

ONDERZOEKSCOMPETENTIES VAN VO-LEERLINGEN

Ontwikkeling van een valide en betrouwbaar meetinstrument dat de perceptie van VO-leerlingen op de eigen onderzoekscompetenties meet

**Leonie A.S. van Eeten
Mei 2009**

Onderzoek in het kader van een Bacheloropdracht aan Universiteit Twente bij de afdeling Curriculumontwerp & Onderwijsinnovatie van de opleiding Onderwijskunde en in opdracht van het lerareninstituut ELAN.

Samenvatting

In dit artikel staat de ontwikkeling van een valide en betrouwbaar meetinstrument centraal. Het gaat om een meetinstrument dat de perceptie van leerlingen in het voortgezet onderwijs (VO-leerlingen) op hun eigen onderzoekscompetenties meet. Het onderzoek concentreert zich op onderzoekscompetenties om hypothesetoetsend onderzoek uit te voeren voor de vakken natuurkunde, scheikunde en biologie. Het doel is dit meetinstrument in te kunnen zetten in combinatie met een meetinstrument dat het daadwerkelijk bezitten van onderzoekscompetenties meet om zo handvaten te verkrijgen voor een kwalitatief goede invulling van der leerlijn 'onderzoeken'.

Als basis voor het meetinstrument is door middel van een literatuurstudie een competentiematrix ontwikkeld. Deze is geëvalueerd met experts en geeft een beschrijving van de onderzoekscompetenties die VO-leerlingen nodig hebben. De beschrijving bestaat uit 12 competenties, verdeeld over 5 onderzoeksfasen. Daarnaast is elke competentie uitgewerkt in een aantal aspecten die de benodigde kennis, vaardigheden en houdingen omschrijven.

Door het meetinstrument op deze competentiematrix te baseren is een eerste stap gezet om de inhouds- en begripsvaliditeit van het meetinstrument te waarborgen. Om deze typen validiteit verder te kunnen waarborgen is het meetinstrument vervolgens ook geëvalueerd met experts.

De op deze manier ontstane vragenlijst is door 133 leerlingen ingevuld. Door de resultaten statistisch te analyseren zijn de betrouwbaarheid, de criterium- en de discriminantvaliditeit bestudeerd. Hieruit bleek dat 6 van de 12 bevraagde competenties voldoende betrouwbaar en criteriumvalide zijn (Cronbach's Alpha > 0,65) om de vragenlijst in te zetten voor onderzoek. Bij 4 competenties is nog enige verbetering nodig (CA 0,60 – 0,65) en bij 2 andere competenties dient nog een grondige verbetering gemaakt te worden (CA 0,35 en 0,52). Deze verbeteringen kunnen in de meeste gevallen gemaakt worden door items toe te voegen. Bij 2 competenties wordt aangeraden een factoranalyse uit te voeren om te controleren of verschillende factoren binnen de competenties de lage waarde van Cronbach's Alpha kunnen veroorzaken.

Uit de uitgevoerde analyses kan niet worden opgemaakt dat de vragenlijst discriminantvalide is. Het is niet duidelijk of het meetinstrument werkelijk geen verschil meet tussen verschillende groepen. Er kan ook geconcludeerd worden dat er geen verschil gemeten wordt, omdat er geen of te weinig verschil is.

Er wordt aangeraden om de vragenlijst aan te passen, opnieuw af te nemen onder een grotere en diversere groep leerlingen en vervolgens nogmaals de criterium- en discriminantvaliditeit te bestuderen.¹

¹ Met dank aan de begeleiders van mijn Bacheloropdracht: vanuit Universiteit Twente, dr. H.B. Westbroek (eerste begeleider) en dr. J.M. Voogt (tweede begeleider) en vanuit de opdrachtgever, ELAN, ir. F.J. Carelsen en dr. ir. H.P. Hendrikse.

Inleiding

Onderzoeksvaardigheden in het curriculum

In 1998 is de Tweede Fase ingevoerd in het voortgezet onderwijs. Daarmee is er in de eindexamenprogramma's meer aandacht gekomen voor de vaardigheden van leerlingen op het gebied van 'onderzoek doen' (Hubers, 2003). Wat die vaardigheden inhouden kan verschillen per vak. Dit komt doordat bij verschillende vakken verschillende typen onderzoek van belang zijn. Zo is bij aardrijkskunde beschrijvend onderzoek vaak van belang. Beschrijvend onderzoek richt zich op het in kaart brengen van (een deel van) de werkelijkheid (Schalk, 2006).

Het onderzoek dat beschreven wordt in dit artikel concentreert zich op hypothesetoetsend onderzoek, dat vooral van belang is voor de natuurwetenschappelijke vakken natuurkunde, scheikunde en biologie. Bij hypothesetoetsend onderzoek wordt een uitspraak gedaan over een mogelijk verband tussen verschijnselen en een mogelijke verklaring van een verschijnsel (hypothese). Die hypothese wordt vervolgens getoetst (Schalk, 2006).

Als in het vervolg over onderzoeksvaardigheden gesproken wordt gaat het om onderzoeksvaardigheden voor hypothesetoetsend onderzoek voor de natuurwetenschappelijke vakken natuurkunde, scheikunde en biologie in het voortgezet onderwijs. De onderzoeksvaardigheden die een eindexamenkandidaat voor die vakken moet bezitten worden in de eindexamenprogramma's sinds de invoering van de Nieuwe Tweede Fase in 2007 als volgt geformuleerd (CEVO², 2005, p.4):

“De kandidaat kan een natuurwetenschappelijk onderzoek voorbereiden, uitvoeren, de verzamelde onderzoeksresultaten verwerken en hieruit conclusies trekken.”

Globale leerlijn

Voor veel eindtermen wordt een leerlijn ontwikkeld, waarin wordt omschreven hoe een leerling op het gewenste eindniveau kan komen. Zo heeft het SLO ook een globale invulling voor de leerlijn onderzoeken en ontwerpen geschreven (SLO, 2004). Er is echter nog weinig bekend over hoe de leerlijn exact en met voldoende kwaliteit ingevuld kan worden. Zo geven docenten in onderzoek van Hubers (2003) aan dat zij ondanks de invoering van de Tweede Fase nog weinig tot niets hebben geleerd over de didactiek van leren onderzoeken.

Volgens Hubers (2003) geven docenten tevens aan dat zij door tijdsgebrek niet toekomen aan meer open onderzoeksvormen dan de zogenaamde 'kookboekpractica', waarbij een experiment exact volgens de voorschriften wordt uitgevoerd. Deze 'kookboekpractica' kunnen wel nuttig zijn voor het veilig en zorgvuldig leren werken en om vertrouwd te raken met practicabehoeften (SLO, 2000 in Hubers, 2003). Van den Berg & Buning (1994, in Hubers, 2003) beredeneren echter dat 'kookboekpractica' geen goede resultaten opleveren als docenten zowel praktische doelen als cognitieve en metacognitieve doelen willen bereiken.

Het laatste onderdeel van de leerlijn 'onderzoeken' is het profielwerkstuk. Hierbij wordt van leerlingen verwacht dat zij zelfstandig vragen kunnen vertalen naar toetsbare hypothesen, een natuurwetenschappelijk onderzoek kunnen opzetten en uitvoeren, een keuze kunnen maken tussen de beschikbare meetinstrumenten en om kunnen gaan met informatiebronnen (SLO, 2004). Dit is een heel andere invulling van 'onderzoek' dan de 'kookboekpractica'. Hierdoor kan geconcludeerd worden dat leerlingen niet voldoende worden voorbereid op de onderzoeksvaardigheden die zij nodig hebben voor het profielwerkstuk.

ELAN, het lerareninstituut van Universiteit Twente, wil docenten graag ondersteunen bij het doceren van onderzoeksvaardigheden. Daarom is een onderzoek gestart naar hoe de leerlijn onderzoeksvaardigheden het beste verder ingevuld kan worden om leerlingen op de juiste manier te begeleiden.

² CEVO heeft de eindexamenprogramma's voor havo en vwo voor het vak *biologie* herschreven voor de invoering van de Nieuwe Tweede Fase in 2007. Het domein 'vaardigheden' waar het subdomein 'onderzoeksvaardigheden' onder valt is echter voor alle natuurwetenschappelijke vakken (natuurkunde, scheikunde, biologie en algemene natuurwetenschappen) identiek (CEVO, 2005).

Invulling leerlijn

Om er achter te komen hoe de leerlijn onderzoeken het beste ingevuld kan worden, wordt door ELAN vanuit verschillende perspectieven gekeken naar de behoefte aan begeleiding bij het ontwikkelen van onderzoeksvaardigheden. Allereerst wordt de *perceptie* van leerlingen in het voortgezet onderwijs (VO-leerlingen) onderzocht: Welke onderzoeksvaardigheden denken zij bij aanvang van het profielwerkstuk voldoende te bezitten en welke nog niet? Daarnaast wordt ook onderzocht welke onderzoeksvaardigheden die VO-leerlingen *daadwerkelijk* bezitten.

De onderzoeksvaardigheden die zij nog niet voldoende (denken te) bezitten zijn een aanwijzing voor de punten die uitgebreider aan bod moeten komen in de leerlijn dan nu het geval is in het curriculum.

Het meten van competenties

Dit onderzoek ondersteunt bij het onderzoeken van het eerste perspectief: de perceptie van VO-leerlingen op hun verschillende onderzoekscompetenties. Om die te kunnen meten is een valide en betrouwbaar meetinstrument nodig.

Er wordt hier gesproken van competenties in plaats van vaardigheden, omdat dit een completer beeld geeft van wat er van leerlingen verwacht wordt. Onder competenties wordt hier namelijk “het geheel van kennis, vaardigheden en houdingen” verstaan (Stokking & Van der Schaaf, 2000).

De ontwikkeling van het meetinstrument wordt in dit artikel beschreven. De hoofdvraag is dan ook: “*Wat is een valide en betrouwbaar meetinstrument voor de perceptie van VO-leerlingen op hun onderzoekscompetenties?*”

1. Theoretische kader

1.1 Competenties

De term ‘competentie’ is tegenwoordig in het onderwijs een veel gebruikt begrip. Er worden in de literatuur, naast bovengenoemde definitie ook veel andere definities van gegeven. In de meeste gevallen wordt het begrip competentie echter gebruikt om te verwijzen naar het geheel van kennis, vaardigheden en attitudes (Hubers, 2003). Zo omschrijven Verloop en Lowyk (2003) een competentie als ‘een bekwaamheid waarin kennisaspecten, vaardigheidsaspecten en attitudeaspecten zijn opgenomen’ (cf. Van Rens, 2005). Omdat het onderscheid tussen kennis, vaardigheid en attitude zo gangbaar is, wordt dit onderscheid ook gemaakt in de in dit artikel gehanteerde definitie.

Bovendien maakt de definitie “het geheel van kennis, vaardigheden en houdingen” het mogelijk een duidelijk beeld te schetsen van wat een bepaalde competentie inhoudt. Er kunnen op deze manier namelijk verschillende componenten binnen een competentie worden onderscheiden. In dit onderzoek worden binnen de componenten kennis, vaardigheden en houdingen nog specifiekere aspecten onderscheiden. Dit gebeurt door onderscheid te maken tussen ‘cognitieve’, ‘procedurele’, ‘metacognitieve’ en ‘communicatieve’ aspecten.

De cognitieve aspecten maken deel uit van de kenniscomponent. Hiermee wordt het conceptueel begrip van de natuurwetenschappen aangeduid dat een leerling moet bezitten om procedurele handelingen correct uit te kunnen voeren.

De procedurele en communicatieve aspecten maken deel uit van de vaardighedencomponent. Met procedurele aspecten worden de onderzoekshandelingen aangeduid die een leerling moet kunnen uitvoeren om een competentie te bezitten. Het gaat hierbij enkel om het *kunnen* uitvoeren van de nodige stappen in onderzoek. Schalk, Van der Schee en Boersma (2007) definiëren dit als het uitvoeren van procedures om zo wetenschap in de praktijk te brengen. De communicatieve aspecten verwijzen ook naar vaardigheden, maar van een ander type. Hierbij gaat het namelijk niet om een onderzoekshandeling, maar om het rapporteren *over* onderzoek.

De metacognitieve aspecten maken deel uit van de houdingen- en de kenniscomponent. Met deze aspecten wordt aangeduid wat een leerling moet doen of welke houding een leerling moet aannemen om de kwaliteit van zijn eigen onderzoek te bewaken (Brown, Bransford, Ferrara & Campione, 1983 in White & Frederiksen, 1998). Als een leerling iets moet doen om de kwaliteit van zijn eigen onderzoek te bewaken, sluit dit aan bij de kenniscomponent. Het gaat dan namelijk om het reguleren van de eigen cognitieve processen door ze te organiseren, te bewaken en indien nodig aan te

passen met als doel bepaalde leeruitkomsten (Smith & Ragan, 2005). Als de leerling een bepaalde *houding* moet aannemen om de kwaliteit van zijn eigen onderzoek te waarborgen, sluit dit aan bij de houdingen- of attitudecomponent.

1.2 Validiteit en betrouwbaarheid meetinstrument

Twee belangrijke aspecten in de hoofdvraag zijn validiteit en betrouwbaarheid. De vraag die dan rijst is wanneer het meetinstrument *valide* en *betrouwbaar* is. De validiteit van het meetinstrument heeft te maken met de betekenis, de bruikbaarheid en de juistheid van de conclusies die uit de resultaten worden getrokken (Van Berkel & Bax, 2006). Van Berkel en Bax (2006) onderscheiden drie typen validiteit, namelijk

- *Inhoudsvaliditeit*: Deze validiteit geeft aan in welke mate het meetinstrument representatief is voor datgene wat men ermee wil nagaan. Dit betekent onder andere dat ieder belangrijk onderwerp in de vragenlijst moet worden bevraagd.
- *Begripsvaliditeit*: Deze validiteit geeft aan in welke mate de resultaten kunnen worden beschouwd als een maat voor het bedoelde theoretisch begrip of de bedoelde vaardigheid. Zo is het niet valide om inzicht te toetsen door te vragen naar een uit het hoofd geleerde definitie.
- *Criteriumvaliditeit*: Deze validiteit geeft aan in welke mate de resultaten samenhangen met soortgelijke metingen. Zo dienen bijvoorbeeld scores op vragenlijsten over hetzelfde onderwerp aan elkaar te zijn gerelateerd.

Dooley (2001) onderscheid tevens discriminantvaliditeit. Deze validiteit is volgens Dooley (2001) gekoppeld aan de criteriumvaliditeit en helpt te beschermen tegen methode-effecten. Er treedt een methode-effect op als bij verschillende metingen met dezelfde methode de resultaten overeenkomen, terwijl er juist verschil verwacht wordt. De overeenkomst in resultaten wordt dan veroorzaakt doordat er dezelfde methode wordt gehanteerd (Dooley, 2001). Er is sprake van discriminantvaliditeit als er een verschil gemeten wordt, als er een verschil verwacht wordt.

De betrouwbaarheid van een vragenlijst geeft aan in hoeverre er vertrouwen kan zijn in de vragenlijst als meting, ongeacht de inhoud van de vragenlijst. Berkel & Bax (2006, p.45) formuleren dit in de vraag ‘Meet de toets echt iets of kan er net zo goed worden gedubbeld om de uitspraken te doen over studenten?’. Betrouwbaarheid is een voorwaarde voor validiteit. Hierdoor zijn betrouwbaarheid en validiteit sterk met elkaar verbonden.

1.3 Onderzoeksvragen

Bij de geformuleerde hoofdvraag kunnen nu de volgende deelvragen worden geformuleerd.

1. Wat is een inhoudsvalide meetinstrument voor de perceptie van VO-leerlingen op hun onderzoekscompetenties?
2. Wat is een begripsvalide meetinstrument voor de perceptie van VO-leerlingen op hun onderzoekscompetenties?
3. Wat is een criteriumvalide meetinstrument voor de perceptie van VO-leerlingen op hun onderzoekscompetenties?
4. Wat is een discriminantvalide meetinstrument voor de perceptie van VO-leerlingen op hun onderzoekscompetenties?

In deze deelvragen komt het begrip ‘betrouwbaarheid’ niet terug. De reden hiervoor is dat betrouwbaarheid sterk samenhangt met de criteriumvaliditeit. Die vorm van validiteit wordt in dit onderzoek bestudeerd door Cronbach’s Alpha, een maat voor de betrouwbaarheid, te berekenen (zie ook paragraaf 2.3). Ook bij de andere vormen van validiteit is betrouwbaarheid als het ware de fundering (Berkel & Bax, 2006). Door de tijdsspanne waarbinnen dit onderzoek plaats diende te vinden, was het echter niet mogelijk de betrouwbaarheid verder te onderzoeken.

1.4 Leeswijzer

De methoden die gehanteerd zijn om de deelvragen te beantwoorden worden in hoofdstuk twee in de opeenvolgende paragrafen toegelicht. De resultaten voor de deelvragen worden vervolgens in hoofdstuk drie in de opeenvolgende paragrafen toegelicht. De conclusies worden per deelvraag toegelicht in hoofdstuk vier. Tot slot worden de bevindingen in het onderzoek besproken in hoofdstuk vijf.

2. Methode

De ontwikkeling van de vragenlijst en de controle op validiteit zijn in dit onderzoek hand in hand gegaan. Telkens als een volgend onderdeel voor de vragenlijst ontwikkeld was, is dit onderdeel geëvalueerd om de validiteit te waarborgen. Hierdoor bestond de ontwikkeling van de vragenlijst uit drie delen. Eerst is een competentiematrix ontwikkeld waarop de vragenlijst vervolgens is gebaseerd. Bij deze twee stappen is de inhouds- en begripsvaliditeit gecontroleerd. De derde stap omvatte de controle op criterium- en discriminantvaliditeit door afname van de vragenlijst onder leerlingen.

2.1 Competentiematrix

Om te zorgen voor een inhouds- en begripsvalide meetinstrument is als basis een competentiematrix ontwikkeld om een overzicht te krijgen van de onderzoekscompetenties die VO-leerlingen nodig hebben. Deze competentiematrix kan gezien worden als een specificatietabel van de onderwerpen die bevestigd dienen te worden in de vragenlijst. Van Berkel & Bax (2006) geven aan dat een dergelijke specificatietabel een hulpmiddel kan zijn om er voor te zorgen dat een vragenlijst inhouds- en begripsvalide is. Uit zo'n specificatietabel zijn dan de belangrijkste onderwerpen die gemeten moeten worden en het niveau waarop de onderwerpen gemeten moeten worden af te leiden.

Voor de begripsvaliditeit is het belangrijk dat het meetinstrument perceptie op *competenties* meet, en bijvoorbeeld niet alleen theoretisch kennis bevestigt. Dit wordt gewaarborgd door een competentiematrix als basis te gebruiken. De competentiematrix geeft een overzicht van vijf fasen in onderzoek, de onderzoekcompetenties die bij elke fase horen en er wordt per competentie onderscheid gemaakt tussen cognitieve, procedurele, metacognitieve en communicatieve aspecten van die competentie.

De competentiematrix is tot stand gekomen op basis van een literatuuronderzoek³. De op deze manier ontstane matrix is vervolgens in interviews geëvalueerd met experts om de inhouds- en begripsvaliditeit zo goed mogelijk te kunnen waarborgen.

Respondenten

Om vanuit verschillende perspectieven feedback op de competentiematrix te krijgen zijn drie docenten uit het voortgezet onderwijs en een wetenschappelijk onderzoeker benaderd.

Twee van de drie docenten nemen actief deel aan het project 3xO (Onderzoeken, Ontwerpen en Ondersteunen) dat ELAN in samenwerking met Universiteit Twente heeft opgezet. Dit project heeft als doel invulling te geven aan de leerlijn 'onderzoeken' en zo de leerlingen uit de natuurprofielen te leren onderzoeken en ontwerpen als voorbereiding op het profielwerkstuk. Doordat de twee docenten actief deelnemen aan dit project kan verwacht worden dat zij veelvuldig hebben nagedacht over de invulling van een leerlijn 'onderzoeken'. Met deze twee docenten zijn de vakken scheikunde en biologie vertegenwoordigd in het geraadpleegde docententeam. Om ook vanuit het perspectief van natuurkunde naar de matrix te kijken is ook een natuurkundedocent bij het docententeam betrokken. Het was helaas niet mogelijk om hiervoor iemand uit het 3xO-project te benaderen, maar de docent is wel werkzaam op één van de scholen die meewerken aan het 3xO-project. De drie docenten komen van drie verschillende scholen in Overijssel en Gelderland van verschillende denominaties.

De geïnterviewde wetenschappelijk onderzoeker is gespecialiseerd in het thema Leren Onderzoeken in het voortgezet onderwijs. Hij is bioloog en vakdidactisch onderzoeker op de ULO afdeling van de Vrije Universiteit (VU) te Amsterdam.

Aanpak

De competentiematrix is geëvalueerd met experts door drie interviews te houden, namelijk

1. een interview met de twee docenten uit het 3xO-project;
2. een interview met de derde docent;
3. en een interview met de wetenschappelijk expert.

In eerste instantie is getracht een groepsinterview met de docenten te plannen. Door het korte tijdsbestek waarbinnen de interviews plaats dienden te vinden, was het echter niet mogelijk om de drie

³ De literatuurstudie die geleid heeft tot de competentiematrix is op te vragen bij de auteur: lvaneeten@gmail.com

docenten op hetzelfde moment te interviewen. Om die reden is het interview met de docenten opgesplitst in twee losse interviews.

Er werd van de experts verwacht dat zij concrete verbeterpunten in de competentiematrix konden aanwijzen en/of concrete aanvullingen op de competentiematrix konden doen, om zo de inhouds- en begripsvaliditeit te verbeteren.

Omdat verwacht werd dat de experts nieuwe informatie konden leveren zijn geen gestandaardiseerde vragenlijsten voor de interviews ontwikkeld. Met gestandaardiseerde vragenlijsten zou te veel nuttige informatie niet aan bod kunnen komen, doordat het niet in de vragen is opgenomen. Er is wel een interviewagenda voor alle interviews gemaakt, die diende als checklist voor de onderwerpen die in ieder geval besproken dienden te worden.⁴ De daadwerkelijke vragen zijn pas tijdens de interviews geformuleerd.

Een aanvullende, methodologische reden om voor een interviewagenda in plaats van een gestandaardiseerde vragenlijst te kiezen is dat hierdoor de gesprekken op een natuurlijkere wijze verlopen en dat op- en aanmerkingen van de geïnterviewden op die manier zo ruim mogelijk aan bod kunnen komen in de interviews (Dooley, 2001).

Alle interviews zijn gestart met een korte uitleg van het onderzoek. Vervolgens is de respondenten gevraagd hun globale indruk van de competentiematrix te beschrijven. Daarna is dieper ingegaan op elke fase van onderzoek die de competentiematrix beschrijft, waarbij alle onderdelen van de matrix zijn besproken. Tot slot is de respondenten gevraagd om tips te geven voor de vertaling van de competentiematrix in een vragenlijst.

De gesprekken zijn opgenomen door middel van audio-apparatuur en er zijn tijdens de gesprekken aantekeningen gemaakt. De aantekeningen zijn na afloop van de interviews met behulp van de audio-opname uitgewerkt en voorgelegd aan de respondenten.

Het was niet nodig de audio-opnames geheel uit te typen, omdat de informatie uit de interviews niet op gestandaardiseerde wijze verwerkt diende te worden met bijvoorbeeld SPSS, zoals soms bij grote datasets het geval is. Het uittypen van met audio-apparatuur opgenomen gesprekken is bovendien erg tijdrovend (Dooley, 2001).

De competentiematrix is vervolgens aangepast op basis van de interviews. De aangepaste competentiematrix is opnieuw voorgelegd aan de respondenten, om na te gaan of zij zich in de aanpassingen en aanvullingen konden vinden. Een deel van de competentiematrix wordt weergegeven in Figuur 1.

⁴ De interviewagenda's voor de interviews over de competentiematrix zijn op te vragen bij de auteur.

Figuur 1 Competentiematrix Onderzoekscompetenties Fase 1: voorbereiden

Onderzoeks-competenties	Procedurale aspecten	Cognitieve aspecten	Metacognitieve aspecten	Communicatieve aspecten
Onderwerp/probleem bepalen	<ul style="list-style-type: none"> Onderwerp/probleem bedenken Onderwerp/probleem verkennen door relevante informatie op te zoeken Onderwerp/probleem verkennen door relevant eerder onderzoek op te zoeken 	<ul style="list-style-type: none"> Natuurwetenschappelijk probleem kunnen herkennen en specificeren Onderscheid tussen beschrijvend en hypothesetoetsend onderzoek kennen Onderzoekbaar probleem kunnen formuleren Gebruik kunnen maken van onderwerpspecifieke concepten 	<ul style="list-style-type: none"> Antwoord op een vraag willen 	<ul style="list-style-type: none"> Onderwerp/probleem helder en eenduidig formuleren De relevantie van je onderzoek/probleem onderbouwen
Onderzoeksvraag bepalen	<ul style="list-style-type: none"> Onderzoeksvraag bedenken 	<ul style="list-style-type: none"> Weten wanneer de onderzoeksvraag voldoende specifiek en ingeperkt is 	<ul style="list-style-type: none"> Zorgen dat de onderzoeksvraag beantwoordbaar is d.m.v. onderzoek 	<ul style="list-style-type: none"> Onderzoeksvraag helder en eenduidig formuleren
Hypothese(n) bepalen	<ul style="list-style-type: none"> Hypothese(n) opstellen 	<ul style="list-style-type: none"> Weten wat de mogelijke antwoorden zijn op de onderzoeksvraag 	<ul style="list-style-type: none"> Zorgen voor toetsbare hypothese(n) 	<ul style="list-style-type: none"> Hypothese(n) helder en eenduidig formuleren
Onderzoeksoptzet en -aankpak bepalen	<ul style="list-style-type: none"> Onafhankelijke, afhankelijke en controlevariabelen bepalen Apparatuur selecteren Werkplan maken voor verwerven en verwerken data 	<ul style="list-style-type: none"> Onderscheid kennen tussen verschillende soorten variabelen Weten op welke manier je hypothesen kunt testen Weten bij welke uitkomst je je hypothese moet verwerpen (en wanneer niet) Weten wanneer een blanco proef nodig is Weten bij welke opzet je meet wat je wilt meten (validiteit) Weten welke metingen je moet doen Weten welke apparaten geschikt zijn 	<ul style="list-style-type: none"> Gebruik maken van <i>geschikte</i> apparatuur Inventief proberen te zijn Willen bewijzen dat de hypothese klopt Zorgen voor een aselechte steekproef die groot genoeg is Zorgen voor voldoende metingen 	<ul style="list-style-type: none"> Duidelijk werkplan schrijven zodat het onderzoek reproduceerbaar is

2.2 Ontwikkeling vragenlijst

De competentiematrix die op bovenstaande wijze tot stand is gekomen, is vervolgens vertaald naar een eerste versie van de vragenlijst. Deze vragenlijst is opnieuw voorgelegd aan experts om de inhouds- en begripsvaliditeit te controleren.

De als eerste geïnterviewde expert was een tweede wetenschappelijk onderzoeker. Deze onderzoeker is scheikundige, en even als de eerste wetenschappelijk onderzoeker, vakdidactisch onderzoeker op de ULO afdeling van de VU en gespecialiseerd in het thema Leren Onderzoeken in het voortgezet onderwijs.

In dit interview is met name aandacht besteed aan het perceptiedeel van de inhoudsvaliditeit. Om inhoudsvalide te zijn, dient de vragenlijst de onderwerpen namelijk op zo'n manier te bevragen dat er mee gemeten wordt wat men wil nagaan: perceptie. In het interview met de wetenschappelijk onderzoeker is daarom vooral aandacht besteed aan de manier waarop de vragen het beste gesteld konden worden om zo goed mogelijk perceptie te meten. De vragenlijst is op basis van dit interview aangepast en vervolgens opnieuw voorgelegd aan de wetenschappelijk onderzoeker, om na te gaan of zij zich kon vinden in de aanpassingen en aanvullingen.

De nieuwe versie van de vragenlijst is vervolgens in twee interviews voorgelegd aan dezelfde docenten die meegeholpen hebben bij de validering van de competentiematrix. Hierbij werd opnieuw bekeken of de vragenlijst inhouds- en begripsvalide was. Ook werd gecontroleerd of de vragen op een voor de leerlingen begrijpelijke manier waren geformuleerd. De vragenlijst is op basis van de interviews wederom aangevuld en aangepast en vervolgens opnieuw voorgelegd aan de respondenten om na te gaan of zij zich konden vinden in de aanpassingen en aanvullingen.

Tot slot is de vragenlijst voorgelegd aan de begeleiders van dit Bacheloronderzoek. Hierbij is vooral aandacht besteed aan de begripsvaliditeit, door te controleren of de verschillende onderdelen van een competentie ongeveer evenveel aan bod kwamen. Ook is bekeken of de vragen eenduidig geformuleerd waren. Op basis hiervan is de vragenlijst opnieuw op een aantal punten aangepast.

Aanpak

De aanpak van de interviews met de docenten en de wetenschappelijk onderzoeker over de vragenlijst is hetzelfde als bij de interviews over de competentiematrix. Enkel de onderwerpen op de interviewagenda⁵ verschilden.

De interviews zijn gestart met een korte beschrijving van de stand van zaken in het onderzoek. Vervolgens is gevraagd of men het eens was met de uiteindelijke versie van de competentiematrix die zij toegestuurd hadden gekregen. Daarna werd ingegaan op de vragenlijst. Hierbij werden achtereenvolgens eerst de inleiding/uitleg, de algemene vragen, de indeling van de vragenlijst, en de gekozen standaard formulering voor de items besproken. Vervolgens werden de vragen per competentie doorlopen. Tot slot is de docenten gevraagd op welke manier de vragen op aantrekkelijkere wijze aan leerlingen zouden kunnen worden gepresenteerd en of zij nog tips hadden voor het afnemen van de vragenlijsten.

2.3 Afname vragenlijst

Voordat de vragenlijst werd voorgelegd aan een groot aantal leerlingen is een proefmeting uitgevoerd. Het doel hiervan was te achterhalen of er nog onduidelijkheden in de vragenlijst zaten. De tot stand gekomen vragenlijst is hiertoe via één van de geïnterviewde docenten voorgelegd aan één Havo4-leerling van zijn school. Haar werd gevraagd het telkens aan te geven als ze een vraag niet begreep, of als ze een fout in de vragenlijst tegenkwam. De vragenlijst is op basis hiervan aangepast.

Vervolgens is de vragenlijst voorgelegd aan een grotere groep leerlingen om de criteriumvaliditeit en de discriminantvaliditeit te kunnen meten.

Criteriumvaliditeit

De criteriumvaliditeit van de vragenlijst is tot nu toe buiten beschouwing gelaten. Deze vorm van validiteit kon namelijk niet in de interviews worden getoetst, omdat er geen soortgelijke meetinstrumenten bekend zijn waarmee de vragenlijst vergeleken zou kunnen worden. De criteriumvaliditeit kan echter wel op een statistische wijze worden bestudeerd.

Dit is gebeurd door de interne consistentie nader te bekijken. Hiermee wordt bedoeld dat de correlatie tussen de verschillende items binnen elke competentie is berekend, door middel van Cronbach's Alpha. Op basis hiervan kon worden vastgesteld welke items hoog met elkaar correleerden en welke niet. Bij sterke correlaties tussen items kan aangenomen worden dat er een sterke link is tussen de items en de latente variabele (DeVellis, 2003). De latente variabele is in dit geval de overkoepelende competentie. Als uit de berekeningen bleek dat bepaalde items onder een competentie heel laag correleerden met andere items onder diezelfde competentie, dienden deze items nader bestudeerd en vervolgens aangepast of verwijderd te worden.

Discriminantvaliditeit

Daarnaast is ook de discriminantvaliditeit bestudeerd. Dit is gedaan door verschillende groepen leerlingen binnen de steekproef waartussen verschillen verwacht kunnen worden met elkaar te vergelijken. Zo is er een vergelijking gemaakt tussen havo4- en vwo5-leerlingen en tussen leerlingen die aangeven wel en leerlingen die aangeven geen ervaring met open onderzoek te hebben.

De eerste vergelijking is gekozen, omdat er verwacht kan worden dat de havo4-leerlingen gemiddeld lager scores op de vragenlijst dan de vwo5-leerlingen, doordat zij een jaar korter ervaring hebben met onderzoek doen en doordat zij de lesstof aangeboden krijgen op havo-niveau, in plaats van vwo-niveau.

Voor de tweede vergelijking is gekozen, omdat verwacht kan worden dat leerlingen die ervaring hebben met het uitvoeren van open onderzoek, gemiddeld hoger scores op de vragenlijst dan leerlingen die aangeven daar geen ervaring mee te hebben.

Respondenten

Om de statistische analyse mogelijk te maken is de vragenlijst ingevuld door zes bestaande klassen op de drie scholen van de docenten in het expertteam.

Op elke school is daarbij gekozen voor een Havo 4 klas en een Vwo 5 klas, omdat deze leerlingen aan de start van het profielwerkstuk staan. Dat is de doelgroep die in een vervolgonderzoek

⁵ De interviewagenda's voor de interviews over de vragenlijst zijn op te vragen bij de auteur.

van ELAN de vragenlijst zal invullen om te onderzoeken welke onderzoekscompetenties nog niet voldoende aan bod zijn gekomen in het huidige curriculum.

In totaal hebben 133 leerlingen meegewerkt aan het onderzoek. De verdeling van leerlingen over de verschillende klassen is weergegeven in Tabel 1.

Tabel 1 Verdeling leerlingen over klassen naar geslacht

Geslacht	Havo 4	Vwo 5	Totaal
Jongens	48	29	77
Meisjes	31	25	56
Totaal	79	54	133

Instrument

De vragenlijst die aan de leerlingen is voorgelegd bestaat uit 98 items. Deze items zijn verdeeld over verschillende categorieën. Binnen de items kan ook onderscheid worden gemaakt tussen items die vragen naar de verschillende aspecten van een competentie, namelijk cognitieve, procedurele, metacognitieve en communicatieve aspecten. Deze verdeling wordt nader gespecificeerd in Tabel 2. Daarbij worden de procedurele en de communicatieve aspecten samengenomen, omdat deze beiden betrekking hebben op vaardigheden.

Tabel 2 Verdeling items over categorieën

Categorie	Aantal items			
	Cognitief	Procedureel en Communicatief	Meta-cognitief	Totaal
A. Algemene stellingen	2	0	7	9
Onderzoeksfase 1: De voorbereiding				
B. Onderzoeksvraag bepalen	2	5	4	11
C. Hypothese bepalen	1	2	2	5
D. Onderzoeksopzet bepalen	7	4	5	16
Onderzoeksfase 2: De uitvoering				
E. Observaties, metingen en/of simulaties uitvoeren	2	3	4	9
F. Waarnemingen registreren	1	3	1	5
Onderzoeksfase 3: De verwerking				
G. Gegevens verwerken	3	4	4	11
Onderzoeksfase 4: De conclusies				
H. Conclusies trekken	2	4	4	10
I. Verklaren	3	2	4	9
Onderzoeksfase 5: De evaluatie				
J. Evalueren onderzoeksproces	0	2	1	3
K. Vergelijken met ander onderzoek/theorie	0	2	3	5
L. Evalueren belang eigen onderzoek	0	1	1	2
M. Voorstellen doen voor verder onderzoek	0	1	2	3

De items werden aangeboden als stellingen (bijv. "Ik kan een hypothese opstellen."). Op elk item dienden leerlingen zichzelf een score te geven. Ze konden daarbij kiezen uit 'helemaal niet mee eens' (1), 'niet mee eens' (2), 'niet mee eens, maar ook niet mee oneens' (3), 'eens' (4) en 'helemaal mee eens' (5). Bij de meeste stellingen geeft een één telkens aan dat een leerling de competentie niet of nauwelijks denkt te bezitten en een vijf dat een leerling de competentie juist wel denkt te bezitten. Bij een aantal stellingen diende de scores om te worden geschaald, doordat een laag cijfer juist aangeeft dat een leerling de competentie denkt te bezitten (bijvoorbeeld 'Ik vind het saai om mijn onderzoeksvraag te verkennen'). Een deel van de vragenlijst is te vinden in Figuur 2.

Figuur 2 Vragenlijst competentie 'Onderzoeksvraag bepalen'

	Helemaal niet mee eens	Niet mee eens	Niet mee eens, maar ook niet mee oneens	Mee eens	Helemaal mee eens
Observaties, metingen en/of simulaties uitvoeren					
Ik vind het belangrijk dat ik nauwkeurig meet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik vind het belangrijk dat ik op de juiste manier meet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik weet hoe mijn meetmethode invloed kan hebben op mijn resultaten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik kan de invloed van mijn meetmethode op mijn resultaten berekenen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik kan de apparaten en het materiaal veilig hanteren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik vind het belangrijk dat ik apparaten en materiaal veilig gebruik.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik weet wat van belang is om de juiste metingen uit te voeren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ik vind het uitvoeren van metingen leuk.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Procedure

De afname van de vragenlijst heeft in alle klassen op dezelfde wijze plaatsgevonden, met uitzondering van het uit elkaar schuiven van tafels op één van de scholen.

Er is telkens gestart met een korte mondelinge uitleg van het onderzoek. Hierin werd hetzelfde verteld als wat in de inleiding van de vragenlijst staat, namelijk het doel van de vragenlijst, dat leerlingen anoniem blijven en dat het belangrijk is dat ze de vragenlijst eerlijk invullen. Dit werd zowel mondeling als schriftelijk vermeld om er zo zeker mogelijk van te zijn dat dit duidelijk was voor leerlingen.

Verder kregen de leerlingen vooraf de instructie om de vragenlijst individueel in te vullen. Hiertoe werden in vier van de zes klassen vooraf de tafels uit elkaar geschoven. Op één school was dit helaas niet mogelijk, doordat de tafels aan elkaar vast zaten.

Hierna werden de vragenlijsten uitgedeeld en in stilte door de leerlingen ingevuld. Telkens als leerlingen ongeveer op de helft van de vragenlijst waren werden kleine marsen of twixen uitgedeeld als blijk van dank naar de leerling toe. De totale tijd die leerlingen nodig hadden voor het invullen van de vragenlijst liep uiteen van 15 tot 30 minuten.

Als leerlingen een begrip in een stelling niet begrepen tijdens het invullen mochten ze hun hand op steken en werd door de docent of de onderzoeker een toelichting gegeven.

Pas als iedereen klaar was werden de vragenlijsten weer ingenomen. Vervolgens is de leerlingen mondeling gevraagd naar hun mening over de vragenlijst.

3. Resultaten

De resultaten worden op dezelfde manier besproken als de gehanteerde methoden, namelijk in drie delen. Eerst wordt besproken hoe inhouds- en de begripsvaliditeit van de competentiematrix is gecontroleerd. Vervolgens wordt dit voor de vragenlijst besproken. Tot slot wordt besproken hoe de criteriumvaliditeit en de discriminantvaliditeit van de vragenlijst is gecontroleerd.

3.1 Competentiematrix

De basis van de competentiematrix zijn de onderzoekscompetenties die gevonden werden bij een literatuurstudie. Bij de evaluatie van deze competentiematrix bleek dat de matrix op een aantal punten aangepast en aangevuld diende te worden om de inhouds- en de begripsvaliditeit van de vragenlijst te kunnen waarborgen.

De aanpassingen betroffen zowel volgens de docenten als volgens de wetenschappelijk onderzoeker de samenvoeging van een aantal competenties en herformuleringen. De competenties ‘doel bepalen’ en ‘probleem bepalen’ zijn bijvoorbeeld samengevoegd, omdat de wetenschappelijk onderzoeker aangaf dat je als je het probleem hebt bepaald ook meteen het doel weet. Dat is dan namelijk je probleem oplossen of je vraag beantwoorden. Een voorbeeld van een herformulering is dat ‘onderzoeksvraag formuleren’ is veranderd in ‘onderzoeksvraag’ bedenken, omdat de competentie op die manier een concretere handeling aanduidt, aldus de wetenschappelijk onderzoeker.

De aanvullingen betroffen met name de aspecten binnen competenties. Er is namelijk getracht om bij iedere competentie cognitieve, procedurele, metacognitieve en communicatieve aspecten te formuleren om de begripsvaliditeit te kunnen waarborgen. Een voorbeeld hiervan is de toevoeging van het cognitieve aspect ‘kennis van een spreadsheetprogramma zoals excel’ bij de competentie ‘gegevens verwerken’.

De verwerking van alle aanpassingen en aanvullingen van de experts heeft geleid tot een uiteindelijke competentiematrix waarop de vragenlijst gebaseerd kon worden⁶. Om een beeld te krijgen van de competentiematrix wordt een deel weergegeven in Figuur 1.

3.2 Ontwikkeling vragenlijst

Nadat de aanpassingen en aanvullingen die boven tafel zijn gekomen in de interviews waren doorgevoerd in de competentiematrix, is deze vertaald naar een eerste versie van de vragenlijst.

Wetenschappelijk onderzoeker

De eerste versie van de vragenlijst is vervolgens in een interview besproken met een tweede wetenschappelijk onderzoeker, om te achterhalen hoe het beste perceptie gemeten kan worden. De volgende punten kwamen naar voren.

- De vragen konden beter geformuleerd worden als stellingen (bijvoorbeeld ‘Ik kan een onderzoekbaar onderzoeksprobleem beschrijven’ in plaats van ‘Weet je wanneer een probleem onderzoekbaar is?’).
- Leerlingen kon het beste gevraagd worden zichzelf te scoren op een vijfpuntantwoordschaal (en niet op een tienpuntsschaal). Hoewel leerlingen gewend zijn aan een tienpuntsschaal, kan zo’n schaal volgens de wetenschappelijk onderzoeker namelijk overkomen alsof ze zichzelf moeten beoordelen. Het kan zijn dat leerlingen de vragen daardoor niet eerlijk beantwoorden. Een leerling zal zich misschien wel een één geven op de vijfpuntsschaal, maar bij dezelfde vraag geen één of twee kiezen op een tienpuntsschaal.
- Verder zou voorafgaand aan de vragenlijst een aantal algemene vragen kunnen worden gesteld die handig zijn bij het verwerken van de data (bijvoorbeeld ‘Heb je ervaring met open onderzoek?’ of ‘Welke natuurwetenschappelijke vakken volg je?’).
- Er is overwogen om vragen over een concrete situatie te stellen, om het voor leerlingen minder abstract te maken. Tijdens het interview is daarom getracht voorbeeldsituaties te bedenken. Dit bleek echter lastig, omdat het doel van dit onderzoek niet is te achterhalen of leerlingen denken te weten wat ze in één specifieke situatie moeten doen, maar juist of ze denken dat ze in het algemeen weten hoe te handelen.

De vragenlijst is in dit interview niet vraag voor vraag besproken, omdat de vragen rechtstreeks afkomstig waren uit de competentiematrix en nog opnieuw geformuleerd zouden worden als stellingen. Hierdoor is er weinig aandacht besteed aan het deel van de inhoudsvaliditeit dat er om gaat dat alle belangrijke onderwerpen bevraagd worden. De wetenschappelijk onderzoeker gaf wel aan zich te kunnen vinden in de competenties en aspecten die in de matrix waren geformuleerd.

Docenten

Nadat deze aanpassingen waren doorgevoerd is de nieuwe versie van de vragenlijst voorgelegd aan de drie docenten. Ook hieruit kwam een aantal aanpassingen naar voren.

- Antwoordcategorieën konden beter op elke pagina herhaald worden, in plaats van één keer in de inleiding uitgelegd te worden. Op die manier werd de inleiding ook wat korter.

⁶ De gehele competentiematrix is op te vragen bij de auteur.

- Algemene vragen (bijvoorbeeld ‘In welke klas zit je?’ en ‘Hoe oud ben je?’) zouden halverwege de inleiding geplaatst kunnen worden. Hierdoor ontstaan twee korte stukken tekst, namelijk voor de vragen en er na. Dit verhoogt volgens de docenten de kans dat leerlingen het daadwerkelijk goed lezen.
- Een aantal woorden in de vragenlijst was waarschijnlijk te moeilijk voor leerlingen. Sommige konden vervangen worden door een simpelere variant. Een aantal kon echter niet omzeild worden. Hierbij zou, aldus de docenten, een korte uitleg op zijn plaats zijn. De docenten raden aan om die uitleg dan zo dicht mogelijk bij het ‘moeilijke woord’ te plaatsen, omdat volgens hen de kans dat leerlingen de uitleg lezen dan het grootst is. Om die reden zijn ‘tips’ toegevoegd aan de vragenlijst. Een voorbeeld hiervan is de stelling ‘Ik weet wanneer het nodig is een blanco proef* op te nemen in mijn onderzoekszopzet en –aanpak’. De asterisk verwijst naar een tip die vlak onder de stelling wordt gegeven: ‘*Tip: Bij een blanco proef voeg je een extra monster aan je opstelling toe, die de te testen substantie niet bevat’.

In de interviews zijn de stellingen in de vragenlijst stuk voor stuk behandeld. Dit leidde tot twee soorten herformuleringen. Allereerst zijn overbodige bijzinnen verwijderd. Daarnaast zijn moeilijke woorden vervangen door een simpelere variant.

Er is in de interviews ook bekeken of er nog nieuwe stellingen aan de vragenlijst toegevoegd dienden te worden om de inhouds- en/of begripsvaliditeit te verhogen. Dit was volgens de docenten echter niet het geval.

Begeleiders Bacheloropdracht

De vragenlijst is aangepast op basis van de interviews en vervolgens besproken met de begeleiders van de Bacheloropdracht. Op basis van dit gesprek zijn nog enkele aanpassingen doorgevoerd.

Allereerst zijn vragen samengevoegd, als ze te veel op elkaar leken. Daarnaast is de standaardformulering ‘Ik streef er naar dat...’ veranderd in ‘Ik vind het belangrijk dat...’, omdat dit een betere weergave geeft van de metacognitieve aspecten van competenties.

Verder zijn vragen met een dubbele betekenis aangepast of opgesplitst in meerdere vragen, om de betrouwbaarheid van de vragenlijst te waarborgen.

Ook is een aantal vragen toegevoegd om er voor te zorgen dat er voldoende vragen in de vragenlijst zitten die betrekking hebben op de houdingen van leerlingen. Het doel hiervan was de begripsvaliditeit zo goed mogelijk te waarborgen.

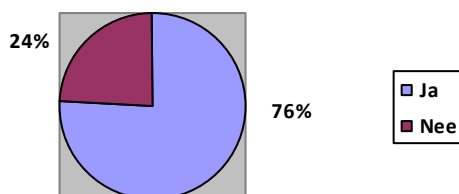
Tot slot zijn de ‘algemene vragen’ aan het einde van de vragenlijst verwijderd. Op al deze vragen werd ergens anders in de vragenlijst namelijk al dieper ingegaan. Alleen de stelling ‘Ik streef er naar systematisch te werken’ is behouden. Deze stelling kon geplaatst worden onder de competentie ‘Onderzoekszopzet bepalen’.

3.3 Afname vragenlijst

De vragenlijst is afgenomen bij 133 leerlingen. De belangrijkste resultaten die uit de analyse blijken zijn dat 6 van de 12 subschalen in de vragenlijst criteriumvalide zijn en dat de vragenlijst geen significant verschil tussen havo4-leerlingen en vwo5-leerlingen, en tussen leerlingen met en zonder ervaring met open onderzoek aangeeft ($p = 0,13$ resp. $p = 0,632$).

Behalve vragen die gerelateerd zijn aan onderzoekscompetenties, zijn er ook algemene vragen gesteld, die relevant zouden kunnen zijn bij het trekken van conclusies. Zo is gevraagd of de leerlingen ervaring hebben met het uitvoeren van open onderzoeken. De verdeling van het antwoord op die vraag is te vinden in Figuur 3.

Figuur 3 *Ervaring met open onderzoek*



Daarnaast zijn enkele algemene stellingen aan de leerlingen voorgelegd over de vakken natuurkunde, scheikunde en biologie en over onderzoek, namelijk:

- Ik vind natuurkunde boeiend.
- Ik vind scheikunde boeiend.
- Ik vind biologie boeiend.
- Ik vind natuurkunde moeilijk.
- Ik vind scheikunde moeilijk.
- Ik vind biologie moeilijk.
- Ik vind het interessant om een natuurwetenschappelijk onderzoek uit te voeren (bijvoorbeeld als profielwerkstuk).
- Ik weet wat beschrijvend onderzoek is.
- Ik weet wat hypothesetoetsend onderzoek is.

Deze vragen zijn niet relevant voor het controleren van de validiteit en betrouwbaarheid van de vragenlijst over perceptie van VO-leerlingen. De vragen kunnen bij later gebruik van de vragenlijst echter wel relevant zijn als conclusies uit de gegevens moeten worden getrokken. Daarom is het toch relevant om ook bij deze vragen de validiteit en betrouwbaarheid te bestuderen.

Criterionvaliditeit

Het eerste doel van de afname van de vragenlijsten was het mogelijk maken om de interne consistentie van de vragenlijst te bestuderen. Hiertoe is per competentie die in de vragenlijst bevraagd wordt Cronbach's Alpha berekend. Cronbach's Alpha is een indicatie van de betrouwbaarheid en geeft de samenhang tussen items weer. Als alle items binnen een competentie een sterke samenhang vertonen, kan aangenomen worden dat er een sterk verband is tussen de items en de overkoepelende competentie. Dit duidt er dan weer op dat de vragenlijst criteriumvalide is. De waarde van Cronbach's Alpha wordt per competentie weergegeven in Tabel 3. Er wordt daarbij ook een waarde weergegeven voor 'competentie A'. Dat onderdeel uit de vragenlijst bestond uit bovengenoemde algemene stellingen.

Tabel 3 Cronbach's Alpha per competentie

Competentie	Aantal items	Cronbach's Alpha
A. Algemene stellingen	9	0,66
B. Onderzoeksvraag bepalen	11	0,69
C. Hypothese bepalen	5	0,60
D. Onderzoekopzet bepalen	16	0,69
E. Observaties, metingen en/of simulaties uitvoeren	9	0,67
F. Waarnemingen registreren	5	0,61
G. Gegevens verwerken	11	0,63
H. Conclusies trekken	10	0,67
I. Verklaren	9	0,71
J. Evalueren onderzoeksproces	3	0,35
K. Vergelijken met ander onderzoek/theorie	5	0,52
L. Evalueren belang eigen onderzoek	2	0,60
M. Voorstellen doen voor verder onderzoek	3	0,77
Competentie B t/m M	89	0,94

DeVellis (2003, pp. 95-96) hanteert de volgende normen voor de waarden van Cronbach's Alpha:

- < 0.60: unacceptable
- 0.60 – 0.65: undesirable
- 0.65 – 0.70: minimally acceptable
- 0.70 – 0.80: respectable
- 0.80 – 0.90: very good
- > 0.90: one should consider shortening the scale.

Hij geeft daarbij aan dat het doel van het meetinstrument belangrijk is bij het kiezen van de norm. Zo is bij individuele assessments een veel hogere Cronbach's Alpha nodig dan bij onderzoek van data over groepen.

In Tabel 3 is te zien dat de competenties J en K een onacceptabele waarde van Cronbach's Alpha hebben (0,35 respectievelijk 0,52). Uit een analyse blijkt dat dit niet op te lossen is door items bij deze competenties te verwijderen. In enkele gevallen levert dit wel een hogere Cronbach's Alpha op, maar nooit een Cronbach's Alpha die groter of gelijk is aan 0,60.

Als per competentie een minimale interne betrouwbaarheid van 0,65 wordt vereist is in Tabel 3 te zien dat ook de competenties C, F, G en L een te lage Cronbach's Alpha hebben (0,60 – 0,65). Uit een analyse blijkt dat dit bij competentie F op te lossen is door een item te verwijderen. Als item 51 uit de vragenlijst wordt verwijderd, krijgt Cronbach's Alpha namelijk een waarde van 0,73. Bij de andere competenties is dit geen oplossing.

Ondanks dat het eerste onderdeel van de vragenlijst 'Algemene stellingen' een Cronbach's Alpha heeft van 0,66 is het nuttig deze vragen nader te bestuderen. De vragen in deze 'categorie' zijn namelijk eigenlijk niet in één categorie te noemen. De naam 'algemene stellingen' geeft dit al aan. Als naar de inhoud van de stellingen wordt gekeken, zou een driedeling in de vragen verwacht kunnen worden, namelijk vragen over hoe men tegen de vakken natuurkunde, scheikunde en biologie aan kijkt, een vraag over hoe men tegen onderzoek aan kijkt en twee vragen over kennis van soorten onderzoek. Omdat op basis van de inhoud van de vragen al vermoed kan worden dat er verschillende factoren een rol spelen binnen de categorie 'Algemene stellingen' zou het interessant zijn om een factoranalyse toe te passen op deze vragen en als er verschillende factoren worden gevonden daar voor opnieuw Cronbach's Alpha te berekenen.

Discriminantvaliditeit

Het tweede doel van de afname van de vragenlijsten was het mogelijk maken om de discriminantvaliditeit te bestuderen. Als de gemiddelde scores van de havo4-leerlingen met die van de vwo5-leerlingen vergeleken worden door middel van een t-toets, blijkt dat vwo5-leerlingen iets hoger scoren op de vragenlijst dan havo4-leerlingen. Dit verschil is bij een significantieniveau $\alpha = 0,05$ echter niet significant ($p = 0,13$). Dit is te zien in Tabel 4 en Tabel 5.

Tabel 4 Gemiddelde scores havo4 vs. vwo5

Klas	N	Gemiddelde score	Standaarddeviatie
Gemiddelde score Havo 4	79	3,49	0,29
Gemiddelde score Vwo 5	54	3,58	0,38

Tabel 5 Uitkomst t-toets havo4 vs. vwo5

T-waarde	-1,527
Df	94,662
Overschrijdingskans	0,13

De gemiddelde scores van leerlingen die aangeven ervaring te hebben met open onderzoek zijn vergeleken met die van leerlingen die aangeven geen ervaring te hebben met open onderzoek. Dit is tevens gebeurd door middel van een t-toets. Hieruit blijkt dat leerlingen die aangeven wel ervaring te hebben gemiddeld iets hoger scoren op de vragenlijst dan leerlingen die aangeven geen ervaring te hebben met open onderzoek. Dit verschil is echter ook niet significant ($p = 0,632$), zoals te zien is in Tabel 6 en Tabel 7.

Tabel 6 Gemiddelde scores ervaren vs. niet ervaren

Open onderzoekservaring	N	Gemiddelde score	Standaarddeviatie
Ja	101	3,53	0,33
Nee	32	3,50	0,33

Tabel 7 Uitkomst t-toets ervaren vs. niet ervaren

T-waarde	0,493
Df	131
Overschrijdingskans	0,623

Er is geen significant verschil gevonden tussen havo4- en vwo5-leerlingen en tussen leerlingen met en zonder ervaring met open onderzoek in de gemiddelde scores op de vragenlijst in zijn geheel. Daarom wordt niet meer specifieker per onderzoekscompetentie gekeken naar de verschillen tussen deze groepen.

4. Conclusie

De vraag die in dit onderzoek centraal staat is ‘Wat is een valide meetinstrument voor de perceptie van VO-leerlingen op hun onderzoekscompetenties?’. Aan het begin van dit artikel is aangegeven dat bij het bepalen van de validiteit gekeken zou worden naar inhouds-, begrips-, criterium- en discriminantvaliditeit. Daarbij wordt betrouwbaarheid gezien als een voorwaarde voor validiteit. Hieronder worden de typen validiteit in relatie tot de ontwikkelde vragenlijst besproken.

4.1 Inhoudsvaliditeit

Voor de inhoudsvaliditeit is het belangrijk dat de vragenlijst representatief is voor datgene wat men er mee wil nagaan, namelijk voor de perceptie van VO-leerlingen op hun onderzoekscompetenties.

Het eerste aspect dat van belang is voor de inhoudsvaliditeit is dat het meetinstrument alle belangrijke onderwerpen bevraagt. Er kan aangenomen worden dat de inhoudsvaliditeit wat betreft dit aspect goed is. De vragenlijst bevraagt namelijk alle onderzoekscompetenties die van belang zijn, doordat de vragenlijst gebaseerd is op een competentiematrix. Die competentiematrix is gebaseerd op een literatuuronderzoek. De matrix is vervolgens in interviews geëvalueerd met docenten en een wetenschappelijk onderzoeker met veel ervaring op het gebied van leren onderzoeken in het voortgezet onderwijs. Op basis van deze interviews is de matrix waar nodig aangepast en aangevuld.

Er zou aan dit aspect van de inhoudsvaliditeit getwijfeld kunnen worden doordat de geïnterviewde wetenschappelijk onderzoeker bioloog is. Men zou kunnen stellen dat er daardoor onvoldoende evaluatie heeft plaatsgevonden vanuit het perspectief van de vakken natuurkunde en scheikunde. Dit is echter gecompenseerd doordat de docenten wel de drie natuurwetenschappelijke vakken natuurkunde, scheikunde en biologie vertegenwoordigen. Daarnaast heeft ook een scheikundig wetenschappelijk onderzoeker de competentiematrix bekeken voorafgaand aan het interview over de vragenlijst en aangegeven dat zij zich kan vinden in de inhoud van de competentiematrix.

Het tweede aspect dat van belang is voor de inhoudsvaliditeit is dat er perceptie van leerlingen wordt gemeten. Op dit punt is de vragenlijst inhoudsvalide, doordat er voor gekozen is *leerlingen* een vragenlijst te laten invullen. Hierop dienen ze aan te geven in hoeverre ze *denken* elke onderzoekscompetentie te bezitten. Door de vragenlijst anoniem te laten invullen worden de leerlingen gestimuleerd de vragenlijst eerlijk in te vullen. De manier waarop het beste perceptie kan worden gemeten is ook besproken met een wetenschappelijk onderzoeker.

4.2 Begripsvaliditeit

Voor de begripsvaliditeit van de vragenlijst is het van belang dat de resultaten een maat zijn voor het meten van de perceptie op *competenties*, en bijvoorbeeld niet enkel voor het meten van kennis. Op basis van het onderzoek kan aangenomen worden dat de begripsvaliditeit van de vragenlijst voldoende is.

De begripsvaliditeit van het meetinstrument wordt gewaarborgd doordat de vragenlijst gebaseerd is op een competentiematrix. Hierdoor is er voor gezorgd dat de vragenlijst bestaat uit vragen over de perceptie op het bezitten van competenties en niet bijvoorbeeld alleen uit vragen over de perceptie op het bezitten van kennis.

Dit kan gesteld worden doordat er in de competentiematrix onderscheid gemaakt wordt tussen verschillende aspecten van een competentie, namelijk cognitieve aspecten, procedurele aspecten, metacognitieve aspecten en communicatieve aspecten. Daarmee worden de drie onderdelen van een competentie kennis, vaardigheden en houdingen gedekt. Cognitieve aspecten gaan namelijk over de kennis die men dient te bezitten voor een bepaalde competentie, procedurele en communicatieve aspecten gaan over de vaardigheden die een leerling dient te bezitten voor een bepaalde competentie en metacognitieve aspecten gaan over de houdingen die leerlingen dienen te hebben om een bepaalde competentie te bezitten en over hoe zij hun kennis dienen te reguleren.

In de vragenlijst komen de cognitieve aspecten aan bod in de stellingen die beginnen met ‘Ik weet...’. De procedurele en de communicatieve aspecten komen aan bod in stellingen die beginnen met ‘Ik kan...’. De metacognitieve aspecten komen aan bod in de overige stellingen. De meeste daarvan beginnen met ‘Ik vind het belangrijk dat...’, maar een aantal begint anders, bijvoorbeeld met ‘Ik vind het interessant om...’ of ‘Ik zie er tegenop om...’.

De volledigheid van de verschillende soorten aspecten is gewaarborgd door zowel de competentiematrix als de vragenlijst te evalueren met experts. Op basis hiervan zijn de aspecten

aangepast en aangevuld. Zo is er naar aanleiding van de evaluatie van de vragenlijst met de begeleiders een aantal metacognitieve stellingen aan de vragenlijst toegevoegd, om er voor te zorgen dat alle onderdelen van een competentie ongeveer gelijk vertegenwoordigd zijn in de vragenlijst.

De volledigheid van de verschillende soorten aspecten kan echter nog wel verbeterd worden. Tijdens de laatste evaluatie van de vragenlijst met de begeleiders is namelijk over het hoofd gezien, dat door eerdere aanpassingen in de vragenlijst, in de evaluatiefase geen kennisaspecten meer bevroegd worden. Het is belangrijk voor de begripsvaliditeit dat er weer kennisaspecten in de vragenlijst worden opgenomen bij de vier competenties in die fase.

4.3 Criteriumvaliditeit

Voor de criteriumvaliditeit van de vragenlijst is het van belang dat scores op items die gaan over dezelfde competentie aan elkaar zijn gerelateerd. Bij sterke correlaties tussen items kan namelijk aangenomen worden dat er ook een sterke link is tussen de items en de overkoepelende competentie (Dooley, 2001).

Uit een statistische analyse van de ingevulde vragenlijsten blijkt dat bij de helft van de competenties die in de vragenlijst bevroegd worden de items binnen een competentie inderdaad aan elkaar zijn gerelateerd. Zo is de waarde van Cronbach's Alpha bij de competenties B, D, E, H, I, en M groter of gelijk aan 0,65. Dit is een acceptabele waarde om met de vragenlijst onderzoek naar de perceptie van een grote groep VO-leerlingen te doen. Bij de overige competenties dient nog een verbetering te worden gemaakt.

Bij competentie F kan de waarde van Cronbach's Alpha verhoogd worden door één item te verwijderen (van 0,61 naar 0,73). Dit is echter niet aan te raden, aangezien dan de enige metacognitieve stelling bij die competentie wordt verwijderd. De lage Cronbach's Alpha zou ook verklaard kunnen worden door het kleine aantal items dat hoort bij de competentie. De competentie voldoet met zijn vijf items, namelijk maar net aan het vaak gehanteerde minimum van vijf items. Het lijkt daarom verstandiger om Cronbach's Alpha te proberen te verhogen door items toe te voegen. Bij de competenties C, J, K en L kan voor dezelfde oplossing worden gekozen. Die competenties bestaan respectievelijk uit vijf, drie, vijf en twee items.

Bij competentie K kan echter ook een andere verklaring gegeven worden voor de lage waarde van Cronbach's Alpha. Als de items binnen deze competentie nader worden bestudeerd, blijkt het mogelijk om binnen deze competentie twee verschillende onderdelen te onderscheiden. Drie van de vijf items gaan namelijk over het daadwerkelijk vergelijken van de eigen resultaten met theorieën uit ander onderzoek. De overige twee items gaan over het omgaan met informatiebronnen. De stellingen worden weergegeven in Tabel 8.

Tabel 8 Stellingen bij competentie K: 'Vergelijken met ander onderzoek/theorie'

Onderwerp stelling	Stellingen bij Competentie K
Vergelijken van de eigen resultaten met theorieën uit ander onderzoek.	Ik kan bepalen met welke theorieën ik mijn resultaten moet vergelijken.
	Ik vind het belangrijk om na te gaan of mijn eigen resultaten overeenkomen met bekende theorieën.
	Ik ben benieuwd of mijn resultaten overeenkomen met bekende theorieën.
Omgaan met informatiebronnen.	Ik kan op de juiste manier naar bronnen verwijzen.
	Ik vind het belangrijk dat ik zorgvuldig verwijs naar bronnen.

Het is mogelijk dat leerlingen op de twee laatste items anders scoren dan op de eerste drie items en dat daardoor de waarde van Cronbach's Alpha daalt. De vragenlijst zou dan verbeterd kunnen worden door extra items aan deze competentie toe te voegen. Er zouden daarbij in ieder geval ook stellingen over cognitieve aspecten moeten worden toegevoegd, aangezien de cognitieve aspecten bij competentie K nu nog niet aan bod komen. Of de competentie inderdaad te verdelen is in 'subcompetenties' zou onderzocht kunnen worden door middel van een factoranalyse.

Voor competentie G dient ook een andere oplossing te worden gezocht dan het toevoegen van items. Bij deze competentie horen namelijk al elf items. Het zou bij deze competentie eveneens nuttig zijn om een factoranalyse uit te voeren, om te kijken of verschillende factoren binnen de competentie de lage waarde van Cronbach's Alpha kunnen verklaren.

De waarde van Cronbach's Alpha voor de gehele vragenlijst is 0,94. Dit duidt op een sterke samenhang tussen de items in de gehele vragenlijst. Dat betekent dat de vragenlijst als geheel een goede weergave is van de perceptie van VO-leerlingen op hun onderzoekscompetenties als geheel.

4.4 Discriminantvaliditeit

Om discriminantvalide te zijn is het voor de vragenlijst belangrijk dat hij verschillende resultaten geeft als verschillende resultaten verwacht kunnen worden. Daarom is door middel van een statistische analyse bekeken of er een significant verschil is tussen de gemiddelde score op de vragenlijst als groepen waartussen verschil verwacht kan worden met elkaar vergeleken worden. Zo is er een vergelijking gemaakt tussen havo4-leerlingen en vwo5-leerlingen en tussen leerlingen die wel aangeven ervaring te hebben met open onderzoek en leerlingen die aangeven hier geen ervaring mee te hebben.

Uit de analyse blijkt dat er geen significant verschil is tussen de gemiddelde score van havo4-leerlingen en de gemiddelde score van vwo5-leerlingen op de vragenlijst. De gemiddelde score van de vwo5-leerlingen is namelijk wel hoger, maar de overschrijdingskans is 0,13. Dit kan betekenen dat de perceptie van havo4-leerlingen niet verschilt van de perceptie van vwo5-leerlingen. Er kan ook geconcludeerd worden dat de vragenlijst niet discriminantvalide is, omdat deze geen significant verschil meet, terwijl dat er wel is. Welke van deze twee betekenissen correct is, is mijns inziens niet duidelijk. Het is mogelijk dat de perceptie van havo4- en vwo5-leerlingen inderdaad niet (significant) verschilt. De havo4- en vwo5-leerlingen staan namelijk beiden aan de start van het profielwerkstuk, het laatste onderdeel van de leerlijn onderzoeken. Bovendien wordt van beide groepen min of meer hetzelfde verwacht; voor havo en vwo gelden dezelfde eisen wat betreft onderzoeksvaardigheden volgens de eindexamenprogramma's.

Uit de analyse blijkt tevens dat er geen significant verschil is tussen de gemiddelde score van leerlingen met ervaring met open onderzoek en leerlingen zonder ervaring met open onderzoek. Leerlingen met ervaring scoren namelijk gemiddeld wel hoger, maar dit verschil is niet significant ($p = 0,632$).

Een mogelijke verklaring hiervoor is dat het verschil in onderzoekservaring tussen leerlingen die aangeven wel ervaring met open onderzoek te hebben, en leerlingen die aangeven geen ervaring met open onderzoek te hebben, te klein is. Als leerlingen gevraagd was of zij ervaring met onderzoek doen hebben, had iedereen waarschijnlijk 'ja' geantwoord, omdat dit in alle klassen die de vragenlijst hebben ingevuld in de les geoefend wordt. Ook nu er gevraagd werd of leerlingen ervaring hebben met *open* onderzoek werd al door een groot aantal leerlingen 'ja' geantwoord (76%). Eén leerling merkte na het invullen van de vragenlijst op dat ze wel ervaring heeft met onderzoek en dat ze meestal wel erg vrij worden gelaten in het uitvoeren van een onderzoek. Het onderwerp van het onderzoek wordt volgens de leerling echter vaak door de docent aangedragen. Daarom had de leerling 'nee' ingevuld op de vraag over open onderzoekservaring. Als het verschil tussen leerlingen met en zonder ervaring met open onderzoek doen enkel bestaat uit het wel of niet zelf mogen aandragen van het onderwerp, zou dit een verklaring kunnen zijn voor het niet significant zijn van het verschil tussen deze twee groepen leerlingen.

De vragenlijst zou op dit punt verbeterd kunnen worden door de vraag over open onderzoekservaring duidelijker te formuleren. Hierbij kan overwogen worden om enkel te vragen naar ervaring met onderzoek, en niet specifiek met *open* onderzoek.

Kortom: er kan op basis van deze analyses niet geconcludeerd worden dat de vragenlijst discriminantvalide is. Er wordt aangeraden de vragenlijst aan te passen op basis van de gevonden resultaten en deze opnieuw voor te leggen aan groepen VO-leerlingen waartussen met meer zekerheid een verschil verwacht kan worden, bijvoorbeeld Vwo4- en Vwo5-leerlingen of Havo3- en Havo4-leerlingen.

5. Discussie

Uit de resultaten van het onderzoek kan geconcludeerd worden dat het ontwikkelde meetinstrument voldoende inhouds- en begripsvalide is. Er zijn echter nog wel een aantal verbeterpunten. Zo dienen er bij de competenties J tot en met M stellingen over cognitieve aspecten toegevoegd te worden om de begripsvaliditeit te verhogen. Bij de competenties J, K en L is toevoeging van extra items ook noodzakelijk om de criteriumvaliditeit te kunnen waarborgen. Dit is ook het geval bij competentie C en F. Bij competentie G dient een factoranalyse te worden uitgevoerd, om te controleren of verschillende factoren binnen de competentie de lage waarde van Cronbach's Alpha kunnen veroorzaken. Tot slot dient de vraag over open onderzoekservaring opnieuw te worden geformuleerd om de discriminantvaliditeit beter te kunnen toetsen.

Als aan deze verbeterpunten is voldaan dienen de criterium- en de discriminantvaliditeit opnieuw getoetst te worden. In dit onderzoek is een steekproef van 133 VO-leerlingen gehanteerd. Het was niet mogelijk binnen het tijdsbestek de vragenlijst aan meer leerlingen voor te leggen. Deze steekproefgrootte is vermoedelijk voldoende om de criteriumvaliditeit van de vragenlijst te kunnen bestuderen. Om de discriminantvaliditeit na aanpassing van de vragenlijst opnieuw te kunnen toetsen, wordt echter aangeraden een grotere steekproef te nemen. Dit maakt statistische vergelijkingen tussen groepen binnen de steekproef en conclusies over verschillen tussen die groepen (zoals havo4- en vwo5-leerlingen) betrouwbaarder.

Er wordt aangeraden niet alleen een grotere steekproef te nemen, maar ook een gevarieerdere steekproef. De vragenlijst is nu namelijk afgenomen op scholen die deelnemen aan het 3xO-project. Op die scholen wordt al tamelijk veel aandacht besteed aan 'leren onderzoeken' in het curriculum. Door ook scholen die (nog) weinig aandacht besteden aan leren onderzoeken in het curriculum in de steekproef op te nemen kan de discriminantvaliditeit beter worden bestudeerd.

De steekproef zou ook gevarieerder kunnen worden gemaakt door zowel leerlingen met weinig onderzoekservaring (bijvoorbeeld Vwo4-leerlingen) als leerlingen die binnenkort aan het profielwerkstuk beginnen (bijvoorbeeld Vwo5-leerlingen aan het einde van het jaar) op te nemen. Een andere optie is leerlingen met een Economie & Maatschappijprofiel of Cultuur & Maatschappijprofiel te vergelijken met leerlingen met een Natuurprofiel.

Als na het aanpassen van de vragenlijst uit de resultaten van de nieuwe steekproef blijkt dat de vragenlijst ook voldoende criterium- en discriminantvalide is, kan de vragenlijst daarna grootschalig worden ingezet om de perceptie van VO-leerlingen op hun onderzoekscompetenties te meten.

Als bij een tweede toetsing van criteriumvaliditeit echter weer een minimale Cronbach's Alpha van 0,65 wordt gehanteerd kan het zijn dat de waarde van Cronbach's Alpha voor de meeste competenties – net als in dit onderzoek – weer tussen 0,60 en 0,70 uitkomt. In dat geval dient er rekening mee te worden gehouden dat de mate waarin de verschillende competenties dan criteriumvalide zijn, niet voldoende is om conclusies te trekken over de perceptie van een individuele VO-leerling. Een waarde van 0,60 tot 0,70 betekent immers dat er nog aardig wat ruis in de vragenlijst zit. Concreet betekent dit dat er op basis van één vragenlijst niets gezegd kan worden over de perceptie van de VO-leerling die de betreffende vragenlijst heeft ingevuld. Als de vragenlijst door een grote groep VO-leerlingen wordt ingevuld, kunnen wel met voldoende zekerheid conclusies getrokken worden over op welke competenties VO-leerlingen in het algemeen hoog scoren en op welke niet (cf. DeVellis, 2003, p.96).

Indien men de vragenlijst wel wil gaan gebruiken om de perceptie van individuele leerlingen op specifieke competenties te bestuderen, dient de betrouwbaarheid van competenties binnen de vragenlijst verhoogd te worden, bijvoorbeeld door extra items toe te voegen aan de competenties met een lage Cronbach's Alpha.

Het doel van dit onderzoek was een valide en betrouwbaar meetinstrument ontwikkelen dat de perceptie van VO-leerlingen op hun eigen onderzoekscompetenties meet. ELAN wil dat meetinstrument inzetten in combinatie met een meetinstrument dat de mate waarin VO-leerlingen onderzoekcompetenties daadwerkelijk bezitten meet. Een vergelijking van de resultaten van beide meetinstrumenten kan ELAN dan handvaten bieden bij het geven van een kwalitatief goede invulling aan de leerlijn onderzoeken. Om de resultaten van beide meetinstrumenten zo goed mogelijk te kunnen vergelijken wordt aangeraden om het meetinstrument dat de daadwerkelijke onderzoekscompetenties van VO-leerlingen meet, te baseren op de voor dit onderzoek ontwikkelde competentiematrix.

Op die manier kan na het inzetten van beide instrumenten bestudeerd worden welke competenties VO-leerlingen bezitten en denken te bezitten, en welke er nog uitgebreider aan bod dienen te komen in het curriculum. De resultaten van beide meetinstrumenten kunnen dan een leidraad vormen voor de invulling van de leerlijn onderzoeken.

In de ontwikkelde vragenlijst staat hypothesetoetsend onderzoek centraal. Als ELAN de vragenlijst ook wil gebruiken om de perceptie op onderzoeksvaardigheden bij andere typen onderzoek te meten, dient de vragenlijst te worden aangepast. Daarbij kunnen de gedefinieerde fasen in onderzoek hetzelfde blijven, maar de onderzoekscompetenties die daarbij horen en de aspecten bij die competenties veranderen. De meest voor de hand liggende verandering is dat de onderzoekscompetentie 'hypothese bepalen' een stuk minder belangrijk wordt, maar ook voor andere competenties dient opnieuw bekeken te worden of ze relevant en goed geformuleerd zijn. Wellicht vraagt een ander type onderzoek, naast aanpassingen ook toevoegingen van onderzoekscompetenties.

Referenties

- Berkel, H. van, & Bax, A. (2006). *Toetsen in het hoger onderwijs*. Bohn Stafleu van Loghum: Houten.
- Centrale Examencommissie Vaststelling Opgaven (2005a). *Concept Examenprogramma Biologie HAVO*. CEVO
- Centrale Examencommissie Vaststelling Opgaven (2005b). *Concept Examenprogramma Biologie VWO*. CEVO
- DeVellis, R.F. (2003). *Scale Development: theory and applications*. Sage Publications, Inc.: Thousand Oaks, California.
- Dooley, D. (2001). *Social Research Methods*. Prentice-Hall, Inc.: New Jersey.
- Hubers, S.T.T. (2003). *Individuele Leertheorieën en het leren onderzoeken in de Tweede Fase*. Proefschrift Technische Universiteit Eindhoven.
- Schalk, H.H. (2006). *Zeker Weten? Leren de kwaliteit van biologie-onderzoek te bewaken in 5 vwo*. Proefschrift Vrije Universiteit Amsterdam.
- Schalk, H.H., Van der Schee, J.A., & Boersma, K.Th. (2007). *The development of understanding of evidence in pre-university biology education in the Netherlands*. Paper presented at the 7th ESERA conference, August 21-25, 2007, Malmö, Sweden.
- Smith, P.L., & Ragan, T.J. (2005). *Instructional Design*. Hoboken, NJ: John Wiley & Sons, Inc.
- Stichting Leerplanontwikkeling. (2004, januari). *Onderzoeken en ontwerpen van PO naar HO in een doorlopende leerlijn*. Enschede: S.T.T. Hubers.
- Stokking, K.M., & Schaaf, M.F. van der (2000). *Ontwikkeling en beoordeling van onderzoeksvaardigheden*. ISOR Rapport 00.13. Utrecht: Universiteit Utrecht (Onderwijskunde/ISOR).
- Verloop, N., & Lowyck, J. (2003). *Onderwijskunde*. Groningen/Houten: Wolters-Noordhoff.
- White, B.Y., & Frederiksen, J.R. (1998). Inquiry, Modeling and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction*. 16, 3-118.