 10, 385-397.
- Van Joolingen, W. R., & De Jong, T. (1997). An extended dual search space model of scientific discovery learning. *Instructional science*, 25, 307-346.
- Veenman, M. V.J. (1993). *Intellectual ability and metacognitive skill: determinants of discovery learning in computerized learning environments*. Thesis. University of Amsterdam.
- Veermans, K. H. & Van Joolingen, W. R. (2004). Combining heuristics and formal methods in a tool for supporting simulation-based discovery learning. *Lecture notes in Computer Science*, 3220, 217-226.
- Wilhelm, P. (2001). *Knowledge, skills and strategies in self-directed inductive learning*. Unpublished doctoral dissertation, Leiden University, Leiden.
- Wilhelm, P., & Beishuizen, J. J. (2003). Content effects in self-directed inductive learning. *Learning and Instruction*, 13, 381-402.