

Metacognitieve vaardigheden in ontwikkelingsperspectief bij onderzoekend leren

Masterthese

Bianca Nijhuis, s1067664

Faculteit Gedragwetenschappen

Psychologie

Instructie Leren en Ontwikkeling

Eerste begeleider: Dr. P. Wilhelm

Tweede begeleider: Dr. A.W. Lazonder

Universiteit Twente

ENSCHDE

15 augustus 2015

Samenvatting

Het eerste doel van dit onderzoek was te ontdekken of er verschillen bestaan in het niveau van metacognitieve vaardigheden bij onderzoekend leren tussen leerlingen uit 2VWO ($N = 27$) en universitaire bachelor studenten ($N = 25$) om zodoende een ontwikkeling te kunnen schetsen. Hiervoor zijn gegevens uit hardopdenkprotocollen van 52 deelnemers geanalyseerd en vergeleken met gegevens over leerprestaties, metacognitie en intelligentie. Een maat voor metacognitieve vaardigheden, rechtstreeks afgeleid van het leergedrag en het denkproces van de lerende, werd berekend en toegepast. Metacognitieve vaardigheden namen toe met de leeftijd en de gebruikte strategieën veranderden in de leeftijdsgroepen. Studenten bleken meer gebruik te maken van metacognitieve leerprocessen die ten gunste kwamen van de leerprestaties.

Het tweede doel was te onderzoeken welke relaties bestaan tussen metacognitieve vaardigheden, intelligentie en leerprestaties. Ondersteuning werd gevonden voor het mixed model van Veenman, Wilhelm en Beishuizen (2004) waarin verondersteld wordt dat metacognitie deels gerelateerd is aan intelligentie. In de leeftijdsgroepen werden middelmatige positieve significante correlaties gevonden tussen metacognitie en intelligentie, waarbij metacognitie en intelligentie beide een eigen meerwaarde hadden voor leerprestaties. De voorspellende waarde van intelligentie voor leerprestaties was in het huidige onderzoek laag. Metacognitieve vaardigheden hadden daarentegen grote voorspellende waarde voor leerprestaties. Ongewogen intelligentiescores namen toe met de leeftijd, wat duidde op intellectuele groei. Significante correlaties werden gevonden tussen metacognitie en intelligentie wat wees op een aan elkaar gerelateerde ontwikkeling. Hiermee is deels bevestiging gevonden voor de door Alexander, Carr en Schwanenflugel (1995) omschreven geleidelijke, parallelle groei van metacognitie met leeftijd samen met intelligentie.

Abstract

The first objective of this study was to discover to what extent differences exist in the level of metacognitive skillfulness in inquiry learning tasks between 2VWO students ($N = 27$) and university bachelor students ($N = 25$) in order to describe metacognitive development. Data from think aloud protocols of 52 participants were analyzed and integrated with data about learning performances, metacognition and intellectual ability. A measure for metacognitive skillfulness directly inferred from learning behavior and the process of thinking of the learner

was calculated and applied. Metacognitive skillfulness showed growth with age and the used strategies altered between age groups. Students used significant more metacognitive learning processes that enhanced learning performances.

The second objective was to examine the relations between metacognitive skillfulness, intellectual ability and learning performances. Support has been found for the mixed model of Veenman, Wilhelm and Beishuizen (2004) supposing that metacognitive skillfulness is related to a certain extend to intellectual ability. Moderate positive significant correlations between metacognition and intellectual ability were found between age groups. Both metacognition and intellectual ability have their own unique contribution to the prediction of learning performance. In the current study the predictive value of intellectual ability was low however the predictive value of metacognitive skillfulness was high. Unweighted scores for intellectual ability increased with age indicating intellectual growth. Between metacognition and intellectual ability significant correlations were found which indicates a related development. Therefore partial support has been found for the continuous and parallel growth of metacognition with age, alongside intelligence described by Alexander, Carr and Schwanenflugel (1995).

Theoretische inleiding

Metacognitie wordt gezien als de meest krachtige voorspeller van leren (Wang, Haertel en Walberg, 1990). Het is een breed begrip waaronder activiteiten vallen zoals plannen, monitoren en evalueren. Doorgaans wordt een onderscheid gemaakt tussen metacognitieve kennis en metacognitieve vaardigheden (Alexander, Carr, & Schwanenflugel, 1995; Veenman & Elshout, 1999). Metacognitieve kennis verwijst naar de declaratieve kennis van een persoon (Flavell, 1979) en metacognitieve vaardigheden bestaan uit de procedurele kennis die nodig is voor de daadwerkelijke regulatie van leeractiviteiten en de controle hierover (Flavell, 1992; Kluwe, 1987). Bij onderzoekend leren worden methodes of principes van wetenschappelijk onderzoek toegepast om problemen op te lossen (Lazonder, 2014). Deze processen van wetenschappelijk redeneren zijn mede afhankelijk van metacognitieve processen.

De intentie van deze these was om binnen de context van onderzoekend leren, het ontwikkelingsniveau van metacognitieve vaardigheden in kaart te brengen, evenals de relaties tussen metacognitieve vaardigheden, intelligentie en leerprestaties. Hiervoor zijn eerder verzamelde gegevens, uit het proefschrift van Wilhelm (2001), geïntegreerd met gegevens van hardopdenkprotocollen die destijds ook waren verzameld, maar nog niet waren geanalyseerd. Deze analyse van hardopdenkprotocollen sloot aan bij de eerder uitgevoerde onderzoeken van Wilhelm (2001), Veenman, Wilhelm en Beishuizen (2004) en Evers (2009).

Veenman, et al. (2004) gebruikten een indirecte maat op basis van logfiles voor het meten van metacognitie, in dit onderzoek werd voor twee leeftijdsgroepen een maat voor metacognitieve vaardigheden berekend en toegepast die rechtstreeks afgeleid is van het leergedrag en het denkproces van de lerende. Op deze manier wordt gekeken of deze bevindingen inhoudelijk kunnen worden ondersteund. Dit is van belang omdat de bevindingen uit het onderzoek van Veenman, et al. (2004) in de sociale wetenschappen regelmatig geciteerd zijn en nog steeds worden.

Onderzoekend leren

Onderzoekend leren is een manier van leren waarbij het zelf verkennen van de natuurlijke of materiële wereld betrokken is. Dit leidt tot het stellen van vragen, het doen van ontdekkingen, en het testen van deze ontdekkingen met als doel het zoeken naar nieuwe inzichten (National Science Foundation, 2000). Wilhelm (2001) stelt dat leerlingen bij onderzoekend leren in de huid van wetenschappers kruipen. Ze krijgen de opdracht zelfstandig experimenten uit te

voeren om iets te weten te komen over een bepaald onderwerp door systematisch te experimenteren. Onderzoekend leren is een actief en zelfgestuurd leerproces: lerenden voeren experimenten uit in een zelfgekozen volgorde om zo uit te zoeken hoe variabelen van een bepaald domein elkaar beïnvloeden (Anderson, 2002). Technologische ontwikkelingen, zoals computersimulaties, hebben ervoor gezorgd dat onderzoekend leren eenvoudig realiseerbaar is en meer effectief geïmplementeerd kan worden (Löhner, Van Joolingen, & Savelsbergh, 2003). Bij onderzoekend leren worden methodes of principes van wetenschappelijk onderzoek toegepast om problemen op te lossen (Lazonder, 2014). Hierbij staan drie processen centraal: het opstellen van hypothesen (opvattingen over het te onderzoeken fenomeen), experimenteren (ontwerpen en uitvoeren van experimenten) en het evalueren van het bewijs (conclusies afleiden op basis van experimenten). Voor het succesvol doorlopen van deze processen is het wenselijk dat deze ondersteund worden met regulatieve processen, ook wel metacognitieve processen genoemd, zoals plannen, monitoren en evalueren (Van Joolingen, de Jong, Lazonder, Savelsbergh, & Manlove, 2005). Regulatieve processen zijn de belangrijkste motor in het proces van onderzoekend leren (Njoo & De Jong, 1993). In de planningsfase stelt de lerende een strategie op over de uitvoering van het proces van onderzoekend leren. Deze strategie zou doelen en subdoelen moeten hebben die bereikt kunnen worden door een bepaalde strategie te volgen. Bij het uitvoeren van de experimenten moet de lerende zijn voortgang monitoren zodat de lerende zijn strategie indien noodzakelijk kan aanpassen. In de evaluatiefase reflecteert de lerende op zijn voortgang en past zijn bevindingen toe op het probleem waarmee hij geconfronteerd wordt.

De coördinatie van theorie en bewijs is kenmerkend voor de meeste benaderingen van wetenschappelijk denken (Klahr, 2000; Klahr, Fay & Dunbar, 1993; Koslowski, 1996; Zimmerman, 2000). Dit proces wordt ook wel datareading genoemd (Kuhn, 2010), waarbij de representatie en de ontcijfering van het bewijs apart van de theorie in beschouwing wordt genomen. Kuhn (2010) beschrijft de coördinatie van theorie en bewijs als een voortgaand ontwikkelingsproces. Onderzoek bij oudere kinderen en adolescenten bevat aanzienlijk bewijs van moeilijkheden op het gebied van coördineren van theorie en bewijs. Zoals Klahr (2000) al eerder concludeerde, is de coördinatie van theorie en bewijs geen afzonderlijke vaardigheid die op een bepaald punt in de ontwikkeling naar boven komt. Het kan bereikt worden door opeenvolgende hogere niveaus van complexiteit in het denkproces over de coördinatie van theorie en bewijs te doorlopen over een langere periode van ontwikkeling. Toch zijn de spreiding en de variabiliteit in de vaardigheid tot wetenschappelijk denken bij volwassenen opvallend (Kuhn, Amsel, & O'Loughlin, 1988; Kuhn & Pease, 2006).

Kuhn (2010) maakte gebruik van gegevens uit micro-genetische onderzoeken, waarbij individuen in meerdere sessies geobserveerd werden in hun proces van kennisverwerving bij het uitvoeren van eenzelfde taak. Bij micro-genetisch onderzoek wordt intensief geobserveerd gedurende enkele weken of maanden, in een snel veranderende leerperiode (Lavelli, Pantoja, Hsu, Messinger, & Fogel, 2005; Siegler, 2006). Kuhn (2010) heeft meerdere micro-genetische onderzoeken samengevat (waaronder Kuhn, Black, Keselman, & Kaplan, 2000; Kuhn & Pease, 2008) waarbij uitingen van deelnemers, variërend van 9 tot 14 jaar, werden ingedeeld in categorieën om een ontwikkeling van het wetenschappelijk denken te kunnen beschrijven in verschillende fasen in het proces van het onderzoekend leren. Een belangrijke bevinding is dat leerlingen uiteenlopende strategieën toepassen bij kennisverwervingstaken. De selectie en het toepassen van deze strategieën (op metacognitief niveau) ontwikkelen zich gedurende de tijd, naar een meer frequent gebruik van meer adequate en gevorderde strategieën (Kuhn, 2010).

Het steeds invloedrijkere construct metacognitie wordt door Kuhn (2000) in een ontwikkelingsraamwerk geplaatst. Het bewustzijn van jonge kinderen van hun eigen en andermans mentale functies vormt het startpunt van dit ontwikkelingsproces, dat kan uitmonden in de beheersing van complexe metacognitieve vaardigheden tegen de tijd dat de volwassenheid is bereikt. Als belangrijke metacognitieve ontwikkelingsdoelen beschrijft Kuhn (2010): ten eerste het bewust worden van hoe en wat men leert en ten tweede het consistent leren toepassen van strategieën, waarbij minder adequate strategieën plaatsmaken voor meer adequate strategieën door middel van inhibitie en versterking.

In het huidige onderzoek proberen we ondersteuning te vinden voor deze bevindingen van Kuhn (2010) betreffende strategieontwikkeling en daarnaast proberen we deze strategieontwikkeling te kunnen generaliseren naar een oudere leeftijdsgroep (bachelor studenten). Als eerste hypothese veronderstellen we dat het gebruik van metacognitieve processen, die sterk samenhangen met de leerprestaties, in frequentie zal toenemen met de leeftijd.

Metacognitieve vaardigheden

Metacognitie is een breed begrip waarbij doorgaans een onderscheid wordt gemaakt tussen metacognitieve kennis en metacognitieve vaardigheden (Alexander et al., 1995; Veenman & Elshout, 1999). Het eerste concept, metacognitieve kennis, verwijst naar de declaratieve kennis van een persoon over het samenspel van persoonlijke kenmerken, taakkenmerken en beschikbare strategieën in een leersituatie (Flavell, 1979). Metacognitieve vaardigheden, aan

de andere kant, bevatten procedurele kennis die nodig is voor de daadwerkelijke regulatie van leeractiviteiten en de controle hierover (Flavell, 1992; Kluwe, 1987). Taakanalyse, planning, monitoring, controle en reflectie zijn manifestaties van deze metacognitieve vaardigheden. Volgens Veenman, Elshout en Meijer (1997) hebben metacognitieve vaardigheden een hoge onderlinge afhankelijkheid. Een zorgvuldige oriëntatie op de taak kan helpen om op relevante informatie te focussen voor het opbouwen van een adequate representatie van de taak. Zodoende kan een gedetailleerd actieplan worden opgesteld met doelen en aanwijzingen voor de uit te voeren leeractiviteiten. Zo'n uitgebreid actieplan kan vervolgens worden gebruikt voor de controle van het leerproces gedurende de uitvoering. Systematisch te werk gaan volgens het plan kan de lerende in staat stellen om voortgang bij te houden. Deze voortgangscontrole (monitoring) is nodig voor het opsporen van fouten en verkeerde procedures, en wordt effectiever uitgevoerd in de context van een dergelijk actieplan. Tenslotte zullen evaluatie en reflectie zoals het trekken van conclusies, samenvatten en leren van ervaringen, nuttiger zijn als ze op duidelijk te volgen activiteiten gebaseerd zijn (Veenman, Elshout, & Meijer, 1997).

De termen metacognitie en zelfregulatie zijn in onderwijskundige literatuur vaak uitwisselbaar (Veenman, 2007), ondanks dat hun conceptuele oorsprong en theoretische achtergrond behoorlijk verschillen (Dinsmore, Alexander, & Loughin, 2008; Fox & Riconscente, 2008). Metacognitieve theorieën vinden hun oorsprong in de ontwikkelingspsychologie met als voorlopers Piaget (Inhelder & Piaget, 1958) en Flavell (1970). Zelfregulerend leren aan de andere kant vloeit zowel voort uit de metacognitieve benaderingswijze als uit Bandura's theorie van zelfregulatie, hoewel met een sterke nadruk op de regulatie van het leerproces en de leeruitkomsten. De theorie van zelfregulerend leren tracht cognitieve, motivationele en contextuele factoren van leren te integreren (Dinsmore, et al., 2008). Onderzoekers van metacognitie beschouwen zelfregulatie als een ondergeschikte component van metacognitie (bv. Brown & DeLoache, 1978; Kluwe, 1987), terwijl andere onderzoekers zelfregulatie zien als een concept ondergeschikt aan metacognitie (bv. Winne, 1996; Zimmerman, 1995). Volgens het laatste sociaal-emotionele perspectief bestaat er naast cognitieve regulatie ook motivationele en affectieve regulatie. Volgens Dinsmore, et al. (2008) is de grens tussen beide theorieën in de loop der tijd vervaagd. In deze these wordt het metacognitieve perspectief gevolgd.

Tussen cognitieve en metacognitieve vaardigheden bestaat een complexe relatie. De meeste opvattingen over metacognitie hebben gemeenschappelijk dat zij het perspectief nemen van "higher-order cognition about cognition". Er is een vertegenwoordiger van hogere

orde die het cognitieve systeem overziet en controleert, terwijl deze tegelijkertijd deel uitmaakt van dit systeem. Dit is het klassieke homunculus probleem: Je kunt jezelf niet in tweeën splitsen waarbij de ene denkt terwijl de ander hem observeert terwijl hij denkt (Veenman et al, 2006). In termen van metacognitieve vaardigheden, iemand kan niet bezig zijn met planning zonder cognitieve activiteiten uit te voeren, zoals het bedenken van probleemoplossende stappen en de volgorde hiervan. Ook kan iemand niet de uitkomst van een berekening controleren zonder de uitkomst te vergelijken met een schatting of door de uitkomst op een andere manier te berekenen. Als metacognitie wordt gezien als (de kennis van) een verzameling van zelfinstructies voor het reguleren van de uitvoering van een taak, dan is cognitie het voertuig van deze zelfinstructies. Cognitieve activiteiten zijn op hun beurt het onderwerp van metacognitie, bijvoorbeeld van het voortdurende proces van monitoring en evaluatie. Dit circulaire proces van metacognitieve en cognitieve activiteiten maakt het moeilijk om ze uit elkaar te houden in de bepaling van metacognitie. Metacognitie is niet altijd expliciet te horen of te zien gedurende de uitvoering van een taak. In plaats daarvan moet het afgeleid worden uit bepaalde cognitieve activiteiten (Veenman et al., 2006).

Nelson (1996) maakt onderscheid in een “object-niveau”, waarin de cognitieve activiteiten plaatsvinden en een “meta-niveau”, dat het object-niveau aanstuurt. Twee algemene stromen van informatie tussen beide niveaus worden verondersteld. Informatie over de status van het object-niveau wordt overgedragen aan het meta-niveau door monitoring, terwijl instructies van het meta-niveau worden overgedragen naar het object-niveau door controlerende processen. Bijna alle metacognitieve vaardigheden dragen ook een cognitieve gedragscomponent in zich. Het genereren van een hypothese, het ontwerpen van een experiment en het trekken van een conclusie zijn cognitieve activiteiten op het object-niveau. Maar die activiteiten zouden niet plaatsvinden zonder aansturing door een regulerend mechanisme op het meta-niveau. Het feit dat men een hypothese genereert alvorens aan de slag te gaan met experimenteren, betekent dat er een meta-beslissing (ook wel metacognitieve zelfinstructie) aan vooraf moet zijn gegaan die aanzette tot hypothese genereren. Hetzelfde geldt voor de Control-of-Variables Strategy (CVS) wat inhoudt dat alleen de te onderzoeken variabele gemanipuleerd wordt, terwijl de andere variabelen constant worden gehouden, zodat valide conclusies kunnen worden getrokken over het effect van onafhankelijke variabelen op de afhankelijk variabele (Chen & Klahr, 1999). CVS is bij uitstek een indicator voor plannen. Het trekken van conclusies is een indicator voor de metacognitieve beslissing om de gevonden resultaten geordend op te slaan in het geheugen. Indien er geen metacognitieve zelfinstructie aan vooraf zou zijn gegaan, dan treedt de cognitieve activiteit van hypothesen

genereren (voorafgaand aan experimenteren), experimenten ontwerpen (voorafgaand aan het uitvoeren van experimenten) en conclusies trekken (na afloop van een experiment of reeks experimenten) in het geheel niet op of in ieder geval niet op systematische wijze. Het is van belang niet alleen te kijken naar het cognitieve aspect van deze drie activiteiten (genereren van hypothesen, experimenteren en evalueren van bewijs) op het object-niveau, maar eveneens naar het meta-niveau mechanisme dat deze cognitieve activiteit aanstuurt.

Metacognitieve vaardigheden kunnen op verschillende manieren worden gemeten. Een onderscheid kan worden gemaakt tussen on-line en off-line methodes (Van Hout-Wolters, 2000; Veenman, 2005). Off-line methodes worden of voor of na de uitvoering van een taak gepresenteerd, terwijl on-line methodes tijdens de uitvoering van een taak worden verkregen. On-line methodes zijn meer voorspellend voor leerprestaties vergeleken met off-line methodes (Veenman, 2005). De scores op bijvoorbeeld vragenlijsten voor metacognitieve activiteiten of strategiegebruik corresponderen nauwelijks met het gemeten gedrag tijdens uitvoering van de taak (Veenman, 2005; Veenman, Prins & Verheij, 2003). Een methode die geschikt is om metacognitieve vaardigheden in beeld te brengen is een multi-methode, waarbij meer dan één techniek voor dataverzameling wordt gebruikt, zoals bij de analyse van hardopdenkprotocollen gecombineerd met logfilegegevens.

Veenman, et al. (2004) veronderstellen dat een klein aantal metingen volstaat om een indicatie te kunnen geven van het niveau van metacognitieve vaardigheden van een individu. Zoals eerder is beargumenteerd door Veenman, et al. (1997) hebben metacognitieve vaardigheden een hoge onderlinge afhankelijkheid. Om deze reden gebruikten Veenman, et al. (2004) als indicatie van metacognitieve vaardigheden het aantal keren dat proefpersonen terug scrollden naar eerdere experimenten en het aantal variabelen dat zij per experiment varieerden. Meer dan één variabele per experiment veranderen werd gezien als een indicator voor zwak systematisch gedrag. Het variëren van slechts één variabele ten opzichte van het vorige experiment (CVS, Chen & Klahr, 1999) werd opgevat als een representatie van goed systematisch gedrag en controle over de experimenten (Schauble, Glaser, Raghavan, & Reiner, 1991). Ten behoeve van de generalisatie van de bruikbaarheid van CVS en scrolgegevens uit de logfiles werden deze indicatoren van metacognitieve vaardigheden door Veenman, et al. (2004) vergeleken met een klein deel van de gegevens uit hardopdenkprotocollen (10% van de planttaak en 5% van de ottertaak; zie de beschrijving van de taken bij de Methode). Deze hardopdenkprotocollen werden uitgetypt en vervolgens beoordeeld met als criteria oriëntatie, systematisch gedrag, evaluatie en elaboratie. Veenman, et al. (2004) concludeerden uit deze vergelijking dat scrollen en het toepassen van CVS

geschikt zijn om het niveau van metacognitieve vaardigheden te bepalen. Echter, het merendeel van de gegevens van de hardopdenkprotocollen werd niet geanalyseerd en was om deze reden niet gedekt. Voor de geldigheid van de bevindingen van Veenman, et al. (2004) is het van belang een groot deel van de hardopdenkprotocollen te analyseren om hun onderzoek naar metacognitieve vaardigheden te onderbouwen. Tevens kan met behulp van deze analyse meer inzicht verkregen worden in het ontwikkelingsniveau van metacognitieve vaardigheden. Evers (2009) heeft al eerder een deel van de hardopdenkprotocollen van kinderen uit groep 6 en 8 van het basisonderwijs geanalyseerd.

In een review onderzoek van Wang et al. (1990) bleek metacognitie de meest krachtige voorspeller van leren te zijn. Metacognitie speelt een belangrijke rol bij het leren maar er zijn nog vele onopgeloste problemen rond het begrip (Veenman, van Hout-Wolters, & Afflerbach, 2006). Ondanks dat er consequente bevestiging is gevonden voor het belang van metacognitie, is inconsistentie in de conceptualisatie van het construct kenmerkend. In Veenman, et al. (2006) wordt deze inconsistentie beschreven vanuit diverse disciplines en achtergronden. Eén van de issues hierbij is de complexe relatie tussen metacognitie en intelligentie. In de volgende paragraaf wordt de relatie met intelligentie nader toegelicht.

Relatie tussen intelligentie en metacognitieve vaardigheden

Over intelligentie bestaan veel verschillende opvattingen (o.a. Brody, 1992; Carroll, 1993; Sternberg, 1990). Hier volgen we de pragmatische zienswijze van Elshout (1983) waarbij intelligentie wordt voorgesteld als een menselijke cognitieve gereedschapskist die cognitieve basishandelingen bevat. De inhoud en kwaliteit van deze gereedschapskist verschilt per persoon en is afhankelijk van biologische invloeden, zoals bijvoorbeeld erfelijkheid, en omgevingsinvloeden, zoals bijvoorbeeld onderwijs (Van der Stel & Veenman, 2008). Volgens Busato, Prins, Elshout en Hamaker (2000) kan intelligentie gezien worden als een belangrijke voorspeller van leren, daar het positief gerelateerd is aan academisch succes. Veenman, et al. (2004) vonden eveneens een positieve relatie tussen intelligentie en leerprestaties bij opdrachten voor onderzoekend leren. Ook Dabbous (2014) vond een positieve correlatie tussen intelligentie en leerprestaties bij taken voor onderzoekend leren bij leerlingen in het eerste jaar van het voortgezet onderwijs. In het huidige onderzoek proberen we de invloed van intelligentie op de leerprestaties bij onderzoekend leren nader te verhelderen voor twee leeftijdsgroepen. Er worden onder meer positieve correlaties verwacht tussen intelligentieniveau en leerprestaties.

Over de relatie tussen intelligentie en metacognitieve vaardigheden bestaat enige discussie. Veenman en Spaans (2005) beschrijven drie verschillende modellen om de relatie tussen intelligentie en metacognitie te duiden: metacognitie als onderdeel van intelligentie (intelligence model), metacognitie als losstaand van intelligentie (independency model) en metacognitie als deels gerelateerd aan intelligentie met een meerwaarde (mixed model). Van der Stel en Veenman (2008) hebben in een onderzoek bij 12 jarigen ondersteuning gevonden voor het mixed model. Intelligentie en metacognitieve vaardigheden zijn middelmatig met elkaar gecorreleerd en bovendien leveren intelligentie en metacognitieve vaardigheden beiden hun eigen unieke bijdrage aan de leerprestaties. Deze resultaten komen overeen met de resultaten uit onderzoek bij oudere leeftijdsgroepen (Veenman, 1993; Veenman & Verheij, 2003; Veenman & Beishuizen, 2004; Veenman et al., 2004). Voor het mixed model is dus bevestiging gevonden bij zowel volwassenen als adolescenten. Als tweede hypothese verwachten we opnieuw ondersteuning te vinden voor het mixed model voor de twee leeftijdsgroepen, waarbij metacognitie deels gerelateerd is aan intelligentie, maar met elk een eigen meerwaarde voor leerprestaties.

Vanuit een ontwikkelingsperspectief zijn door Alexander, et al. (1995) drie belangrijke hypothesen opgesteld. Ten eerste de ‘ceiling’ hypothese waarbij metacognitie bij jonge kinderen geassocieerd is met de ontwikkeling van intelligentie maar waarbij dit effect zal verdwijnen bij het ouder worden. Ten tweede de ‘acceleration’ hypothese waarbij het effect van intelligentie op metacognitie toeneemt met de leeftijd, er is sprake van wederzijdse versterking. Als laatste de ‘monotonic development’ hypothese waarbij er sprake is van geleidelijke groei van metacognitie met leeftijd samen met intellectuele groei, er is dus sprake van parallelle ontwikkeling. Alexander, et al. (1995) vonden bevestiging voor de ‘monotonic development’ hypothese en lieten zien dat metacognitieve kennis zich ontwikkelt in een geleidelijke stijgende lijn gedurende de schooljaren, parallel met de ontwikkeling van intelligentie van studenten. De impact van intelligentie neemt niet toe en neemt niet af tijdens de jaren. Veenman et al. (2004) vonden overeenkomstige resultaten voor de ontwikkeling van metacognitieve vaardigheden in relatie met intelligentie. Van der Stel en Veenman (2008) hebben gevonden dat het niveau van metacognitieve vaardigheden toeneemt met de leeftijd en daarbij hebben zij geen significante interacties gevonden tussen intelligentie en leeftijd zodat zowel de ‘ceiling’ hypothese als de ‘acceleration’ hypothese zijn verworpen. Veenman en Spaans (2005) concluderen dat metacognitieve vaardigheden zich naast intelligentie ontwikkelen, maar niet volledig afhankelijk zijn van intelligentie. Bovendien voorspellen metacognitieve vaardigheden leerprestaties beter dan intelligentie. Het onderzoek van

Veenman en Spaans (2005) heeft echter als beperking dat het een kleine onderzoeksgroep betreft en slechts de groei bekijkt tussen twee leeftijdsgroepen (12- en 15- jarigen). Ondanks de nauw verbonden relatie met cognitieve processen, kunnen metacognitieve vaardigheden niet gelijk worden gesteld met intelligentie (cf. Sternberg, 1990). In een overzicht van verschillende onderzoeken door Veenman et al. (2014) van deelnemers van verschillende leeftijden, verschillende soorten taken en verschillende domeinen wordt verondersteld dat intelligentie gemiddeld genomen .45 correleert met metacognitieve vaardigheid. Intelligentie alleen verklaart 10% van de variantie in leren, metacognitieve vaardigheden alleen verklaart 18% van de variantie in leren, terwijl beide voorspellers samen nog 22% in de variantie delen in leren (voor een overzicht, zie Veenman, et al., 2004; Veenman & Spaans, 2005; Veenman, 2008).

Als derde hypothese verwachten we in het huidige onderzoek bevestiging te vinden van deze parallelle ontwikkeling, waarbij er sprake is van geleidelijke intellectuele groei samen met metacognitie in twee leeftijdsgroepen. We verwachten net zoals bij Veenman en Spaans (2005) dat metacognitieve vaardigheden en intelligentie toenemen met de leeftijd en dat hierbij de invloed van intelligentie niet toeneemt of afneemt met de leeftijd. Van de Stel en Veenman (2010) onderzochten de ontwikkeling van metacognitie nader met een longitudinaal onderzoek bij leerlingen van 12 tot 14 jaar, in twee opeenvolgende schooljaren. Er werd zowel kwantitatieve als kwalitatieve groei gevonden van metacognitieve vaardigheden. Verder laten de resultaten zien dat metacognitieve vaardigheden (deels) onafhankelijk van intelligentie bijdragen aan leerprestaties.

Onderzoeksvragen en hypotheses

De onderzoeksvragen die in dit onderzoek centraal staan zijn: Welke verschillen bestaan er in het ontwikkelingsniveau van metacognitieve vaardigheden tussen leerlingen uit groep 2VWO en universitaire studenten bij onderzoekend leren en welke relaties bestaan er tussen metacognitieve vaardigheden, intelligentie en leerprestaties? Voor de beantwoording van deze onderzoeksvragen zijn een aantal hypotheses opgesteld.

Als eerste hypothese veronderstellen we dat het gebruik van metacognitieve leerprocessen, die sterk samenhangen met de leerprestaties, in frequentie zal toenemen met de leeftijd.

Als tweede hypothese verwachten we opnieuw ondersteuning te vinden voor het mixed model van Veenman et al. (2004) voor de twee leeftijdsgroepen, waarbij metacognitie deels gerelateerd is aan intelligentie met elk een eigen meerwaarde voor leerprestatie.

Als derde hypothese, verwachten we in het huidige onderzoek bevestiging te vinden van de geleidelijke groei van metacognitie samen met intellectuele groei (parallele ontwikkeling) in de twee leeftijdsgroepen. We verwachten net zoals bij Veenman en Spaans (2005) dat metacognitieve vaardigheden toenemen met de leeftijd en dat hierbij de invloed van intelligentie niet toeneemt of afneemt met leeftijd.

Methode

De in het onderzoek verwerkte data zijn verzameld door Wilhelm (2001) ten behoeve van zijn proefschrift. In het huidige onderzoek werden de gegevens van hardopdenkprotocollen van 2VWO-ers en de universitaire studenten geanalyseerd. Deze protocolanalyse sloot aan bij de eerder uitgevoerde onderzoeken van Wilhelm (2001) en Veenman et al. (2004). Met behulp van de analyse van hardopdenkprotocollen werd onderzocht welke metacognitieve leerprocessen gerelateerd zijn aan leerprestaties en of er verschillen tussen de leeftijdsgroepen bestaan. Het onderzoek van Veenman, et al. (2004) werd deels gerepliceerd om de door hen veronderstelde relatie tussen intelligentie, metacognitieve vaardigheden en leerprestaties, volgens het mixed model, nader te onderzoeken. Veenman, et al. (2004) maakten gebruik van logfile gegevens voor het meten van metacognitieve vaardigheid. Tijdens hun onderzoek werden de proefpersonen gevraagd om hardop te denken. In het huidige onderzoek werden deze protocollen van de groepen 2VWO en universitaire studenten geanalyseerd om op deze wijze een maat voor metacognitieve vaardigheden te verkrijgen die rechtstreeks afgeleid is van het leergedrag en het denkproces van de lerende. De gebruikte scoringsmethode ontworpen door Wilhelm & Beishuizen (2003) wordt beschreven in de Methode in paragraaf *Metacognitieve leerprocessen/vaardigheden*.

Proefpersonen

De proefpersonen waren 30 leerlingen van 2VWO van een school voor voortgezet onderwijs in het stedelijk gebied van Den Haag (gemiddelde leeftijd 14.1, SD: 0.45) en 30 bachelor studenten van de Universiteit Leiden, van de Faculteit Sociale Wetenschappen (gemiddelde leeftijd 22.5 jaar, SD: 3.37). In de 2VWO groep werden leerlingen met ernstige leerproblemen uitgesloten van deelname aan het onderzoek. De sociaal economische status werd geschat op iets hoger dan gemiddeld, de verdeling in geslacht was in de groep 2VWO redelijk in balans (18 jongens en 12 meisjes). Ouders werden per brief op de hoogte gesteld van de inhoud van het onderzoek en gaven toestemming voor de deelname van hun kinderen.

Aangezien de universiteiten in Nederland verplicht zijn alle studenten toe te laten die het hoogste niveau voorgezet onderwijs hebben afgerond, zonder verdere toelatingscriteria, zal de deelnemende groep een relatief brede range van intelligentie vertegenwoordigen (zie Veenman, 1993). De vrouwelijke studenten waren in de meerderheid, in lijn met de verdeling van geslacht voor studenten sociale wetenschappen. Alle studenten namen vrijwillig deel aan het onderzoek.

Van de in totaal 60 proefpersonen was bij 7 proefpersonen maar één taak beschikbaar voor analyse. De redenen hiervoor waren slechte kwaliteit of het ontbreken van de audio-opname van de tweede taak. In totaal werden van 53 proefpersonen de hardopdenkprotocollen van twee taken geanalyseerd waarvan 27 uit de groep 2VWO en 26 uit de groep studenten. Doordat van één student een score ontbrak in de logfile, zijn in totaal de scores van 52 proefpersonen (27 uit de groep 2VWO en 25 uit de groep studenten) meegenomen in de verdere analyse.

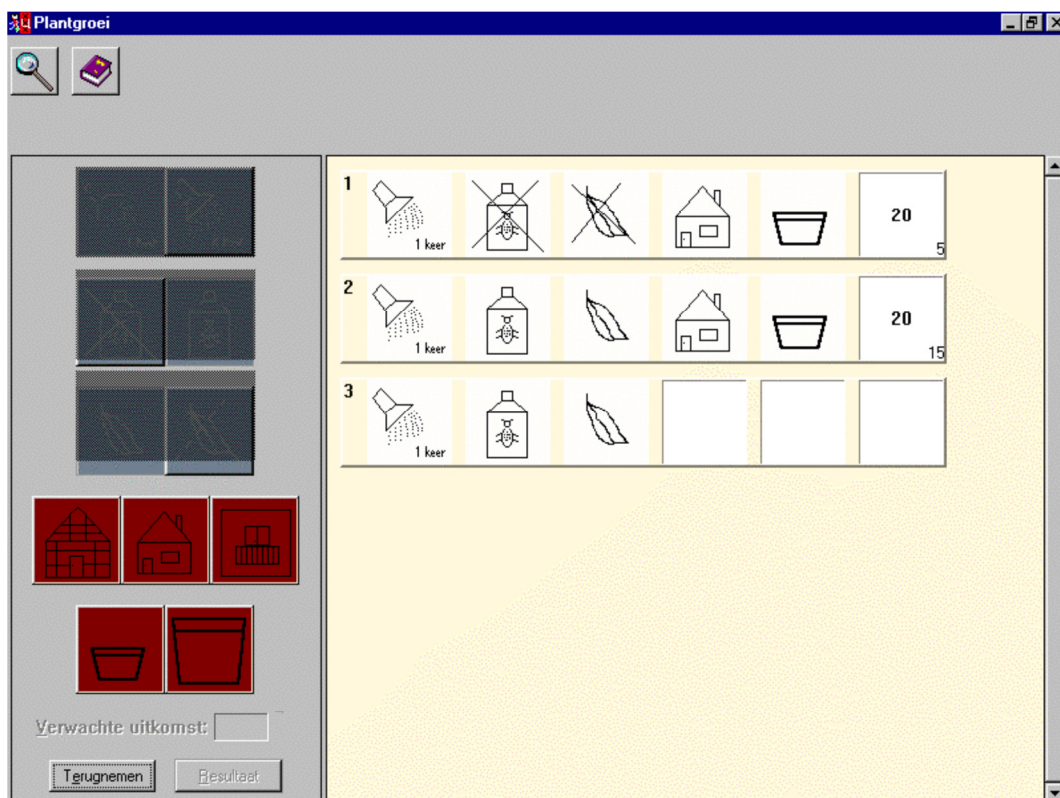
Materialen

Onderzoekend leren taken

De proefpersonen werkten aan vier simulatietaken voor onderzoekend leren die gecreëerd zijn in een leeromgeving die FILE wordt genoemd (Flexible Inquiry Learning Environment; Hulshof, Wilhelm, Beishuizen, & Van Rijn 2005). FILE is een gecomputeriseerde leeromgeving voor zelfgestuurd inductief leren, een andere benaming voor onderzoekend of ontdekkend leren. In FILE kunnen leerlingen van diverse leeftijdsgroepen (van 8 jaar tot volwassenen) experimenten uitvoeren om effecten te ontdekken van meerdere onafhankelijke variabelen op een afhankelijke variabele. Het model dat de onderliggende relaties weergeeft tussen de onafhankelijke variabelen en de afhankelijke variabele was voor alle taken hetzelfde. Proefpersonen werden geïnstrueerd hoe ze een niveau van een onafhankelijke variabele in een rij op het scherm konden plaatsen door op de corresponderende pictogrammen te klikken, hoe ze een voorspelling konden invoeren en hoe ze het resultaat konden bekijken. Er werd gewezen op de scrol- en selectiefunctie om eerdere experimenten te kunnen bekijken.

In het huidige onderzoek zijn twee verschillende simulatieopdrachten geanalyseerd uit het domein biologie en aardrijkskunde. In het onderzoek van Wilhelm (2001) werd gebruik gemaakt van vier simulatieopdrachten, waarvan werd besloten alleen de eerst afgenomen taak van een domein nader te analyseren. In het huidige onderzoek werd gebruik gemaakt van een Plant taak (biologie) en een Otter taak (aardrijkskunde).

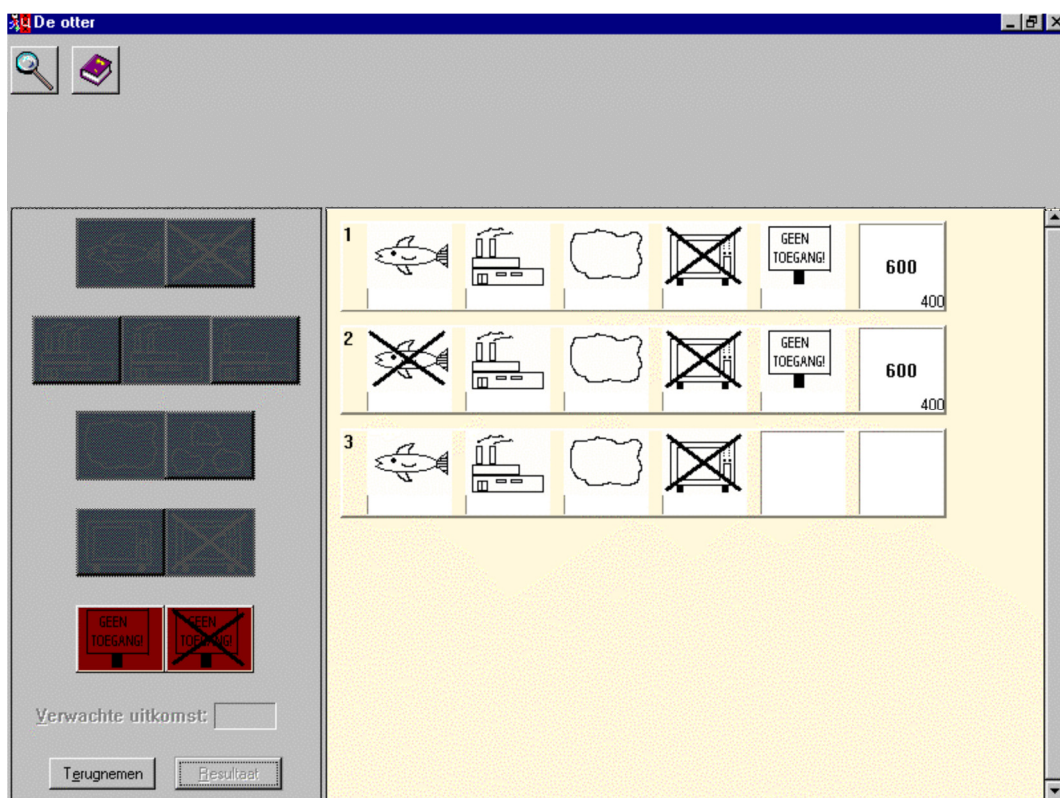
In de Plant taak (zie Figuur 1) werden de proefpersonen uitgenodigd om uit te zoeken hoe vijf factoren de groei van een plant beïnvloeden. Deze factoren zijn de grootte van de bloempot (klein of groot), het aantal keren water geven (één keer of twee keer per week), de plaats waar de plant staat (huis, kas of balkon), dode bladeren in de bloempot leggen (wel of geen) en bladluismiddel gebruiken (wel of geen). De lengte van de plant is de afhankelijke variabele en kan de volgende waarden aannemen: 5, 10, 15, 20 of 25 cm. Het wel of niet neerleggen van dode bladeren en het wel of niet gebruiken van bladluismiddel heeft geen effect (irrelevant). Eén of twee keer water geven en de grootte van de bloempot hebben een interactie effect. In een kleine bloempot heeft de plant maar één keer per week water nodig, dan wordt de plant het grootst. Krijgt de plant twee keer per week water in een kleine pot dan wordt hij heel klein. In een grote bloempot maakt het niet uit of de plant één of twee keer per week water krijgt, de plant wordt dan gemiddeld groot. De plaats waar je de plant neerzet, heeft een curvilinear hoofdeffect, waarbij één van de drie niveaus van de variabele een effect heeft (huis) en de andere twee niet (kas of balkon). De plant groeit namelijk in een huis minder goed dan in een kas of balkon.



Figuur 1. Interface van de Plant taak.

In de Otter taak (Figuur 2) werden de proefpersonen gevraagd naar de invloed van vijf factoren op de groei van de otterpopulatie in Nederland. Deze factoren zijn het leefgebied

(één groot gebied of meerdere kleinere gebieden), de toegang tot het gebied voor publiek (wel of geen), het aantal afvallozingen van fabrieken (één, twee of drie), informatie geven over de otter op televisie (wel of geen) en vis bijvoeren (wel of niet). De grootte van de otterpopulatie is de afhankelijke variabele en kan de volgende waarden aannemen: 100, 250, 400, 450 of 600 otters. Het wel of geen informatie geven over de otter op televisie en het wel of geen vis bijvoeren heeft geen effect (irrelevant). Eén groot leefgebied of meerdere kleine leefgebieden en de toegang voor publiek hebben een interactie effect. Wanneer de otters in een groot gebied leven zonder toegang voor publiek, dan wordt de otter populatie het grootst. Wordt er wel toegang gegeven in een groot gebied dan wordt de otterpopulatie heel klein. In meerdere kleine gebieden maakt het niet uit of er wel of geen toegang wordt gegeven voor publiek, de otterpopulatie wordt dan gemiddeld groot. Het aantal afvallozingen van fabrieken heeft een curvilinear hoofdeffect, bij drie afvallozingen blijft de otterpopulatie kleiner dan bij één of twee afvallozingen.



Figuur 2. Interface van de Otter taak.

Intelligentiematen

Voor het bepalen van intelligentie zijn de onderstaande testen gekozen: (1) Getallen series (Elshout, 1976) voor inductief redeneren; (2) Concrete Syllogismen voor deductief redeneren; (3) Verborgene Figuren Test (Flanagan, 1951) voor ruimtelijk inzicht; (4) Ruimtelijk inzicht

(Differentiële Aanleg Test, Evers & Lucassen, 1983) eveneens voor ruimtelijk inzicht en (5) een test met redactiesommen die items bevatten van Elshout (1976) en van Mommers, de Visser en Warnars (1977). Om er zeker van te zijn dat de testen bij de verschillende leeftijdsgroepen uit het onderzoek (groep 6 en 8, klas 2VWO en universitaire studenten) gebruikt konden worden, zijn sommige testen aangepast om vloer- en plafondeffecten te voorkomen (zie Wilhelm, 2001). De keuze voor deze tests is gebaseerd op de bevindingen van Elshout (1976). Hij correleerde de scores van verschillende (sub)testen voor intelligentie met de leerprestaties op een taak voor onderzoekend leren en selecteerde hiervan de vijf testen die het hoogste correleerden met deze leerprestaties. De totale intelligentiescore werd verkregen door het berekenen van het ongewogen gemiddelde van de gestandaardiseerde scores van de vijf subtests (Cronbach's $\alpha = 0.83$). De intelligentiescores werden niet omgezet in leeftijdsgebonden IQ scores omdat ze bedoeld waren om een intellectuele groei te kunnen meten parallel met metacognitieve vaardigheden (Veenman et al., 2004).

Procedure

De intelligentietesten werden afgenomen in een groep. Vervolgens maakten de proefpersonen in een individuele sessie de onderzoekend leren taken achter de computer. De eerste sessie begon met een instructie op de computer, waarbij gebruik werd gemaakt van een andere FILE taak over een dagelijkse bezigheid (de Peter taak, zie Wilhelm 2001). Voordat de proefpersonen aan de eerste leertaak begonnen, werd gevraagd om te laten zien hoe de interface werkte. Wanneer de proefpersonen een bepaalde functie niet konden demonstreren, herinnerde de testleider hen aan deze functie. Alle computerhandelingen die door de proefpersonen werden uitgevoerd, zijn opgeslagen in een logfile. Tijdens de taken werden daarnaast audio opnames gemaakt van de proefpersonen terwijl zij hardop nadachten.

In het begin van een taak kregen de proefpersonen een korte uitleg over het doel van de taak en ze werden geïnformeerd over de vragen die na afloop van de taak gesteld zouden worden. Deze vragen hadden betrekking op de effecten van de onafhankelijke variabelen en hadden de volgende vorm: “Welk verschil denk je dat de (onafhankelijke variabele) maakt voor de (afhankelijke variabele)?” Er werd aan de proefpersonen gevraagd om hardop te denken. Een bandopname van twee minuten van iemand die hardop dacht, werd als instructie afgespeeld. Voorafgaand aan de uitvoering van een taak werden de preconcepties van de proefpersonen gemeten door te vragen: “Wat denk je nu dat de (onafhankelijke variabele) uitmaakt voor de (afhankelijke variabele)?”

De proefpersonen in het oorspronkelijke onderzoek voerden twee taken uit in het biologie domein en twee taken in het aardrijkskunde domein. Ze werden willekeurig verdeeld over twee condities die verschilden in of ze eerst de biologietaken gingen maken of eerst de aardrijkskundetaken. Met de protocolanalyse werd in het huidige onderzoek alleen de eerste taak van elk domein geanalyseerd. De proefpersonen mochten maximaal 35 minuten aan één taak werken maar ze konden stoppen wanneer ze het gevoel hadden dat ze klaar waren. Wel werden de proefpersonen door de testleider gestimuleerd minimaal tien experimenten uit te voeren. Ze mochten ook aantekeningen maken. Zoals gezegd werden de proefpersonen gevraagd om hardop te denken. Ter ondersteuning stelde de testleider vragen met het oog op de onderzoeksplannen van de proefpersonen ('Wat ga je proberen uit te vinden?'), de voorspellingen ('Wat denk je dat de uitkomst zal zijn?') en de conclusies ('Wat heb je uitgevonden?'). Na afloop van alle experimenten werd per variabele de volgende vraag gesteld: "Welk verschil denk je dat de (onafhankelijke variabele) maakt voor de (afhankelijke variabele)?" De antwoorden op deze begripsvragen werden vergeleken met het onderliggende model van de betreffende taak en leverde een begripsscore op.

Scoring en Data-analyse

Leerprestaties

In het huidige onderzoek werd als maat voor leerprestatie een begripsscore per proefpersoon en per taak berekend. De basis van de scoringsmethode (Wilhelm, 2001) bestaat uit negen uitspraken die de correcte effecten van de onafhankelijke variabelen in de leertaak weergeven. Twee uitspraken hebben betrekking op de non-causale verbanden, drie uitspraken hebben betrekking op het hoofdeffect en de overige vier uitspraken hebben betrekking op het interactie-effect. Twee punten worden toegekend voor elke bewering die correspondeert met deze uitspraken, nul punten als de bewering incorrect of afwezig is. Als een proefpersoon de beperkende conditie niet benoemde bij het beschrijven van de effecten van interacterende variabelen, werd één punt toegekend. Een voorbeeld van een uitspraak over een interactie-effect is "één of twee keer water geven maakt niet uit *in een grote pot* (beperkende conditie)". Deze uitspraak levert twee punten op, als de grote pot niet wordt genoemd levert deze uitspraak slechts één punt op. De maximale score is 18 punten. Indien beide taken werden gemaakt, werd een gemiddelde begripsscore berekend over twee taken. Wanneer een proefpersoon slechts één van de twee taken had gemaakt, werd alleen de bijbehorende begripsscore gebruikt. De begripsscores werden zowel in het originele onderzoek van Wilhelm (2001) als in het huidige onderzoek bepaald, zodat zij overeenkwamen met de taken

die beschikbaar waren voor analyse van de hardopdenkprotocollen. Voor de scoring van de begripsscores zie Bijlage 1.

Metacognitieve leerprocessen/vaardigheden

Als theoretische basis voor de analyse van de hardopdenkprotocollen is het Scientific Discovery as Dual Search (SDDS) model gebruikt (Klahr & Dunbar, 1988). Met behulp van een scoringsprotocol werden leerprocessen geëvalueerd. Dit scoringsprotocol van Wilhelm (2001) werd ontworpen om het plangedrag, het opstellen van hypothesen, het testen van hypothesen en het trekken van conclusies te meten over de uitgevoerde experimenten. In het SDDS model van Klahr en Dunbar (1998) wordt het wetenschappelijk ontdekkingsproces gezien als een zoekproces in de hypotheseruimte (alle mogelijke hypothesen over een onderwerp of systeem) en in de experimentruimte (alle experimenten die uitgevoerd kunnen worden in het systeem). De hypotheseruimte bevat de kennis van een lerende over de relaties tussen de variabelen in het domein; de experimentruimte bevat alle mogelijke experimenten die uitgevoerd kunnen worden met de beschikbare middelen. Wetenschappelijk denken verloopt door in de hypotheseruimte te zoeken naar toetsbare hypothesen en door in de experimentruimte te zoeken naar een combinatie van experimenten om deze hypothesen te toetsen en het resultaat te evalueren om de hypothese te kunnen bevestigen of bij te stellen. Wanneer een lerende niet in staat is om toetsbare hypothesen te vinden in de hypotheseruimte dan kan deze zoeken in de experimentruimte door verkennende experimenten uit te voeren die kunnen helpen om nieuwe hypothesen te formuleren. De leerprocessen van het SDDS model zijn in vier hoofdcategorieën ondergebracht: hypothesen genereren, experimenten ontwerpen, interpretatie van data en regulatie van leren. Deze categorieën zijn grotendeels vergelijkbaar met de criteria van Veenman et al. (2004) die gebruikt werden om de kwaliteit van metacognitieve vaardigheden te beoordelen in de hardopdenkprotocollen. Deze zijn: oriëntatie (genereren van hypothesen voor elk experiment), plannen en systematisch gedrag (plannen van een reeks van experimenten, vermijden van onsystematisch gedrag zoals het variëren van twee onafhankelijke variabelen tussen twee opeenvolgende experimenten), evaluatie (opmerken en corrigeren van fouten) en elaboratie (conclusies trekken, uitkomsten met elkaar in verband brengen, verklaringen geven en samenvatten). Hieronder worden de door Wilhelm (2001) afgeleide hoofdcategorieën en leerprocessen van het SDDS model, zoals die in het huidige onderzoek werden gebruikt voor de scoring, toegelicht.

Bij het genereren van hypothesen werd gekeken of de lerende hypothesen opstelde. Een uitspraak werd als hypothese beschouwd als het veronderstelde effect van een specifieke variabele werd genoemd (bijvoorbeeld “Ik ga variabele X checken want ik denk dat deze variabele verschil uitmaakt”). Er werden hypothesen gescoord ten aanzien van één variabele (eenvoudige hypothesen) en hypothesen ten aanzien van twee variabelen waarin een gezamenlijk effect werd onderzocht (complexe hypothesen).

Bij het ontwerpen van de experimenten werden twee variabelen gemeten, het verkennen van de experimentruimte en het gebruik van CVS. Wat een lerende kan ontdekken, is afhankelijk van de informatie die een lerende verzamelt. Om deze reden werd het aantal uitgevoerde experimenten gescoord. Wanneer het aantal unieke experimenten dat uitgevoerd kan worden beperkt is, zal het percentage van het totaal aantal mogelijke experimenten dat wordt uitgevoerd, een reflectie zijn van de mate waarin de lerende de experimentruimte heeft verkend (Klahr & Dunbar, 1988). Bij het toepassen van CVS werd het aantal variabelen gemeten dat een lerende per experimentregel varieerde.

Bij de interpretatie van de data werd gevraagd wat de lerende had ontdekt. De conclusies van de lerenden werden vastgelegd (formuleren van inferenties). Hierbij werd een onderscheid gemaakt tussen inclusie inferenties (dat een variabele verschil maakt), exclusie inferenties (dat een variabele geen verschil maakt) en interactie inferenties (er is wel/geen interactie tussen variabele x en y) (zie Kuhn et al., 1995). Bij het formuleren van inferenties werd gekeken naar validiteit van deze conclusies, in hoeverre deze inferenties gemaakt konden worden op basis van het gegenereerde bewijs. Tevens werd het percentage van valide inferenties bepaald als voorspeller van positieve leeruitkomsten. De voorspellingfout was een andere manier om de vaardigheid van de lerende te meten bij de interpretatie van data. Aan lerenden werd gevraagd om de uitkomst van hun experimenten te voorspellen in grootte of aantal. De voorspellingsfout werd bepaald door het gemiddelde gestandaardiseerde verschil tussen de voorspelde uitkomst en de werkelijke uitkomst van een experiment.

De regulatie van het leren bestond uit het formuleren van plannen. Er werden verschillende plannen onderscheiden: inductieve plannen (plannen gericht op het onderzoeken van het effect van één variabele of van een combinatie van meerdere variabelen/interactie), deductieve plannen (plannen gericht op het checken van een effect van een variabele in een ander experiment of het voorspellen van een uitkomst door regels ten aanzien van alle variabelen te toetsen), plannen gericht op een specifieke uitkomst (indicatie voor engineering benadering (Schauble et. al, 1991)) en vage of ongerichte plannen (Wilhelm & Beishuizen,

2003). De exacte scoring van de vier hierboven beschreven categorieën staat in het scoringsprotocol (zie Bijlage 2).

Ter controle van de bruikbaarheid van het scoringsprotocol lieten Wilhelm en Beishuizen (2003) 20% van de hardopdenkprotocollen door twee onafhankelijke beoordelaars scoren om de betrouwbaarheid vast te stellen voor de leerprocesmaten en de begripsscores. In Evers (2009) was de interbeoordelaars-betrouwbaarheid voor de score van de leerprocesmaten van onderzoekend leren 82%. In het huidige onderzoek is gebruik gemaakt van hetzelfde scoringsprotocol als in Evers (2009). De protocollen van drie studenten en drie 2VWO-ers zijn door Wilhelm en Nijhuis met elkaar vergeleken tot om een eenduidige interpretatie van de categorieën te komen.

Het gemiddelde percentage van overeenstemming over de begripsscores, verkregen in het onderzoek van Wilhelm (2001), was 81%. In het huidige onderzoek zijn de begripsscores opnieuw bepaald voor de leeftijdsgroepen 14- en 23-jarigen. De overeenstemming met de begripsscores van Wilhelm (2001) was 92%. Daarnaast is uit het onderzoek van Wilhelm en Beishuizen (2003) gebleken dat het stellen van de hierboven beschreven vragen over onderzoeksplannen, voorspellingen en conclusies, geen effect had op de leerresultaten.

Om de scores op de hiervoor beschreven leerprocessen te kunnen bepalen, was het nodig om de ruwe data van de hardopdenkprotocollen te aggregeren. Dit werd zowel voor beide taken samen als per taak apart gedaan, zodat een totaalscore per leerprocesmaat en per taak kon worden berekend. Vervolgens zijn deze drie geaggregeerde bestanden samengevoegd met de algemene data van het onderzoek van Wilhelm (2001) waarin ook de intelligentiescores waren opgenomen. Er werd een gemiddelde (gewogen score) berekend voor elke leerprocesmaat per taak en per proefpersoon. De frequentie van de leerprocessen werd gedeeld door het aantal experimenten dat een proefpersoon voor de betreffende taak had uitgevoerd. Op deze wijze kon een vergelijking plaatsvinden tussen de groep 2VWO en universitaire studenten. Hiervoor werd gebruik gemaakt van Anova en de Mann-Whitney U toets om verschillen tussen gemiddelden van de twee groepen te bepalen. Om de relatie tussen leerprocessen uit de analyse van hardopdenkprotocollen en de leerprestaties te onderzoeken, werd in eerste instantie de Pearson correlatie coëfficiënt gebruikt. Leerprocessen (zonder overlap) die significant correleerden met de leerprestaties, werden geselecteerd als mogelijke indicator voor metacognitieve vaardigheid. Vervolgens werd een multiple regressieanalyse uitgevoerd met de geselecteerde leerprocessen om te bepalen of de leerprocessen geschikt waren om metacognitieve vaardigheden aan te duiden. Er werd gecontroleerd voor verschillende combinaties van 6 tot 4 variabelen. Uiteindelijk werden de variabelen met de

beste voorspellende waarde voor leerprestaties voor beide groepen gecombineerd tot één variabele om metacognitieve vaardigheden te representeren. Deze samengestelde maat van metacognitieve vaardigheden werd daarna gebruikt voor de correlatieanalyse met leerprestaties en intelligentie waarbij tevens (semi)partiële correlaties werden berekend.

Resultaten

Ten eerste werd onderzocht of het gebruik van leerprocessen, die sterk samenhangen met de leerprestaties, in frequentie toenemen met de leeftijd, door een vergelijking te maken tussen de groep 2VWO en bachelor studenten. Hiervoor werden eerst de leerprestaties per groep met elkaar vergeleken en vervolgens werden de gemiddelde scores op de leerprocessen naast elkaar gezet, zodat deze daarna met de leerprestaties gecorreleerd konden worden.

Om te bepalen of er verschillen bestonden in leerprestaties tussen de groepen 2VWO en studenten zijn de gemiddelde begripsscores van beide taken met elkaar vergeleken. Om te controleren of de variantie per leertaak gelijk was, werd de Levene's test gebruikt. Uit Tabel 1 blijkt dat er geen significant verschil bestaat in de totale leerprestaties tussen de groepen 2VWO en studenten. Wanneer gekeken werd naar de afzonderlijke domeinen, is na Bonferroni correctie ($\alpha = .025$) geen significant verschil gevonden in leerprestaties tussen de groepen 2VWO en studenten voor de otter taak in het aardrijkskunde domein en de plant taak in het biologie domein.

Tabel 1

Gemiddelde en Standaarddeviatie van de Leerprestaties per Taak en Groep

	Groep 2VWO: Begripsscore		Groep studenten: Begripsscore		df	F	p
	n	Gem (SD)	n	Gem (SD)			
Aardrijkskunde: Otter taak	29	10.48 (2.68)	28	11.89 (2.02)	1.52	4.53	.029
Biologie: Plant taak	28	11.18 (3.44)	28	11.57 (2.02)	1.54	0.00	.677
Totaal	30	10.72 (2.65)	30	11.75 (2.34)	1.58	1.01	.114

Noot. De maximale begripsscore is 18 punten.

Om verschillen in het toepassen van leerprocessen te kunnen bepalen tussen de groepen 2VWO en studenten zijn in Tabel 2 gemiddelden per groep per leerproces weergegeven. De score op CVS is omgeschaald omdat een hoge score op CVS betekende dat er veel variabelen per experiment zijn veranderd, wat echter als negatief wordt beschouwd. Om te controleren

Tabel 2

Gemiddelden in Leerprocessen per Groep met One Way Anova of Mann-Whitney U

	2VWO (n=27)	Studenten (n=26)	Anova			Mann - Whitney U		
	<i>gem (sd)</i>	<i>gem (sd)</i>	df	<i>F</i>	<i>p</i>	<i>z</i>	<i>U</i>	<i>p</i>
Hypothese genereren								
Formuleren van hypothesen	.43 (.20)	.29 (.23)	1.51	5.64	.025			
Eenvoudige hypothesen stellen	.38 (.18)	.24 (.19)	1.51	7.61	.009			
Complexe hypothesen opstellen	.05 (.08)	.05 (.08)	1.51	0.00	.097			
Experimenten ontwerpen								
Experimentruimte	26.49 (7.16)	29.61 (7.26)	1.56	2.71	.105			
CVS (omgeschaald)	3.22 (.48)	3.49 (.18)				2.25	564.50	.025
Interpretatie van data								
Formuleren van inferenties	.95 (.16)	.89 (.11)	1.51	2.54	.118			
Valide inferenties formuleren	.68 (.24)	.79 (.12)				1.68	445.50	.093
Percentage valide inferenties	.72 (.23)	.89 (.09)				3.14	527.50	.002
Voorspellingsfout	1.03 (.23)	1.13 (.31)	1.56	2.04	.158			
Regulatie van leren								
Formuleren van plannen	.80 (.12)	.89 (.05)				3.77	563.00	<.001
Eén variabele checken	.51 (.19)	.53 (.13)	1.51	0.11	.738			
Interactie checken	.07 (.08)	.15 (.12)				3.31	536.50	.001
Specifieke uitkomst genereren	.10 (.13)	.05 (.05)				-1.03	293.50	.303
Effect checken van een variabele in een ander experiment	.09 (.11)	.14 (.11)	1.51	2.33	.133			
Uitkomst voorspellen	.02 (.04)	.02 (.04)	1.51	0.00	.971			
Inductieve plannen	.58 (.20)	.68 (.11)				1.32	425.00	.188
Deductieve plannen	.11 (.12)	.16 (.12)	1.51	2.10	.153			
Ongericht of geen plan	.20 (.12)	.11 (.05)				-3.77	139.00	<.001

Noot. De scores voor de leerprocessen zijn gewogen door het aantal uitgevoerde experimenten per taak.

of er een ANOVA mocht worden uitgevoerd bij de diverse leerprocessen is per leerproces de Levene's test gebruikt om te toetsen of de variantie tussen de groepen gelijk is. Bij gelijke variantie is ANOVA gebruikt en bij ongelijke variantie is de non-parametrische versie van ANOVA voor twee groepen, de Mann-Whitney U toets uitgevoerd. Uit deze toetsen blijkt dat er significante verschillen bestaan in het gebruik van metacognitieve leerprocessen tussen 2VWO-ers en studenten.

De studenten scoren hoger op CVS dan 2VWO-ers wat betekent dat zij minder variabelen per experiment variëren. Dit verschil is significant. De studenten scoren significant hoger dan 2VWO-ers op het percentage valide inferenties. Ook scoren de studenten significant hoger dan 2VWO-ers op het formuleren van plannen en op het checken van interacties. De 2VWO-ers scoren significant hoger op het formuleren van hypothesen dan de studenten. Ook scoren zij significant hoger op het formuleren van eenvoudige hypothesen dan studenten. Tevens scoren de 2VWO-ers significant hoger op het gebruiken van ongerichte of geen plannen dan studenten. Tussen de twee leeftijdsgroepen lijken dus significante verschillen in strategieën waar te nemen. Echter, na het toepassen van een Bonferroni correctie voor meervoudige vergelijkingen ($\alpha 0.05/18 = 0.003$) blijken tussen beide groepen nog vier leerprocessen significant van elkaar te verschillen (zie Tabel 2). De studenten scoren significant hoger op het percentage valide inferenties, het formuleren van plannen en het checken van interacties. De 2VWO-ers scoren significant hoger op het maken van ongerichte of geen plannen (wat negatief correleert met leerprestaties, zie uitleg bij correlaties).

Om het verband tussen leerprestaties voor beide leeftijdsgroepen en de leerprocessen uit de analyse van hardopdenkprotocollen te bepalen, is de Pearson correlatie coëfficiënt gebruikt. Correlaties tussen de leerprestaties en de leerprocessen worden in Tabel 3 weergegeven. Vervolgens is met behulp van regressieanalyse een maat voor metacognitieve vaardigheden berekend, waarna deze maat gerelateerd is aan intelligentie en leerprestaties.

In de totale groep zijn de correlaties tussen de gemiddelde leerprestaties over beide taken en de metacognitieve leerprocessen significant voor de volgende acht indicatoren: het toepassen van CVS, valide inferenties formuleren, het percentage valide inferenties, het formuleren van plannen, een specifieke uitkomst genereren, het effect checken van een variabele in een ander experiment, deductieve plannen en ongericht of geen plan. Twee van deze correlaties met de leerprestaties zijn negatief, namelijk een specifieke uitkomst genereren en een ongericht plan. Dat betekent dat zowel een hogere score op specifieke uitkomsten genereren als een ongericht plan samengaan met een lagere score op leerprestaties.

Tabel 3

Pearson Correlaties Coëfficiënten tussen Leerprocessen en Leerprestatie per Leeftijdsgroep

	Totaal (n=53)		2VWO (n=27)		Studenten (n=26)	
	r	p	r	p	r	p
Hypothese genereren						
Formuleren van hypothesen	.10	.496	.12	.557	.19	.356
Eenvoudige hypothesen stellen	.18	.196	.39	.044	.10	.619
Complexe hypothesen opstellen	-.18	.201	-.63	<.001	.30	.143
Experimenten ontwerpen						
Experimentruimte	.19	.145	.18	.341	.11	.567
CVS (omgeschaald)	.56	<.001	.62	<.001	.30	.123
Interpretatie van data						
Formuleren van inferenties	.17	.221	.21	.293	.22	.284
Valide inferenties formuleren	.68	<.001	.80	<.001	.46	.017
Percentage valide inferenties	.66	<.001	.81	<.001	.45	.020
Voorspellingsfout	-.20	.128	-.41	.026	-.14	.485
Regulatie van leren						
Formuleren van plannen	.30	.031	.35	.071	.07	.747
Eén variabele checken	.08	.567	.33	.097	-.36	.070
Interactie checken	.11	.419	-.37	.059	.42	.033
Specifieke uitkomst genereren	-.43	.001	-.36	.070	-.71	<.001
Effect checken van een variabele in een ander experiment	.36	.009	.46	.016	.20	.323
Uitkomst voorspellen	.20	.161	.12	.547	.28	.164
Inductieve plannen	.15	.287	.16	.433	.03	.898
Deductieve plannen	.40	.003	.46	.017	.30	.144
Ongericht plan	-.30	.031	-.35	.071	-.07	.747

In de studenten groep zijn in aantal minder significante correlaties gevonden tussen de leerprestaties en de leerprocessen dan in de 2VWO groep. Positieve significante correlaties tussen leerprestaties en leerprocessen in de studenten groep zijn: het formuleren van valide inferenties, percentage valide inferenties en een interactie checken. Een negatieve significante correlatie is gevonden tussen leerprestaties en het genereren van een specifieke uitkomst.

In de 2VWO groep zijn de volgende positieve correlaties tussen de leerprestaties en de leerprocessen significant: eenvoudige hypothesen stellen, het toepassen van CVS, valide inferenties formuleren, het percentage valide inferenties, het effect checken van een variabele in een ander experiment en deductieve plannen. Significant negatieve correlaties in de groep 2VWO tussen de leerprocessen en de leerprestaties zijn: complexe hypothesen opstellen en de voorspellingsfout.

De correlatiepatronen tussen leerprestaties en leerprocessen zijn verschillend voor de groepen 2VWO en studenten. Tevens is het correlatiepatroon van de totaalgroep anders dan die van beide groepen. Positieve correlaties tussen de leerprestaties en leerprocessen die zowel in de groepen 2VWO en studenten als bij de totaalgroep voorkomen zijn het formuleren van valide inferenties en het percentage valide inferenties.

Zoals eerder uit Tabel 3 is gebleken, zijn in de totale groep bij acht leerprocessen significante correlaties gevonden met de leerprestaties. Voor het bepalen van de relatie tussen de gebruikte leerprocessen en de leerprestaties is een multiple regressieanalyse uitgevoerd, weergegeven in Tabel 4. De variabelen met de beste voorspellende waarde voor leerprestaties voor beide groepen zijn gecombineerd in één model om metacognitieve vaardigheden te representeren.

Tabel 4

Regressieanalyse van Leerprocessen met Voorspellende Waarde voor Leerprestaties

	Model 1		Model 2	
	<i>B</i>	95% BI	<i>B</i>	95% BI
Constant	38.487**	[10.33, 66.65]	38.465**	[10.65, 66.28]
CVS (omgeschaald)	1.69	[-.21, 3,60]	1.67*	[.06, 3,29]
Valide inferenties	8.84***	[5.39, 12,30]	8.84***	[5.43, 12,24]
Deductieve plannen	4.26	[-.15, 8.67]	4.25	[-.08, 8.59]
Ongericht plan	-8.20*	[-14.55, -1.85]	-8.20*	[-14.48, -1.93]
Specifieke uitkomst genereren	-.12	[-6.63, 6,39]		
R^2	.61		.61	
F	14.65***		18.71***	
ΔR^2			.00	
ΔF			0.00	

Noot. BI = betrouwbaarheidsinterval.

* $p < .05$. ** $p < .01$ *** $p < .001$.

Voorafgaand aan deze analyse zijn de leerprocessen met negatieve correlaties met leerprestaties omgezet van negatief naar positief. Aangezien een ongericht plan oftewel geen plan tegengesteld is aan het formuleren van een plan, is deze laatste niet in de analyse opgenomen. De reden hiervoor is dat een specifieke uitkomst genereren en deductieve plannen ook zijn opgenomen in de analyse en dat zij al deel uitmaken van het overkoepelende begrip ‘het formuleren van een plan’. Aangezien het formuleren van valide inferenties ten opzichte van alle experimenten en het percentage van de valide inferenties van het totaal aantal gemaakte inferenties hoog met elkaar correleren ($r = .86$, $p < .001$) omdat zij elkaar

deels overlappen, is ervoor gekozen om het formuleren van valide inferenties, dat het hoogst correleerde met de leerprestaties, samen met de overige vier leerprocessen op te nemen in de regressieanalyse.

Het genereren van een specifieke uitkomst voegt als vijfde variabele in de regressieanalyse van de totale groep weinig voorspellende waarde toe. Wanneer we de voorspellende waarde per groep bekijken, dan heeft het checken van een specifieke uitkomst voor de groep studenten wel degelijk een toegevoegde waarde in het voorspellen van leerprestaties. De verschillen tussen model 1 en 2 worden weergegeven in Tabel 5. Om deze reden is ervoor gekozen deze vijfde variabele toch op te nemen in het model.

Bij het uitvoeren van de regressieanalyse met de geselecteerde vijf variabelen voor beide groepen apart, blijkt dat de voorspellende waarde van het gezamenlijke model voor de groepen verschillend is, 2VWO ($R^2 = .79$), studenten ($R^2 = .58$). Voor de groep 2 VWO is het percentage verklaarde variantie hoger dan voor de groep studenten, wat betekent dat het model meer representatief is voor de eerste groep. De vijf leerprocessen die zijn opgenomen in de maat van metacognitieve vaardigheden zijn: het toepassen van CVS, het formuleren van valide inferenties, het genereren van specifieke uitkomsten (negatieve correlatie, omgezet), deductieve plannen en een ongericht oftewel geen plan (negatieve correlatie, omgezet). De homogeniteit van deze maat van metacognitieve vaardigheden, gemeten met Cronbach's Alpha, is .72.

Tabel 5

Verkorte Weergave van de Regressieanalyse van Leerprocessen met Voorspellende Waarde voor Leerprestaties per Groep

	Model 1		Model 2	
	R^2	F	R^2	F
2VWO	.79	16.01***	.78	18.97***
Studenten	.58	5.24**	.38	3.02*

Noot. Model 1 (5 variabelen: CVS omgeschaald, valide inferenties, deductieve plannen, ongericht plan, specifieke uitkomst genereren). Model 2 (4 variabelen: CVS omgeschaald, valide inferenties, deductieve plannen, ongericht plan).

* $p < .05$. ** $p < .01$ *** $p < .001$.

Tussen de leeftijdsgroepen 2VWO en studenten zijn significante verschillen gevonden in de scores op de samengestelde maat van metacognitieve vaardigheden ($U = 477.00$, $p = .011$). De 2VWO-ers hebben een gemiddelde score op deze maat van 9.73 (SD: 0.89) en de studenten

hebben een significant hogere gemiddelde score van 10.27 (SD: 0.32). Dit betekent dat de studentengroep meer gebruik maakt van metacognitieve leerprocessen die ten gunste komen van de leerprestaties.

Hiernaast zijn tussen de leeftijdsgroepen 2VWO en studenten ook significante verschillen gevonden in de ongewogen gemiddelde scores op intelligentie ($F(1, 50) = 0.49, p < .001$). De 2VWO-ers hebben een ongewogen gemiddelde intelligentiescore van 107.93 (SD: 27.27) en de studenten hebben een significant hogere ongewogen gemiddelde score van 140.12 (SD: 28.65), wat wijst op intellectuele groei.

Ten tweede onderzochten we welke relaties er bestaan tussen metacognitieve vaardigheden, en intelligentie. We probeerden ondersteuning te vinden voor het mixed model van Veenman, et al. (2004) waarbij verondersteld wordt dat metacognitie deels gerelateerd is aan intelligentie met elk een eigen meerwaarde voor de leerprestatie.

Om te onderzoeken of metacognitieve vaardigheden en intelligentie deels aan elkaar zijn gerelateerd in het huidige onderzoek zijn Pearson correlatie coëfficiënten berekend voor de samenhang in de totale groep ($N = 52$) tussen metacognitie, intelligentie en leerprestaties (zie Tabel 6). In het onderzoek is in de totale groep een positieve significante correlatie gevonden tussen metacognitie en intelligentie ($r = .51, p < .001$). Ook bestaat er een positieve correlatie tussen metacognitie en leerprestaties ($r = .65, p < .001$) en tussen intelligentie en leerprestaties ($r = .45, p < .001$).

Om de bijdrage, het unieke deel, van respectievelijk metacognitie en intelligentie ten opzichte van leerprestaties te verklaren, zijn semi-partiële correlatieanalyses uitgevoerd. Voor de totale groep is de semi-partiële correlatie tussen metacognitie en leerprestaties (intelligentie uitgepartieerd, ook wel uitgefilterd) $.49$ ($p < .01$). Deze semi-partiële correlatie reflecteert de samenhang van metacognitie met leerprestaties, onafhankelijk van intelligentie (zie Veenman, 2008). De semi-partiële correlatie tussen intelligentie en leerprestaties (metacognitie uitgepartieerd) is $.13$ (n.s.). Deze semi-partiële correlatie reflecteert de samenhang van intelligentie met leerprestaties, onafhankelijk van metacognitie.

Door het gebruik van analyse technieken voor het delen van de variantie (Pedhazur, 1982), werd de unieke en de gezamenlijke bron van variantie in leerprestaties onderverdeeld voor intelligentie en metacognitie. De totale variantie is berekend door het kwadraat van de meervoudige correlatie van intelligentie en metacognitie, voor het voorspellen van leerprestaties, te bepalen uit de correlatie tussen intelligentie en leerprestaties en de semi-partiële correlatie van metacognitie en leerprestaties (zie Tabel 6: $R^2 = (.45)^2 + (.49)^2 = .44$).

De gedeelde variantie van beide voorspellers samen voor leerprestaties is berekend door het kwadraat van beide semi-partiële correlaties van het kwadraat van de meervoudige correlatie af te trekken (gedeelde variantie $r^2 = .44 - (.13)^2 - (.49)^2 = .18$). De hierbij berekende uniek verklaarde variantie van metacognitie voor leerprestaties is $(.49)^2 = 24\%$ en de uniek verklaarde variantie van intelligentie voor leerprestaties is $(.13)^2 = 2\%$. Beide variabelen samen hebben 18% gedeelde variantie zodat de totaal verklaarde variantie 44% bedraagt. Voor de berekeningswijze wordt verwezen naar Veenman (2008) en Veenman, Bavelaar, De Wolf, en Van Haaren (2014).

Tabel 6

Correlaties tussen Metacognitieve Vaardigheden, Intelligentie en Leerprestaties en Semi-partiële Correlaties

	Intelligentie	Metacognitie	Semi-partieel Metacognitie	Semi-partieel Intelligentie
Totale Groep				
Leerprestaties	.45***	.65***	.49**	.13
Metacognitie	.51***			
Groep 2VWO				
Leerprestaties	.45*	.71***	.56**	.15
Metacognitie	.45*			
Groep studenten				
Leerprestaties	.36	.60**	.49**	.12
Metacognitie	.42*			

* $p < .05$. ** $p < .01$ *** $p < .001$.

In het onderzoek is bij de groep 2VWO ($N = 27$) een positieve significante correlatie gevonden tussen metacognitie en intelligentie ($r = .45$, $p = .019$). Ook bestaat er een positieve correlatie tussen metacognitie en leerprestaties ($r = .71$, $p < .001$) en tussen intelligentie en leerprestaties ($r = .45$, $p = .012$), (zie Tabel 6). Voor de groep 2VWO is de semi-partiële correlatie tussen metacognitie en leerprestaties (intelligentie uitgepartieerd) $.56$ ($p < .01$). De semi-partiële correlatie tussen intelligentie en leerprestaties (metacognitie uitgepartieerd) is $.15$ (n.s.). De hierbij berekende uniek verklaarde variantie van metacognitie voor leerprestaties is 32% en de uniek verklaarde variantie van intelligentie voor leerprestaties is 2%. De gedeelde variantie van beide variabelen samen voor leerprestaties bedraagt 18% en totaal verklaarde variantie is 52%.

Bij de groep studenten ($N = 25$) is in het onderzoek een positieve significante correlatie gevonden tussen metacognitie en leerprestaties ($r = .60$, $p = .003$). Ook bestaat er een

positieve correlatie tussen metacognitie en intelligentie ($r = .42$, $p = .036$). Intelligentie lijkt geen betekenisvolle rol meer te spelen ten opzichte van leerprestaties (zie Tabel 6). Voor de groep studenten is de semi-partiële correlatie tussen metacognitie en leerprestaties (intelligentie uitgepartieerd) $.49$ ($p < .001$). De semi-partiële correlatie tussen intelligentie en leerprestaties (metacognitie uitgepartieerd) is $.12$ (n.s.). De hierbij berekende uniek verklaarde variantie van metacognitie voor leerprestaties is 24% en de uniek verklaarde variantie van intelligentie voor leerprestaties is 1%. De gedeelde variantie van beide variabelen samen voor leerprestaties bedraagt 11% en totaal verklaarde variantie is 37%.

Conclusies en Discussie

In dit onderzoek is onderzocht welke verschillen bestaan in het ontwikkelingsniveau van metacognitieve vaardigheden bij onderzoekend leren en hoe deze metacognitieve vaardigheden gerelateerd zijn aan intelligentie en leerprestaties. Hierbij zijn gegevens van 2VWO leerlingen en universitaire bachelor studenten uit hardopdenkprotocollen geanalyseerd en geïntegreerd met eerder verzamelde gegevens, over leerprestaties, metacognitie en intelligentie uit het proefschrift van Wilhelm (2001).

De eerste hypothese waarin verondersteld wordt dat het gebruik van metacognitieve leerprocessen, die positief samenhangen met de leerprestaties, in frequentie zal toenemen met de leeftijd, kan op basis van de gevonden verschillen tussen leeftijdsgroepen grotendeels worden ondersteund.

Uit de resultatensectie volgt dat studenten significant hogere scores op metacognitieve leerprocessen laten zien. Daarnaast zijn tussen de twee leeftijdsgroepen een aantal verschillen in het toepassen van strategieën waar te nemen. Deze bevindingen sluiten aan bij het onderzoek van Kuhn (2000) waarin zij stelt dat de selectie en het toepassen van strategieën zich ontwikkelen naar een meer frequent gebruik van adequate en gevorderde strategieën. In het huidige onderzoek laten de studenten en de 2VWO-ers significante verschillen zien in het gebruik van diverse metacognitieve leerprocessen. De studenten behalen een hoger percentage van valide inferenties, wat betekent dat zij betere, meer valide, conclusies trekken uit hun experimenten in vergelijking met 2VWO-ers. Zij werken doelgerichter door het formuleren van plannen en ook veronderstellen zij vaker interacties dan 2VWO-ers. Studenten lijken zich te richten op het aansturen van de uitvoering van hun experimenten. De 2VWO-ers

daarentegen werken minder doelgericht en maken vaker ongerichte of geen plannen van tevoren.

Om te kunnen bepalen of er een positieve samenhang bestaat tussen de gebruikte metacognitieve leerprocessen en de leerprestaties in de verschillende groepen is het van belang de correlatiepatronen te beschouwen. Deze correlatiepatronen tussen leerprocessen en leerprestaties uit de resultatensectie laten een aantal verschillen tussen de groepen zien. Er kan worden afgeleid dat bij de groep 2VWO het stellen van eenvoudige hypothesen samengaat met goede leerprestaties. Wanneer 2VWO-ers echter complexe hypothesen stellen, pakt dit negatief uit; dit hangt namelijk samen met lagere leerprestaties. Bij de studentengroep heeft het opstellen van hypothesen geen betekenisvolle invloed op de correlatie met de leerprestaties. Mogelijk kunnen studenten goed presteren zonder het gebruik van hypothesen. Het opstellen van complexe hypothesen correleert in de studentengroep echter wel positief met het checken van interacties ($r = .74, p < .001$). Mogelijk speelt hypothesevorming in deze groep wel een rol in het proces van het onderzoeken van interacties.

In de groep 2VWO blijkt bij het ontwerpen van experimenten het netjes experimenteren van belang voor de leerprestaties. Wanneer 2VWO-ers minder zorgvuldig zijn met het experimenteren en meer variabelen per experiment variëren, correleert dat met lagere leerprestaties. Bij de studentengroep lijkt minder zorgvuldig experimenteren weinig van invloed op de leerprestaties.

Bij de interpretatie van de data valt het op dat in de groep 2VWO, het percentage van valide inferenties formuleren, het hoogst correleert met de leerprestaties, waardoor het trekken van gegronde conclusies bij hen een sterke samenhang heeft met de leerprestaties.

Bij de regulatie van het leren blijkt voor de studenten een positief verband te bestaan met de leerprestaties wanneer zij interacties checken en een negatief verband wanneer zij een specifieke uitkomst genereren. De 2VWO-ers laten een ander patroon zien, bij hen gaat het checken van interacties juist samen met lagere leerprestaties. Studenten zijn blijkbaar in staat om het checken van interacties op een zinvolle wijze toe te passen. In de groep 2VWO gaat het effect checken van een variabele in een ander experiment wel samen met hogere leerprestaties; dit lijkt een soort tussenstap naar het volledig ontdekken van interacties. Het gebruik maken van deductieve plannen, waarbij ideeën gecheckt worden in de experimenten, bestaande uit een combinatie van het effect checken van een variabele in een ander experiment samen met het voorspellen van de uitkomst van een experiment, correleert in de 2VWO groep ook positief met leerprestaties.

Uit deze bevindingen volgt dat studenten meer gebruik maken van metacognitieve leerprocessen die hun leerprestaties ondersteunen of vergroten. Samengevat betekenen bovenstaande verbanden dat de studentengroep vaker gebruik maakt van metacognitieve leerprocessen die ten gunste komen van de leerprestaties, waarmee bevestiging is gevonden voor de eerste hypothese. De nuances in het proces zijn hierboven nader toegelicht.

De tweede hypothese, waarbij verwacht werd ondersteuning te vinden voor het mixed model waarbij metacognitie deels gerelateerd is aan intelligentie met elk een eigen meerwaarde voor de leerprestatie voor de twee leeftijdsgroepen, kan worden bevestigd.

In de totale groep, de 2VWO groep en de studentengroep wordt ondersteuning gevonden voor het mixed model van Veenman, et al. (2004). Zowel in de totale groep als in de 2VWO groep en studentengroep is een middelmatige positieve significante correlatie gevonden tussen metacognitie en intelligentie, waarbij metacognitie en intelligentie beide een eigen meerwaarde hebben voor de leerprestaties. Dit komt overeen met de bevindingen van Van der Stel en Veenman (2008) waarin intelligentie en metacognitie bij 12-jarigen middelmatig aan elkaar zijn gecorreleerd. In het huidige onderzoek wordt in de studentengroep echter geen betekenisvolle relatie gevonden tussen intelligentie en leerprestaties. Dit verschil kan niet verklaard worden uit verschillen in spreiding in de intelligentiescores van beide groepen. Toch wordt intelligentie als belangrijke voorspeller van het leren beschouwd, daar het positief gerelateerd is aan academisch succes (Busato, et al., 2000). Bij andere onderzoekend leren taken zijn eveneens positieve correlaties gevonden tussen intelligentie en leerprestaties (Veenman, et al., 2004 en Dabbous, 2014).

Uit de semi-partiële correlatieanalyse blijkt dat de relatie tussen intelligentie en leerprestaties niet significant is in beide groepen, wanneer metacognitie uitgepartieerd wordt. De unieke bijdrage van intelligentie voor de leerprestaties is in het huidige onderzoek laag en nagenoeg gelijk in beide groepen, 2% voor 2VWO-ers en 1% voor studenten. In een overzicht van verschillende onderzoeken door Veenman et al. (2014) van deelnemers van verschillende leeftijden, verschillende soorten taken en verschillende domeinen werd de unieke bijdrage van intelligentie hoger ingeschat, namelijk 10% (Veenman, 2008). De unieke bijdrage van intelligentie aan het leren blijkt ook afhankelijk te zijn van het soort leertaak. Bij wiskundetaken werd bijvoorbeeld een hogere correlatie tussen intelligentie en leerprestaties gevonden dan bij geschiedenistaken (Van der Stel & Veenman, 2008 en 2010). Aardijkskunde- en biologie taken zoals gebruikt in het huidige onderzoek hebben meer raakvlak met geschiedenistaken.

Daarentegen is de unieke bijdrage van metacognitie voor het leren in het huidige onderzoek iets hoger in de totale groep 24% ten opzichte van 18% in het onderzoeksoverzicht van Veenman (2008). Kortom in het huidige onderzoek lijken 2VWO-ers en studenten bij de onderzoekend leren taken vooral een beroep te doen op hun metacognitieve vaardigheden. De unieke bijdrage van metacognitie voor de leerprestaties is bij 2VWO-ers en studenten respectievelijk 32% en 24%. In de studentengroep is de unieke bijdrage van metacognitie aan de leerprestaties lager. Een verklaring kan gevonden worden in de samenstelling van de maat van metacognitie, die voor beide groepen samen is opgesteld. Uit de multiple regressieanalyse voor beide groepen samen volgt dat de leerprocessen die in deze maat zijn opgenomen, iets minder goed overeenkomen met de leerprocessen die in de studentengroep zelf hoog correleren met de leerprestaties. Bij het uitvoeren van een multiple regressieanalyse voor de groep studenten apart blijkt dat in deze groep een andere combinatie van leerprocessen een hogere voorspellende waarde heeft voor de leerprestaties dan in de totaalgroep. De studenten zetten een aantal andere metacognitieve processen in dan die opgenomen zijn in de maat voor de totaalgroep. Blijkbaar vindt er in de tussenliggende periode van 2VWO naar student een ontwikkeling plaats in het toepassen van metacognitieve processen waardoor de algemene maat van metacognitie minder aansluit bij de studentengroep.

De derde hypothese waarin werd verwacht bevestiging te vinden van de geleidelijke groei van metacognitie samen met intellectuele groei (parallele ontwikkeling) in de twee leeftijdsgroepen wordt deels ondersteund.

Het gebruik van metacognitieve vaardigheden neemt toe met de leeftijd en de gebruikte strategieën veranderen in de leeftijdsgroepen. In de studentengroep worden andere metacognitieve strategieën aangesproken die meer ten gunste komen van de leerprestaties. Er is zowel sprake van kwantitatieve als kwalitatieve groei in metacognitieve leerprocessen. Hiernaast neemt de ongewogen gemiddelde intelligentiescore toe met de leeftijd, wat duidt op intellectuele groei. Ook zijn in beide groepen significante middelmatige correlaties gevonden tussen metacognitie en intelligentie wat wijst op een ontwikkeling die deels aan elkaar gerelateerd is. Deze bevindingen komen overeen met die van Alexander, et al. (1995) waarin verwezen wordt naar ontwikkeling van metacognitieve kennis in een geleidelijk stijgende lijn met leeftijd, waarbij de impact van intelligentie niet toeneemt of afneemt tijdens de jaren. Echter deze geleidelijke, gezamenlijke groei kan met de huidige onderzoeksgegevens niet geheel worden getoetst aangezien intelligentie en metacognitie niet longitudinaal, maar cross-

sectioneel zijn onderzocht en de ontwikkelingslijnen van proefpersonen niet geheel gevolgd konden worden.

Er zijn een aantal beperkingen die in dit onderzoek kunnen worden genoemd. De proefpersonen zijn niet afkomstig uit een aselechte steekproef en namen vrijwillig deel aan het onderzoek. Het generaliseren van gegevens naar alle 14-jarigen en jongvolwassenen is daarom niet gewaarborgd. De proefpersonen vertegenwoordigen waarschijnlijk meer de hogere intelligentieniveaus omdat zij afkomstig zijn uit het Voorbereidend Wetenschappelijk Onderwijs en uit het Wetenschappelijk Onderwijs. Psychologiestudenten representeren mogelijk niet de minder talige leerlingen waardoor de resultaten mogelijk minder van toepassing zijn op bijvoorbeeld technische studenten. Hier staat echter tegenover dat andere onderzoeken met vergelijkbare groepen zijn uitgevoerd (studenten en VWO-ers) en dat onderzoekend leren taken in de hogere onderwijsniveaus regelmatig worden toegepast.

In het onderzoek is geen longitudinaal onderzoek uitgevoerd maar er is een vergelijking gemaakt tussen leeftijdsgroepen. De tussenliggende periode is niet gedekt, waardoor het ontwikkelingsverloop niet geheel kon worden gevolgd en uitspraken over ontwikkeling dus met voorzichtigheid dienen te worden geïnterpreteerd.

De gegevens uit het proefschrift van Wilhelm zijn verouderd, maar de data uit het huidige onderzoek sluiten wel grotendeels aan bij bevindingen uit andere meer recentere onderzoeken (Veenman & Verheij, 2003; Veenman & Beishuizen, 2004; Veenman et al., 2004; Veenman & Spaans, 2005; Veenman, 2008; Van der Stel en Veenman, 2008; Van de Stel en Veenman, 2010; Veenman et al., 2014).

Proefleider effecten kunnen bij de afname van het onderzoek een rol hebben gespeeld. Wanneer een proefleider meer of minder informatie geeft over een taak of bijvoorbeeld extra uitleg geeft, kan dit het leerproces en de leerresultaten beïnvloeden. In het onderzoek werd echter gebruik gemaakt van een uitgebreid afname- en scoringprotocol zodat proefleider effecten in principe verwaarloosbaar zouden moeten zijn. Voorafgaand aan de taken is telkens dezelfde informatie gegeven en tijdens en na afloop van de taken zijn dezelfde vragen gesteld voor het bepalen van de leerprocessen en leerprestaties. Volgorde effecten zijn zoveel mogelijk uitgesloten door de proefpersonen afwisselend in conditie A en B te plaatsen, waarbij in conditie A eerst de biologie taken aan bod kwamen en in conditie B eerst de aardrijkskundetaken.

De interpretatie van het scoringprotocol zou een punt van aandacht kunnen zijn. Wanneer een beoordelaar de scoringscategorieën niet goed kent of verkeerd toepast, kan dit een

vertekend beeld geven van de geobserveerde leerprocessen. Aangezien de opgestelde categorieën uitgebreid zijn omschreven en doorgesproken, zou dit echter geen belemmering voor het onderzoek moeten vormen. In overeenkomstig onderzoek was de interbeoordelaars-betrouwbaarheid voor de verschillende categorieën van onderzoekend leren vrij hoog, namelijk 82% (Evers, 2009).

Een andere bevinding in dit onderzoek is dat de algemene maat voor metacognitie minder goed aansluit bij beide groepen, aangezien er duidelijke verschillen in strategieën zijn gevonden tussen beide groepen. Mogelijk zou voor beide leeftijdsgroepen een aparte maat van metacognitie kunnen worden opgesteld. Een nadeel hiervan is echter dat deze maten onderling niet vergelijkbaar zijn waardoor de voorspellende waarde niet groot zal zijn.

De wijze waarop metacognitie is gemeten in dit onderzoek zou verbeterd kunnen worden met een andere werkwijze. Door metacognitieve vaardigheden minder af te leiden van het denkproces (hardopdenken) en meer direct te observeren door oriënterende, plannende, monitorende en evaluatieve activiteiten in kaart te brengen. Het nadeel hiervan kan zijn dat metacognitief gedrag niet altijd expliciet zichtbaar hoeft te zijn en in plaats daarvan afgeleid moet worden uit bepaalde cognitieve activiteiten (Veenman et al., 2006). Combinaties van het observeren van gedrag (en mogelijk direct registreren) met hardopdenkprotocollen zijn ook denkbaar.

Een aanbeveling is om longitudinaal onderzoek uit te voeren om meer inzicht te krijgen in de ontwikkeling in het toepassen van metacognitieve leerprocessen gedurende de periode van 14 tot 23 jaar. Op deze wijze kunnen proefpersonen gedurende een langere periode worden gevolgd en kan de ontwikkeling van metacognitieve vaardigheden nauwkeurig in kaart worden gebracht. Een sequentieel design behoort ook tot de mogelijkheden, waarbij longitudinaal onderzoek wordt gecombineerd met een cross-sectioneel ontwerp met verschillende leeftijdsgroepen. Een andere mogelijkheid is om nog één of twee tussenliggende groep(en) op te nemen in het cross-sectionele onderzoek met leeftijdsgroepen. In het huidige onderzoek is de stap van 14-jarigen naar 23-jarigen namelijk vrij groot en beslaat nagenoeg 10 jaar. Tussen beide groepen zijn verschillen waargenomen in de toepassing van metacognitieve vaardigheden. De toevoeging van één of meer leeftijdsgroepen kan zicht bieden op de ontwikkeling van metacognitieve vaardigheden in de tussenliggende periode en kan verhelderen welke stappen waar te nemen zijn. Op deze wijze kan de veronderstelde ontwikkeling in het huidige onderzoek meer onderbouwd worden. Zoals eerder beschreven, kon het ontwikkelingsverloop niet geheel worden gevolgd en zullen uitspraken over de ontwikkeling dus met voorzichtigheid dienen te worden geïnterpreteerd.

Een andere aanbeveling zou kunnen zijn om het proces naar het ontdekken van interacties te verhelderen door subcategorieën op te nemen in het scoreprotocol. In het huidige onderzoek is gevonden dat het checken van interacties een rol speelt bij het behalen van positieve leerprestaties in de studentengroep. Bij de 2VWO-ers heeft dit echter nog een negatief effect op de leerprestaties. Dit zou met kennis te maken kunnen hebben en/of met een bepaalde cognitieve ontwikkeling. De vraag is nu waar het omslagpunt ligt voor het al of niet ontdekken van interacties en hoe dit tot stand kan komen. Ter Braak (2013) heeft in haar onderzoek naar interacties al een aanzet gegeven door bij de scoring van interacties een onderverdeling te maken in het ontdekken van een interactie, het vermoeden hebben van een interactie en het vermoeden hebben of veronderstellen van een verkeerde interactie. Mogelijk zou een dergelijke indeling in categorieën ook in vervolgonderzoek opgenomen kunnen worden zodat het proces van het wel of niet ontdekken van interacties meer geleidelijk in kaart kan worden gebracht. Hierdoor kan het proces van de ontwikkeling van metacognitie in kleinere stappen worden weergegeven. Voor de dagelijkse praktijk zou dit zicht kunnen geven op hoe het leren ontdekken van interacties op een goede manier kan worden gestimuleerd zodat dit op jongere leeftijd al kan worden toegepast.

Tot slot zou er ook gekeken kunnen worden naar in hoeverre metacognitieve processen correleren met het concept executieve functies. Zoals Kuhn en Pease (2006) veronderstelden, komt leren steeds meer onder bewuste controle te staan in de jaren vanaf het midden van de kindertijd tot het begin van de volwassenheid. Deze stelling is consistent met het groeiend bewijs van de late ontwikkeling van de prefrontale cortex die betrokken is bij executief functioneren (Casey, Giedd, & Thomas, 2000). Het is ook consistent met het groeiende besef van het belang van de metacognitieve functies en de executive controle in cognitieve prestaties (Kuhn & Pease, 2006). Mogelijk heeft dit concept de nodige raakvlakken met de ontwikkeling van metacognitieve vaardigheden en zijn beide aan elkaar gerelateerd. Voor het meten van executieve functies zijn relatief nieuwe testen beschikbaar. De vraag die hierbij gesteld zou kunnen worden, is in hoeverre metacognitieve vaardigheden samenhangen met de ontwikkeling van executieve functies en of indirect de rijping van de prefrontale cortex een voorwaarde is voor het kunnen bereiken van een hoger metacognitief niveau bij onderzoekend leren. Voor de praktijk zou dit mogelijk kunnen betekenen dat het stimuleren van metacognitieve vaardigheden pas zinvol kan zijn als een bepaalde rijping van de hersenen heeft plaatsgevonden. Andersom is ook mogelijk: dat het stimuleren van metacognitieve vaardigheden juist de ontwikkeling van de prefrontale cortex bevordert en daarmee mede de ontwikkeling van executieve functies.

Referenties

- Alexander, J., Carr, M., & Schwanenflugel, M. (1995). Development of metacognition in gifted children: Directions for future research. *Developmental Review*, 15, 1–37. doi:10.1006/drev.1995.1001.
- Anderson, S. (2002). Working together to develop a professional learning community. In *Proceedings of the 25th HERDSA Annual Conference* (pp. 20–27), Perth, Australia. Published by the Higher Education Research and Development Society of Australia.
- Braak, A. M. ter (2013). *Het gebruik van analogieën bij het ontdekken van interacties tijdens onderzoekend leren* (Masterthese). University of Twente, Enschede.
- Brody, E. B. (1992). *Intelligence. Nature, determinants, and consequences*, (2nd ed.) New York: Academic Press.
- Brown, A. L., & DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. S. Siegel (Ed.), *Children's thinking: What develops?* (pp. 3–35). Hillsdale, NJ: Erlbaum.
- Busato, V.V., Prins, F.J., Elshout, J.J. & Hamaker, C. (2000). Learning style, personality, achievement motivation, intellectual ability and academic success of psychology students in higher education. *Personality & Individual Differences*, 29, 1057-1068.
- Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Casey, B. J., Giedd, J., & Thomas, K. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, 54, 241–257.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, 70, 1098–1120. doi:10.1111/1467-
- Dabbous, S. (2014). *The role of analogies in discovering interactions in inquiry learning tasks: a comparison between students from low, middle and high level of secondary school* (Masterthese). University of Twente, Enschede.
- Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, selfregulation, and self-regulated learning. *Educational Psychology Review*, 20, 391–409.
- Elshout, J. J. (1976). *Karakteristieke moeilijkheden in het denken [Characteristic difficulties in thinking]*. Amsterdam: University of Amsterdam.
- Elshout, J. J. (1983). Is measuring intelligence still useful? In S. B. Anderson, & J. S. Helmick (Eds.), *On educational testing* (pp. 45–56). San Francisco: Jossey-Bass.

- Evers, W. F. (2009) *The relation between metacognitive skillfulness, intellectual ability, prior knowledge, causal mechanisms and inquiry learning performance in fourth- and sixthgraders* (Masterthese). University of Twente, Enschede. Verkregen van: <http://purl.utwente.nl/essays/59469>.
- Evers, A., & Lucassen, W. (1983). *Differentiele Aanleg Testserie*. Lisse: Swets & Zeitlinger.
- Flanagan, J. C. (1951–1956). *Flanagan aptitude classification tests*. SRA.
- Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese, & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 5, pp. 181–211). New York: Academic Press.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906–911.
- Flavell, J. H. (1992). Perspectives on perspective taking. In H. Beilin, & P. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 107–141). Hillsdale, NJ: Erlbaum.
- Fox, E., & Riconscente, M. (2008). Metacognition and self-regulation in James, Piaget, and Vygotsky. *Educational Psychology Review*, 20, 373–389.
- 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.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge & Kegan Paul.
- Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, MA: MIT Press.
- Klahr, D. & Dunbar, K. (1988) Dual space search during scientific reasoning. *Cognitive Science*, 12 (1), 1-55.
- Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, 25, 111–146. doi: 10.1006/cogp.1993.1003.
- Kluwe, R. H. (1987). Executive decisions and regulation of problem solving behavior. In F. E. Weinert, & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 31-64). Hillsdale, NJ: Erlbaum.
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge: MIT press.
- Kuhn, D. (2000). Why development does (and doesn't) occur: Evidence from the domain of inductive reasoning. In R. Siegler & J. McClelland (Eds.), *Mechanisms of cognitive development: Neural and behavioral perspectives* (pp. 221-249). Mahwah, NJ: Erlbaum.

- Kuhn, D. (2010). What is Scientific Thinking and How Does it Develop? In U. Goswami (Ed.), *Handbook of Childhood Cognitive Development* (2nd ed., pp. 497–519). Oxford: Blackwell.
- Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. Orlando, FL: Academic Press.
- Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*, 18 (4), 495-523. Verkregen van: <http://www.jstor.org/stable/3233891>.
- Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Society for Research in Child Development Monographs*, 60 (4), serial no. 245.
- Kuhn, D., & Pease, M. (2006). Do children and adults learn differently? *Journal of Cognition and Development*, 7, 279-293.
- Kuhn, D., & Pease, M. (2008). What needs to develop in the development of inquiry skills? *Cognition and Instruction*, 26 (4), 512-559. doi: 10.1080/07370000802391745.
- Lavelli, M., Pantoja, A. P. F., Hsu, H., Messinger, D., & Fogel, A. (2005). Using microgenetic designs to study change processes. In D. M. Teti (Ed.), *Handbook of research methods in developmental science* (pp. 40-65). Malden, MA: Blackwell Publishing.
- Lazonder, A. W. (2014). Inquiry learning. In M. Spector, M. D. Merrill, J. Elen & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 453-464). Berlin: Springer. doi: 10.1007/978-1-4614-3185-5_36.
- Löhner, S., Van Joolingen, W. R., & Savelsbergh, E. R. (2003). The effect of external representation on constructing computer models of complex phenomena. *Instructional Science*, 31(6), 395–418.
- Mommers, J. C., Visser, C. de, & Warnars, R. L. J. (1977). *Uitkomst. Rekenmethode door een werkgroep onder leiding van C. Klavier. 6B. Repetitieboek: boek voor de onderwijzer*. Tilburg: Uitgeverij Zwijsen B.V.
- National Science Foundation (2000). Inquiry: Thoughts, Views, and Strategies for the K-5 Classroom, *Foundations*, 2, 1–5 (www.nsf.gov/pubs/2000/nsf99148/intro.htm).
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, 51, 102–116. <http://dx.doi.org/10.1037/0003-066X.51.2.102>

- Njoo, M., & de Jong, T. (1993). Exploratory learning with a computer simulation for control theory: learning processes and instructional support. *Journal of Research in Science Teaching*, 30, 821–844.
- Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart and Winston.
- Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences*, 1, 201–238.
- Siegler, R.S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed.) (pp. 464-510). Hoboken, NJ: Wiley.
- Sternberg, R. J. (1990). *Metaphors of the mind: Conceptions of the nature of intelligence*. Cambridge: Cambridge University Press.
- Van der Stel, M., & Veenman, M. V. J. (2008). Relation between intellectual ability and metacognitive skillfulness as predictors of learning performance of young students performing tasks in different domains. *Learning and Individual Differences*, 18, 128–134.
- Van der Stel, M., & Veenman, M. V. J. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences*, 20, 220–224.
- Van Hout-Wolters, B. (2000). Assessing active self-directed learning. In R. Simons, J. van der Linden, & T. Duffy (Eds.), *New Learning* (pp. 83–101). Dordrecht: Kluwer.
- Van Joolingen, W.R., De Jong, T., Lazonder, A.W., Savelsbergh, E.R., & Manlove, S. (2005). Co-Laboratory: research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, 21, 671-688.
- Veenman, M. V. J. (1993). *Intellectual ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*. Amsterdam: University of Amsterdam.
- Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multimethod designs? In C. Artelt, & B. Moschner (Eds), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 75–97). Berlin: Waxmann.
- Veenman, M. V. J. (2007). The assessment and instruction of self-regulation in computer-based environments: A discussion. *Metacognition and Learning*, 2, 177–183.
- Veenman, M. V. J. (2008). Giftedness: Predicting the speed of expertise acquisition by intellectual ability and metacognitive skillfulness of novices. In M. F. Shaughnessy,

- M. V. J. Veenman, & C. Kleyn-Kennedy (Eds.), *Meta-cognition: A recent review of research, theory, and perspectives* (pp. 207–220). Hauppauge: Nova Science Publishers.
- Veenman, M. V. J., Bavelaar, L., De Wolf, L., & M. G. P. Van Haaren (2014). The on-line assessment of metacognitive skills in a computerized environment. *Learning and Individual Differences*, 29, 123-130.
- Veenman, M.V.J., & Beishuizen, J.J. (2004). Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraint. *Learning and Instruction*, 14, 621–640.
- Veenman, M. V. J., & Elshout, J. J. (1999). Changes in the relation between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education*, 14, 509–523.
- Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs. domain specificity of metacognitive skills in novice learning across domains. *Learning and Instruction*, 7, 187–209.
- Veenman, M. V. J., Prins, F. J., & Verheij, J. (2003). Learning styles: Self-reports versus thinkaloud measures. *British Journal of Educational Psychology*, 73, 357–372.
- Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences*, 15, 159–176.
- Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition Learning*, 1, 3–14. doi: 10.1007/s11409-006-6893-0.
- Veenman, M. V. J., & Verheij, J. (2003). Identifying technical students at risks: Relating general versus specific metacognitive skills to study success. *Learning and Individual Differences*, 13, 259–272.
- Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction*, 14, 89–109. doi: 10.1016/j.learninstruc.2003.10.004.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*, 84, 30–43.
- Wilhelm, P. (2001). *Knowledge, skills and strategies in self-directed inductive learning*. (Proefschrift). Universiteit Leiden, Leiden.
- Wilhelm & Beishuizen (2003) Content effects in self-directed inductive learning. *Learning and Instruction*, 13, 381–402. doi:10.1016/S0959-4752(02)00013-0.

- Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, 8, 327–353.
- Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist*, 30, 217–221.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, 20, 99–149. doi: 10.1006/drev.1999.0497.

Bijlage 1: Scoringsprocedure begripsscore

Stap 1:

Zoek op het bandje het moment op waar de proefpersoon zijn/haar eindconclusie geeft (dat is het moment waarop (a) de proefpersoon aangeeft klaar te zijn of (b) de testtijd van 35 minuten voorbij is. Als de proefpersoon bij de eindconclusie aangeeft nog experimenten te willen doen om bepaalde effecten te onderzoeken en dat ook doet merk je de laatste conclusie op de band aan als deel uitmakend van de eindconclusie.

Stap 2:

Scoor de uitspraken die de proefpersonen doen aan de hand van de negen uitspraken die voor elke taak zijn beschreven.

Uitspraken kunnen **correct**, **gedeeltelijk correct** of **incorrect zijn**

- Iedere uitspraak krijgt twee punten

N.B.: het gaat erom dat de proefpersonen de correcte relaties weergeven.

Een uitspraak als: “Het blauwe huisje levert minder op dan het bruine huisje” (Vormen en Kleuren – II) is inhoudelijk hetzelfde als “Het bruine huisje levert meer op dan het blauwe huisje”, en krijgt dus ook twee punten.

- Gedeeltelijk correcte uitspraken krijgen 1 punt

Deze uitspraken hebben allemaal betrekking op uitspraken 6 t/m 9, deze kunnen nl. gedeeltelijk worden geuit, wat met 1 punt wordt gewaardeerd.

Uitspraken die niet te classificeren zijn schrijf je letterlijk uit in een personal file, deze files kunnen we later scoren op inhoud.

- Incorrect: 0 punten

Geef bij ieder van de negen uitspraken het aantal punten aan.

Je krijgt dus per proefpersoon een rijtje van negen scores:

	t1	t2	t3	t4	t5	t6	t7	t8	t9
pp 1	0	1	2	1	0	1	2	0	0

Maak een logboek waarin je bijhoudt welke scoringsbeslissingen/problemen etc. er in de loop van de tijd ontstaan.

S. Begripsscore berekenen

Otter

(non-causaal)

1. Vis bijvoeren of geen vis bijvoeren maakt niet uit.
2. Wel of geen media-aandacht maakt niet uit.

(hoofdeffect)

3. Eén of twee fabriekspijpen maakt niet uit.
4. Drie fabriekspijpen geeft minder otters dan twee.
5. Drie fabriekspijpen geeft minder otters dan één.

(interactie)

- Wel of geen toegang
- 6. Geen toegang of wel toegang maakt bij meerdere gebieden niet uit.
- 7. Geen toegang of wel toegang maakt wel uit bij één gebied.
- Gebieden
- 8. Eén gebied is beter dan meerdere gebieden zonder toegang.
- 9. Meerdere gebieden is beter dan één gebied met toegang.

Plant

(non-causaal)

1. Wel of geen bladluismiddel maakt niet uit.
2. Wel of geen dode bladeren in de pot maakt niet uit.

(hoofdeffect)

3. In een kas of op het balkon maakt niet uit.
4. In huis wordt de plant kleiner dan in de kas.
5. In huis wordt de plant kleiner dan op het balkon.

(interactie)

- Water geven
6. Eén keer per week of twee keer per week water maakt niet uit in een grote pot.
 7. Eén keer per week of twee keer per week water maakt wel uit in een kleine pot.
- Pot
8. Kleine pot is beter dan een grote pot met één keer per week water geven.
 9. Kleine pot is slechter dan een grote pot met twee keer per week water geven.

T. Focus variabele en hun codes (1, 2, 3, 4 en 5)

Otter

1	2	3	4	5
Gebieden	Toegang	Fabrieken	Vis	Media
(één groot = 0)	(geen = 0)	(drie = 0)	(wel = 0)	(wel = 0)
(meerdere = 1)	(wel = 1)	(twee = 1)	(niet = 1)	(niet = 1)

Plant

1	2	3	4	5
Pot	Water	Groeiplek	Bladluis	Dode bladeren
(klein = 0)	(1 keer = 0)	(kas = 0)	(niet = 0)	(wel = 0)
(groot = 1)	(2 keer = 1)	(huis = 1)	(wel = 1)	(niet = 1)

Bijlage 2: Scoringsprotocol

➔ De naam van de variabele in SPSS staat tussen haakjes

A. Proefpersoonnummer (ppn)

B. Sexe

Man: 1
Vrouw: 2

C. Groep

Groep 9: 2VWO
Groep 10: Studenten

D. Conditie

Aardrijkskunde - Biologie: 1
Biologie - Aardrijkskunde: 2

E. Starttijd nieuw experiment (startEXP)

Definitie: een experiment begint en eindigt op het moment dat een proefpersoon een niveau van de eerste variabele in een nieuw experiment aanklikt.

- voer het tijdstip (in minuten en seconden) in waarop bovenstaande gebeurt.
- meestal (maar niet altijd) is de vraag “Wat ga je nu uitzoeken?” het begin van een nieuw experiment

F. Experiment (experiment)

Voer het volgnummer in.

G. Plan (plan)

Scoor 1 van de volgende 6 categorieën:

Algemeen: overtuiging is belangrijk

1. Onderzoek het effect van een variabele
 (“Ik ga uitzoeken wat de fietsen uitmaken”)

De variabele waar het om draait (“focus”) moet hierbij ook gescoord worden (categorie H.).

2. Onderzoek het gezamenlijk effect van twee variabelen.

(“Ik ga uitzoeken wat de fietsen te maken hebben met de schoenen”)

De variabelen waar het om draait (“focus”) moeten hierbij ook gescoord worden (categorie H.).

- Plan 1 of 2 worden gescoord op basis van het antwoord op de vraag: “Wat ga je nu uitzoeken?” Als in het voorafgaand experiment gezegd wordt wat er nog moet worden uitgezocht en een ppn. gaat dat (op basis van de logfiles) ook doen, dan scoor je ook plan 1 of 2.

- Als iemand netjes aan het experimenteren is (bv. als door de keuze van de variabelen duidelijk wordt welke plan er wordt gevolgd ook als de variabele niet expliciet wordt genoemd) wordt eerder plan 1 of 2 gescoord dan plan 6. Echter, de log-files dienen alleen ter ondersteuning van de protocollen. Enkel en alleen van de log-files kan niet afgeleid worden of er een plan en welk plan er wordt gevolgd.

3. Een specifieke uitkomst genereren (= doel).

Middel: geleerde regels toepassen.

(als het plan is om een specifieke uitkomst te genereren, dan moet die ook genoemd worden. Er mag niet alleen gezegd worden dat er een heel goede of slechte uitkomst gegenereerd gaat worden).

4. Verificatie - I

Doel: kijken of het effect van een bepaalde variabele hetzelfde is in een ander experiment (regel t.a.v. één variabele verifiëren). Middel: nieuw experiment.

Deze categorie heeft een focus en zeer waarschijnlijk ook een hypothese.

5. Verificatie - II

Doel: de uitkomst van een rijtje te voorspellen op basis van wat je weet (regels t.a.v. alle variabelen toetsen). Middel: nieuw experiment

Deze categorie heeft geen focus!

6. Overig

- Informatie verzamelen

Doel: nieuwe informatie verzamelen, bijv. omdat iemand is vastgelopen

Voorbeeld: "Kijken wat er hier uitkomt."

- Als de plan-vraag niet is gesteld of uit de experimenten niet duidelijk is wat pp. van plan is.
- Idiosyncratisch plan (plan wat niet te scoren is in 1 - 5)

H. Focus plan: welke variabelen zijn bij het plan betrokken? (focusVAR)

(geldt alleen als plan = 1, 2 of 4, anders focus plan = 99).

Focus variabele x: x (1,2,3,4,5)

Focus variabele x en y: xy

Geen focus (of n.v.t.): 99

X en Y verwijzen naar de nummers die onder punt T. worden genoemd!

I. Hypothesen stellen (hypothese)

Ten aanzien van variabele x: x

Ten aanzien van variabele x en y: xy

Geen hypothese (of n.v.t.): 99

Wanneer is een uitspraak een hypothese?

Algemeen: Overtuiging en duidelijkheid zijn belangrijk

Criterium: er moet worden verwezen naar specifieke variabelen en aangegeven worden dat (minimaal 1 van onderstaande situaties):

- een variabele een verschil maakt.
- een variabele geen verschil maakt.
- het effect van een variabele een bepaalde richting heeft.
- het effect een bepaalde grootte heeft.

De uitspraak: "als dit experiment een andere uitkomst oplevert dan maakt die variabele verschil", merk je niet aan als hypothese.

Niet scoren: hypothesen ten aanzien van meer dan twee variabelen (= 99)

X en Y verwijzen naar de nummers die onder punt T. worden genoemd!

J. Inferentie: conclusies (inferentie)

Algemeen: overtuiging is belangrijk

Inclusie inferentie: 1

Exclusie inferentie: 2 (t.a.v. variabele 3: ook scoren)

Interactie: 3 (de uitspraak: “Er is geen interactie tussen...” ook zo scoren)
Geen inferentie 0

- Er moet een verwijzing naar een variabele zijn (dus niet enkel een niveau) of een indicatie dat er sprake is van twee niveaus (bijv. door gebruik vergrotende trap, “verschil” of “bepalend”). Als er bijv. wordt gezegd “Meerdere gebieden zijn beter” is dat al een inferentie.

- Scoor alleen duidelijke/overtuigde inferenties, die volgen op de inferentie-vraag. Scoor samenvattingen van eerdere conclusies (bijv. bij het maken van aantekeningen) niet, tenzij een eerdere conclusie wordt veranderd of een nieuwe inferentie wordt gegeven voor de (nieuwe) inferentie-vraag is gesteld. Wanneer iemand netjes experimenteert scoor je sneller een inferentie.

Een interactie effect gaat voor een hoofdeffect op één van de betrokken variabelen (eerdere conclusie wordt dan a.h.w. veranderd).

- Wanneer een eerdere inferentie correct is maar niet valide en daarna volgt dezelfde inferentie maar dan valide, dan scoor je deze allebei.

- Wanneer iemand zegt: “Ik heb uitgevonden dat de uitkomsten verschillend zijn”, of “Deze (dit experiment) is minder dan die”, dan scoor je bij inferentie: 0. Het gaat nl. alleen maar om inferenties ten aanzien van één of meerdere variabelen. Als er geen variabelen worden genoemd is er dus geen inferentie gemaakt.

- Wanneer iemand zegt: “X en Y variabele zijn bepalende factoren”, dan scoor je een inferentie. Dit is een inclusie-inferentie

Wanneer een pp. bij een experiment meerdere inferenties maakt moeten deze inferenties in een bepaalde volgorde worden ingevoerd, nl. de variabele met de laagste focuswaarde eerst. Bij een interactie staat het kleinste getal altijd vooraan in de focus (dus altijd 12 en nooit 21 schrijven). Het eerste en het tweede getal zijn bepalend voor de volgorde waarin de inferenties worden opgetekend in de datafile.

Voorbeeld: 1 1 1
 12 2 4
 23 3 5
 3 4 etc.

Bij twee inferenties t.a.v een variabele (komt voor bij variabele 3) bepaalt de code van het type inferentie de volgorde in de datafile. Dat betekent dat een inclusie-inferentie (1) eerst wordt ingevoerd en daarna de exclusie-inferentie (2).

Wanneer iemand een inclusie/exclusie inferentie maakte en daarbij zei: “in deze combinatie”, “met de rondjes” etc. heb ik niet anders gescoord. Deze toevoegingen zijn conservatief en dus prima.

K. Focus inferentie (effect codes) (focusINF)

Focus variabele x: x

Focus variabele x en y: xy

N.v.t.: 99

Criterium: wordt alleen gescoord als Plan 1, 2 of 4 gevolgd worden.

X en Y verwijzen naar de nummers die onder punt T. worden genoemd!

L. Valid: gegeven verzamelde evidentie (validINF)

Indien valide: 1

Indien invalide: 0

N.v.t. (als er geen inferentie werd gemaakt) 99

Invalide zijn die inferenties die niet stroken met gegenereerde evidentie. Meestal betekent dit dat geen rekening wordt gehouden met het constant houden van variabelen (veronderstelde irrelevante effecten worden geacht ook niet te interacteren). Soms wordt een eerdere (valide) inferentie veranderd. Dit gebeurt voornamelijk bij de interacterende variabelen. Een zodanig inferentie is invalide als er elders het (valide) tegendeel ligt. Om de validiteit te beoordelen moet naar alle eerdere experimenten gekeken worden.

M. Correctness (t.a.v. het model in de taak) (correctINF)

Correct:	1
Incorrect:	0
N.v.t.:	99

Bij deze categorie staat vaak een “00” bij valide inclusie/exclusie inferenties ten aanzien van de interacterende variabelen. Die vind ik niet zo interessant, het geeft alleen maar aan dat pp. op dat moment op een dwaalspoor kunnen zitten. Met name interessant zijn de interactie - inferenties (12) die valide zijn en correct volgens het model. Als iemand een interactie onderzoekt en concludeert: er is geen interactie tussen variabele 4 en 5 (klopt) en dat valide weet af te leiden, dan scoor je een interactie - inferentie die valide is en correct volgens het model.

N. Causaal Mechanisme (CM)

Aanwezig:	1
Afwezig:	0

Algemeen: overtuiging is belangrijk.

Definitie: een causaal mechanisme ondersteunt een hypothese en biedt een verklaring voor een verband tussen variabelen. Er moet verwezen worden naar een specifieke hypothese over een of twee variabelen.

- Bij het genereren van een causaal mechanisme wordt aanvullende informatie gegeven om de relatie tussen twee variabelen te verklaren (bijv. kleine pot en veel water is slecht omdat de plant dan verdrinkt).
- Het causale mechanisme kan aangepast worden aan de uitkomsten van de experimenten.
- Het causale mechanisme kan spontaan of als reactie op een prompt gegenereerd worden.

O. Frequentie Causaal Mechanisme (CMfrequentie)

Aantal CM:	1, 2, 3, ...
Nvt.:	0

Het gaat hier om het aantal gegenereerde causale mechanismen per experiment.

- Als er geen causaal mechanisme wordt gegenereerd, scoor je een “0”.
- De causale mechanismen worden per experiment gescoord (zie definitie punt E)

P. Uniekheid van het causaal mechanisme (CMrang)

Niet uniek:	0
Uniek:	rangnummer geven (1, 2, 3,)

Elk nieuw causaal mechanisme (in relatie tot alle vorige) krijgt een nieuw rangnummer.

- Als een unieke causaal mechanisme wordt gegenereerd wordt een rangnummer gegeven, altijd een punt hoger dan de laatste waarde in de SPSS-kolom
- Als er meerdere (unieke) causale mechanisme worden gegenereerd in een experiment wordt dat aantal opgeteld bij het hoogste rangnummer dat in de SPSS-kolom staat. Het resulterende aantal wordt ingevoerd bij het experiment.
- Als een causaal mechanisme niet meer uniek is wordt het genegeerd

Q. Evidentie betrekken bij het opstellen van een theorie (hypothese + causaal mechanisme)

Betrekken: 1
Niet betrekken: 0
Nvt: 99

Een causaal mechanisme kan het experimenteelgedrag beïnvloeden.

Criterium: het experimenteelgedrag in het eerstvolgende experiment of daarna wordt aangepast op basis van een gegenereerd causaal mechanisme (deductief proces).

- Het is niet voldoende een causaal mechanisme te genereren op basis op het resultaat van het rijtje waarmee de proefpersoon bezig is. Dit wordt al door punt N uitgedrukt.
- Als er geen causaal mechanisme wordt gegenereerd is het niet van toepassing en er wordt 99 gescoord.