

# Ontwikkeling en validering van een instrument om bij medisch specialisten de transfer van geleerde didactische vaardigheden naar de werkplek in kaart te brengen

**Vanessa Werners**

UMCG, Wenckebach Instituut  
Rijksuniversiteit Groningen, Onderwijskunde

Groningen, november 2008

**Wenckebach Ontwikkelplatform**

Universitair Medisch Centrum Groningen





# Ontwikkeling en validering van een instrument om bij medisch specialisten de transfer van geleerde didactische vaardigheden naar de werkplek in kaart te brengen

'Evaluatieonderzoek naar het 'Teach the Teacher' programma'

Groningen, november 2008

Auteur  
Studentnummer

mw. V. Werners  
1552996

Afstudeerscriptie in het kader van

Masterthesis Onderwijskunde, Faculteit  
Gedrags- en Maatschappijwetenschappen, RUG

Opdrachtgever

dr. P.M. Boendermaker  
Wenckebach Instituut, UMCG

Begeleiders  
Tweede beoordelaar

dr. R.Maslowski  
dr. S. Doolaard  
Gedrags- en Maatschappijwetenschappen, RUG

Begeleider UMCG

prof.dr. R.P. Zwierstra, emeritus hoogleraar  
Medisch Onderwijs  
dr. P.M. Boendermaker, Senior stafmedewerker  
Wenckebachinstituut

ISBN 978-90-8827-032

NUR 841

Trefw ontwikkeling, validiteit, betrouwbaarheid, didactische vaardigheden medisch specialisten, meetinstrument, vragenlijst, medisch onderwijs, leertheorieën, leren van volwassenen, evaluatieonderzoek

Omslag: Wenckebach Instituut, Universitair Medisch Centrum Groningen

© 2008 Wenckebach Ontwikkelplatform Publicaties Groningen, Nederland.

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enige andere manier, zonder voorafgaande toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikel 16B Auteurswet 1912 j° het Besluit van 20 juni 1974, St.b. 351, zoals gewijzigd in Besluit van 23 augustus 1985, St.b. 471 en artikel 17 Auteurswet 1912, dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht. Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) dient men zich tot de uitgever te wenden.

## Voorwoord

*Elk kent zijn eigen deur,  
Toegang tot het bekende  
Wat wacht ons  
Achter een vreemde deur?*

S. Bolhuis, 2004

Met deze masterthesis komt een einde aan mijn studententijd in de bruisende stad Groningen. Ik kreeg met dit onderzoek een kans om een kijkje achter de schermen van het Universitair Medisch Centrum Groningen (UMCG) te nemen. Een kijkje achter een '*vreemde onbekende deur*'. In dit onderzoek deed ik een aanzet om een betrouwbaar en valide instrument te ontwikkelen voor het meten van het effect van het 'Teach the Teacher' programma module (1) daarmee werd een bijdrage geleverd aan het bedrijven van 'best evidence medical education'. Tijdens deze scriptieperiode leerde ik ontzettend veel over het medisch onderwijs, dankzij mijn enthousiaste begeleiders vanuit het UMCG. Daarnaast leerde ik ook veel over mijn eigen persoon, over mijn zwakke en sterke kanten. Het scriptieproces heeft hoge bergen en lage dalen gekend, ik ben dan ook heel trots dat ik nu het eindresultaat van deze scriptieperiode kan presenteren in deze masterthesis. Ik heb bereikt wat ik wilde bereiken.

Ten eerste bedank ik mijn begeleiders vanuit het UMCG. De heren prof.dr. R.P. Zwierstra en dr. P.M. Boendermaker bedank ik voor het vertrouwen in mijn kunnen, voor de bemoedigende woorden als ik het even niet meer zag zitten, voor de opbouwende feedback en de vragen die mij aanzetten tot kritisch denken. Jullie zijn

geweldige 'teachers'. De heer dr. R. Maslowski dank ik voor zijn begeleiding vanuit de Master Onderwijskunde. Ik dank dr. J. Pols voor de enthousiaste ontvangst in het UMCG, en het inzicht dat een onderzoek in het UMCG helemaal niet zo'n gek idee was. Mevrouw A. Muurman dank ik voor de praktische ondersteuning vanuit het Ontwikkelplatform. Daarnaast wil ik alle medisch specialisten en 'leerlingen' bedanken die als respondent hebben meegewerkt aan het onderzoek. Graag wil ik alle docenten van het 'Teach the Teacher' programma bedanken die ik ben tegengekomen tijdens de scriptieperiode en bij wie ik een cursus mocht bijwonen. Ik dank Mariska Leliveld en Anja Karg voor de praktische tips op onderzoeksgebied, tips over het medisch onderwijs en het 'meedenken' met mijn scriptie. Dank aan de heer dr. K.H. Groenier bij de methodologische ondersteuning en dé oplossing. Bedankt medestudenten van het Ontwikkelplatform voor de gezellige uurtjes bij het Stinfi! Dank aan Daan van Boggelen en Richard Werners voor de technische ondersteuning bij dit onderzoek. Bedankt Huize Spilsluize! Mijn lieve huisgenoten die zorgden voor aangename onderbrekingen tijdens het scriptieproces en de emotionele ondersteuning op moeilijke momenten. Bedankt oppasgezin voor de ontspannen tochtjes naar Haren! Dank aan mijn commissiegenootjes, vrienden en muziekvrienden die mij af en toe stevig toespraken, en mij achter de computer weghaalden omdat er echt nog meer in het leven is dan alleen 'de scriptie'. Tot slot bedank ik mijn ouders, mijn broertjes Richard en Martijn voor hun nimmer aflatende steun, die altijd in mij *zijn* en *blijven* geloven....

Vanessa Werners  
Groningen, november 2008



# INHOUDSOPGAVE

<b>SAMENVATTING .....</b>	<b>1</b>
<b>1 INLEIDING VAN HET ONDERZOEK .....</b>	<b>3</b>
1.1 INLEIDING .....	3
1.2 ONTWIKKELINGEN IN HET MEDISCH ONDERWIJS .....	3
1.3 ONTWIKKELINGEN MET BETREKKING TOT COMPETENTIEGERICHT ONDERWIJS .....	4
1.4 CONTEXT VAN DE INSTELLING WAAR HET ONDERZOEK WORDT UITGEVOERD .....	8
1.5 AANLEIDING EN PROBLEEMSTELLING VAN HET ONDERZOEK .....	8
1.6 DOELSTELLING EN OPERATIONALISATIE VAN DE ONDERZOEKSVRAAG .....	10
1.7 EFFECTSTUDIES NAAR SOORTGELIJKE TRAININGEN ALS 'TEACH THE TEACHER' .....	13
1.8 SAMENVATTING HOOFDSTUK 1 EN BESCHRIJVING VAN DE OPZET VAN DE MASTERTHESIS .....	16
<b>2 THEORETISCH KADER .....</b>	<b>17</b>
2.1 INLEIDING .....	17
2.2 INSTRUCTIONAL EVENTS VAN GAGNÉ .....	17
2.3 MODEL 'DIDACTISCHE ANALYSE' VAN VAN GELDER .....	19
2.4 SOCIAAL-COGNITIEVE LEERTHEORIE VAN BANDURA .....	20
2.5 HET CREËREN VAN EEN VEILIG LEERKLIMAAT IN HET MEDISCH ONDERWIJS .....	22
2.6 LEREN VAN VOLWASSENEN .....	22
2.7 SAMENVATTING HOOFDSTUK 2 .....	23
<b>3 ONDERZOEKSMETHODEN .....</b>	<b>31</b>
3.1 INLEIDING .....	31
3.2 ONDERZOEKSMETHODEN MET BETREKKING TOT VRAAGSTELLING 1 .....	31
3.3 ONDERZOEKSMETHODEN MET BETREKKING TOT VRAAGSTELLING 2 .....	32
3.3.1 Onderzoeksmethode .....	33
3.3.2 Validiteit .....	33
3.3.3 Betrouwbaarheid .....	34
3.4 SAMENVATTING HOOFDSTUK 3 .....	35
<b>4 ONDERZOEKSRISULTATEN .....</b>	<b>37</b>
4.1 INLEIDING .....	37
4.2 ONDERZOEKSRISULTATEN VRAAGSTELLING 1 .....	37
4.2.1 Analyse instrumenten .....	39
4.3 ONTWIKKELING VAN EEN BETROUWBAAR EN VALIDE INSTRUMENT .....	49
4.3.1 Inhoudsvaliditeit van het ontwikkelde instrumentarium .....	49

4.3.2	Constructvaliditeit van het ontwikkelde instrumentarium .....	51
4.4	SAMENVATTING HOOFDSTUK 4 .....	56
<b>5</b>	<b>CONCLUSIE EN DISCUSSIE.....</b>	<b>57</b>
5.1	INLEIDING.....	57
5.2	CONCLUSIES VRAAGSTELLING 1.....	57
5.3	CONCLUSIES VRAAGSTELLING 2.....	57
5.4	DISCUSSIE .....	58
5.4.1	Discussie deelvraag 1 en aanbevelingen voor verder onderzoek .....	59
5.4.2	Discussie deelvraag 2 en aanbevelingen voor verder onderzoek .....	60
5.4	SAMENVATTING HOOFDSTUK 5 .....	63
<b>LITERATUURLIJST .....</b>		<b>65</b>
<b>BIJLAGEN .....</b>		<b>70</b>



## Samenvatting

Binnen het UMCG zijn de afgelopen jaren trainingen ontwikkeld voor medisch specialisten die een opleidingstaak hebben, het zogenaamde 'Teach the Teacher' programma. Het is het doel van het 'Teach the Teacher' programma, naast de door de beroepsverenigingen vastgestelde inhoudelijke doelen, te zorgen voor voldoende competenties bij specialisten om hun opleidingstaak naar behoren te kunnen uitvoeren. Binnen het 'Teach the Teacher' programma zijn drie modules ontwikkeld: (1) training in didactische basisprincipes, (2) onderwijs tijdens de dagelijkse patiëntenzorg en (3) het begeleiden, toetsen en beoordelen van artsen in opleiding tot specialist (aios). Uit de cursusevaluaties blijkt dat de deelnemers zeer tevreden zijn over de cursussen. Onduidelijk is echter in welke mate specialisten het geleerde toepassen in hun dagelijkse (opleidings)praktijk. De onderzoeksvraag luidt: *Is het mogelijk om een valide en betrouwbaar instrument te ontwikkelen, gericht op het meten van de door de medisch specialist geleerde vaardigheden in module (1) van het 'Teach the Teacher' programma?* Deze vraagstelling werd in twee stappen beantwoord. Eerst werd gekeken of er instrumenten aanwezig waren om het effect van module (1) te bepalen en indien er geen instrumenten aanwezig waren of er wel schalen uit de gevonden instrumenten bruikbaar waren voor de ontwikkeling van het nieuwe instrument. Het tweede doel van deze studie beoogde te bepalen of het mogelijk was om een betrouwbaar en valide instrument te ontwikkelen om het effect van module (1) te meten.

Voor de theoretische fundering van dit onderzoek werd gebruik gemaakt van onderwijskundige- en psychologische theorieën. Met behulp van deze theorieën werd een conceptueel kader geschetst van kernpunten die aandacht

behoefden bij de ontwikkeling van het meetinstrument. Een eerste versie van het instrument (vragenlijst) werd ontworpen op basis van de resultaten van het literatuuronderzoek en observaties tijdens de training. De inhoudsvaliditeit werd vastgesteld door observaties tijdens de training en de beoordeling van de opgestelde items door twee cursusleiders van het 'Teach the Teacher' programma. Voor het bepalen van de constructvaliditeit werd een exploratieve factoranalyse uitgevoerd. Daarnaast werd nagegaan in hoeverre de verschillende schalen betrouwbaar gemeten waren. In totaal hebben twee rondes van dataverzameling plaatsgevonden om voldoende gegevens te verzamelen voor de statistische analyses. In eerste instantie werden gegevens verzameld onder aios van medisch specialisten die hadden deelgenomen aan module (1) van het 'Teach the Teacher' programma (n=6). In een later stadium werd besloten, vanwege de lage respons bij de eerste vorm van dataverzameling, om ook vragenlijsten te laten invullen door aios binnen het UMCG (n=38).

Geconstateerd kan worden dat er op basis van het conceptueel kader geen volledige instrumenten geschikt zijn die het effect van het 'Teach the Teacher' programma module (1) kunnen meten. Op basis van het conceptueel kader zijn wel een aantal items geselecteerd die bruikbaar waren bij de ontwikkeling van het nieuwe instrument. Uit de resultaten van de factoranalyse blijkt dat er acht factoren onderscheiden worden: *'creëren van een veilig leerklimaat', 'begeleiding van het ontwikkelingsproces van de leerling', 'organisatie van het onderwijsmoment', 'bewustzijn rolmodel', 'vertrouwen en integriteit', 'creëren van veiligheid tijdens een feedbacksessie', 'controle en herhaling', 'toetsing'*. Op basis van de correlatieanalyse kan gesteld worden dat alle subschalen positief met elkaar samenhangen. Het instrument bevat zeer betrouwbare

schalen daarmee is een betrouwbaar instrument ontwikkeld.

Er komen een aantal aandachtspunten naar voren in de discussie waar in verder onderzoek rekening mee dient te worden gehouden. Met betrekking tot de data die werd verzameld voor dit onderzoek onder aios, kan het volgende worden opgemerkt. In de analyse werd voor de missende waarden gebruik gemaakt van een schatting. Bij de missende waarden werd het gemiddelde op het item ingevuld. In de vooranalyses werden verschillende items verwijderd vanwege de vele missende waarden, in vervolgonderzoek dient te worden nagegaan waarom de items niet goed zijn ingevuld. De steekproef was klein. In dit onderzoek werden alleen de inhouds- en constructvaliditeit bepaald, in vervolgonderzoek zal ook de criteriumvaliditeit van het instrument moeten worden bepaald. Het is gelukt om een basisinstrument te ontwikkelen voor het meten van het effect van het 'Teach the Teacher' programma module (1). Er wordt expliciet gesproken over een basisinstrument omdat bijvoorbeeld nog niet aan alle eisen van validiteit is voldaan en geanalyseerd moet worden of dezelfde resultaten ook optreden bij een andere onderzoeksgroep.

# 1 Inleiding van het onderzoek

## 1.1 Inleiding

*"Ik zweer/beloof dat ik de geneeskunst zo goed als ik kan zal uitoefenen ten dienste van mijn medemens. Ik zal zorgen voor zieken, gezondheid bevorderen en lijden verlichten. Ik stel het belang van de patiënt voorop en eerbiedig zijn opvattingen. Ik zal aan de patiënt geen schade doen. Ik luister en zal hem goed inlichten. Ik zal geheim houden wat mij is toevertrouwd. Ik zal de geneeskundige kennis van mijzelf en anderen bevorderen. Ik erken de grenzen van mijn mogelijkheden. Ik zal mij open en toetsbaar opstellen, en ik ken mijn verantwoordelijkheid voor de samenleving. Ik zal de beschikbaarheid en toegankelijkheid van de gezondheidszorg bevorderen. Ik maak geen misbruik van mijn medische kennis, ook niet onder druk. Ik zal zo het beroep van arts in ere houden."*

*"Dat beloof ik."*

of

*"Zo waarlijk helpe mij God almachtig."*

Nederlandse artseneed KNMG en VSNU (2003)

Binnen het hoger onderwijs neemt de medische opleiding een bijzondere plaats in. Hippocrates besteedde omstreeks 400 jaar voor Christus al aandacht aan de opleiding van toekomstige artsen, deze tekst staat bekend als de eed van Hippocrates (KNMG, 2004). In de tekst van de huidige artseneed, die geëvolueerd is uit de eed van Hippocrates, wordt nog steeds aandacht besteed aan het overbrengen van de 'kunst' aan leerlingen (Metz, Scherpbier en Van der Vleuten, 1995). Met de verankering in de artseneed 'van het overbrengen van kennis aan anderen', wordt een basis gecreëerd voor de start van deze masterthesis. De arts dient anderen, artsen in opleiding tot specialist (aios),

artsen niet in opleiding tot specialist (anios) en co-assistenten, te doceren in de Geneeskunde en daarnaast dient de arts zichzelf voortdurend te blijven ontwikkelen, door middel van bij- en nascholing, om op deze manier nieuwe kennis te verwerven.

De laatste jaren wordt binnen de medische opleidingen steeds meer aandacht besteed aan de samenwerking tussen onderwijskundigen en artsen. Het voordeel van deze samenwerking is dat ontwikkelingen in de algemene Onderwijskunde goed kunnen worden vertaald naar de onderwijspraktijk van de medische wereld (Metz, Scherpbier en Van der Vleuten, 1995). Ook in deze thesis is sprake van samenwerking tussen de algemene Onderwijskunde en de medische wereld.

In dit hoofdstuk worden in de eerste paragrafen de laatste ontwikkelingen met betrekking tot het medisch onderwijs besproken zowel voor de basisopleiding als voor de medische vervolgoopleidingen. Paragraaf 1.4 schetst de context van de instelling waar het onderzoek wordt uitgevoerd, het UMCG. Paragraaf 1.5 beschrijft de aanleiding voor het onderzoek en de probleemstelling. Gevolgd door paragraaf 1.6 waar de doelstelling van het onderzoek wordt beschreven en de operationalisatie van de onderzoeksvraag plaats vindt. Tot slot wordt in paragraaf 1.7 een kort overzicht gegeven van effectstudies naar soortgelijke trainingen. In de afsluitende paragraaf 1.8 wordt vooruitgeblikt op de inhoud van deze masterthesis.

## 1.2 Ontwikkelingen in het medisch onderwijs

De laatste decennia is er veel veranderd in de structuur van de opleiding tot basisarts. "Na tien jaar tweefasenstructuur is met ingang van het studiejaar 1994-1995 de wet op het Hoger Onderwijs en het Wetenschappelijk Onderzoek (WHW) van kracht geworden" (Metz, Scherpbier en Van der Vleuten, 1995, p. 5-6). De aanhoudende wijzigingen in de wettelijke regelingen zorgen voor grote problemen bij de medische faculteiten. De wijzigingen die worden voorgeschreven ten aanzien van de opleidingsstructuur, in

praktijk komt dit neer op verkorting van de studieduur, zijn moeilijk overeen te stemmen met de ontwikkelingen die in de Geneeskunde zelf gaande zijn. Deze ontwikkelingen zouden eerder uitbreiding dan verkorting van de opleidingsduur veronderstellen (Metz, Scherpbier en Van der Vleuten, 1995). “De medische faculteiten bestuurlijk verenigd in het Discipline-overlegorgaan Medische Wetenschappen (DMW) van de Vereniging van Samenwerkende Nederlandse Universiteiten (VSNU), hebben dit voorstel niet overgenomen en zich voorhands op het standpunt gesteld, dat eerst getracht moet worden om met behulp van onderwijskundige maatregelen de opleiding zodanig te optimaliseren, dat – binnen de bestaande zesjarige cursusduur – de noodzakelijke verbeteringen in de wetenschappelijke vorming en in de beroepsvoorbereiding van de arts toch gerealiseerd kunnen worden” (Metz, Scherpbier en Van der Vleuten, 1995, p. 6). In de daarop volgende jaren werd tevens de numerus fixus voor de studie Geneeskunde verhoogd. De verhoging van het aantal studenten vraagt om een andere organisatie van onderwijs en dat in het licht van bovenstaande ontwikkelingen.

Met het teruglopen van de voor het medisch onderwijs beschikbare middelen en daarmee de toenemende belangstelling voor effectiviteit en efficiëntie van de opleiding hebben mede geleid tot de aandacht voor kwaliteitszorg in het onderwijs. Een bijzondere impuls werd gegeven door het invoeren van universitaire onderwijsvisitaties. De visitatierapporten geven inzicht in de onderwijssituatie van de acht medische faculteiten in Nederland. (Metz, Scherpbier en Van der Vleuten, 1995). Uit het visitatierapport dat in 2004 door de Quality Assurance Netherlands Universities (QANU) werd uitgevoerd blijkt dat er in de rapporten ook aandacht wordt besteed aan de docentprofessionalisering van de staf binnen de medische faculteiten. In het rapport over de medische faculteit van de Rijksuniversiteit Groningen wordt het volgende aangehaald over de didactische scholing van docenten: “De commissie heeft met genoegen vastgesteld dat veel docenten enthousiast betrokken zijn bij het onderwijs. Daartoe heeft ongetwijfeld het ruime

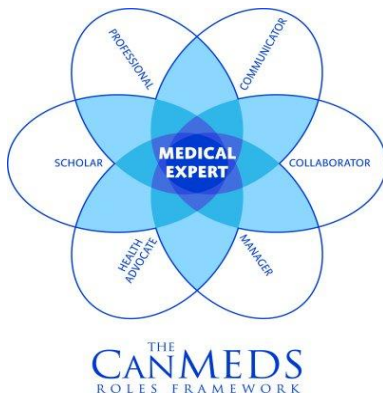
aanbod aan trainingen voor docenten bijgedragen. De opleiding is wel van mening dat de didactische professionalisering nog voor verbetering vatbaar is. De cursussen die worden aangeboden zijn niet verplicht (met uitzondering van de training voor tutoren), maar soms wel overtekend” (QANU, 2004, p. 125). Uit deze opmerking blijkt dat het enthousiasme onder artsen om te doceren aanwezig is maar dat de didactische scholing van docenten nog wel voor verbetering vatbaar is, dit punt behoeft dus nog meer aandacht. Zwierstra, Venekamp & Koppelman (2006) verbazen zich in 2006 nog over het feit dat in de meeste onderwijssectoren een kwalificatiesysteem bestaat dat is gekoppeld aan de onderwijsbevoegdheid van docenten maar dat een dergelijk kwalificatiesysteem ontbreekt in het academisch onderwijs. Bij verschillende universiteiten zijn wel kwalificatiesystemen aanwezig maar deze zijn elders niet toepasbaar. In dit artikel wordt wel geconstateerd dat er een landelijke werkgroep actief is vanuit de medische faculteiten die voorstellen voorbereidt voor een landelijk kwalificatiesysteem. Inmiddels is op woensdag 23 januari 2008 de overeenkomst 'Wederzijdse erkenning basiskwalificatie onderwijs' getekend door de veertien Rectores Magnifici van de Nederlandse universiteiten. “De basiskwalificatie onderwijs (bko) kan worden beschouwd als een bewijs van bekwaamheid voor docenten in het wetenschappelijk onderwijs” (Open Universiteit Nederland, z.j.). Met deze overeenkomst laten de universiteiten zien dat de trajecten voor docentprofessionalisering op orde zijn en onderling erkend worden (Open Universiteit Nederland, z.j.). Op deze manier worden ook de medische faculteiten verplicht om aan de professionalisering van docenten te werken en is het geen vrijblijvende activiteit meer.

### 1.3 Ontwikkelingen met betrekking tot competentiegericht onderwijs

De voorgaande paragraaf richtte zich vooral op de ontwikkelingen in de basisopleiding tot arts. In deze paragraaf zal vooral worden gefocust op de

vervolgopleiding van de arts. In het Kaderbesluit van het Centraal College Medisch Specialismen (CCMS) is besloten tot de modernisering van de medisch specialistische vervolgopleidingen (CCMS, 2004). Het doel van deze modernisering is te zorgen voor een competentiegericht curriculum dat leidt tot transparant en flexibel onderwijs (Medische vervolgopleidingen, geraadpleegd 10 mei 2008; Zwierstra et al., 2006). De opleiding tot arts richt zich op het ontwikkelen van algemene- en vakspecifieke competenties. De algemene competenties, geldend voor alle medisch specialisten, zijn gebaseerd op een onderzoek dat in Canada werd uitgevoerd, genaamd CanMeds 2000 (Canadian Medical Education Directions for Specialists). In 1990 rees de volgende vraag bij het Royal College of Physicians and Surgeons of Canada: 'Hoe kunnen we toekomstige artsen voorbereiden om in de toekomst effectief te werk te gaan, met het oog op de politieke en maatschappelijke veranderingen, en daarnaast tegemoet komen aan de zorg voor patiënten?' Aan het einde van de jaren '80 werden voor het eerst de zorgen uitgesproken over de veranderingen die optraden in de medische praktijk. Ook educatieve commissies rapporteerden over nieuwe onderwerpen en trends in het medisch onderwijs. Bij veranderingen moet gedacht worden aan; de informatievoorziening via internet, politieke regulaties, technologische ontwikkelingen in de medische wereld enz. Om een antwoord te kunnen geven op bovenstaande vraag werd een commissie opgericht, de Societal Needs working Group (SNGW). Deze commissie kreeg als opdracht om de belangrijkste competenties te identificeren die van toepassing waren op alle artsen, ongeacht het specialisme van de arts, om op deze manier aan de vraag van 'de toekomstige maatschappij' tegemoet te komen. De resultaten van dit onderzoek hebben uiteindelijk geleid tot een competentiemodel, de zogenaamde 'CanMeds-bloem'. Het CanMeds 2000 onderzoek heeft verschillende fasen doorlopen; fase (1): raamwerk-ontwikkeling, fase (2): pilot-project, fase (3): implementatie van het CanMeds-model. Op dit moment wordt gewerkt aan de laatste fase van het CanMeds onderzoek, fase (4): het doorvoeren van het

CanMeds-model binnen de initiële opleiding tot arts, in het programma van alle medische faculteiten. De 'CanMeds-bloem' wordt nu dus ook gebruikt om vorm te geven aan het medisch onderwijs. In het CanMeds-model worden de doelstellingen verwoord waaraan een medisch specialist moet voldoen. Het model staat hieronder afgebeeld (zie figuur 1.). De volgende competenties worden onderscheiden; medisch handelen, communicatie, samenwerking, kennis en wetenschap, maatschappelijk handelen, organisatie en professionaliteit (Frank, 2005).



**Figuur 1:** CanMeds 2000 competentiegebieden medisch specialist

6

De medisch specialisten die een opleidingstaak hebben moeten functioneren naar het CanMeds-model. In de ‘Scholar’-rol van de CanMeds-bloem moeten de medisch specialisten waarmaken dat ze een echte ‘docent’ zijn. De inhoud waarover de medisch specialist onderwijs geeft moet aan alle rollen van het CanMeds-model voldoen. De Scholar-rol komt dus later in de opleiding aan de orde. In de beschrijving van deze rol wordt duidelijk aandacht besteed aan de didactische vaardigheden waar de arts over moet beschikken. In bijlage (1) van deze scriptie is een uitgebreide beschrijving terug te vinden van deze ‘Scholar’ rol en de bijbehorende competenties.

Het CCMS heeft per competentie beschreven aan welke eisen de medisch specialist moet voldoen. Van de medisch specialist wordt gevraagd deze competenties op peil te houden. De medisch specialisten, tezamen met de wetenschappelijke verenigingen, zijn verantwoordelijk voor het houden van toezicht op het professioneel handelen van de medisch specialist en het bieden van ondersteuning bij de continue professionele ontwikkeling van de medisch specialist (Orde van medisch specialisten, geraadpleegd 10 mei 2008). “Competentiemanagement vormt daarmee de

kern van het professioneel kwaliteitsbeleid” (Orde van medisch specialisten, geraadpleegd 10 mei 2008). Naast de algemene competenties worden per medisch specialisme de vakinhoudelijke competenties beschreven. Om tot de ontwikkeling van de algemene- en vakspecifieke competenties te komen dienen er competentiegerichte opleidingsprogramma’s ontwikkeld te worden. Er zal bijvoorbeeld gebruik worden gemaakt van nieuwe toetsinstrumenten, zoals korte praktijk beoordelingen (KPB) of een portfolio, om te kijken of de competenties worden behaald. Of de innovaties in de medische vervolgoledingen slagen is mede afhankelijk van de professionalisering van de betrokken medisch specialisten als docent. Het ‘Teach the Teacher’ programma speelt in op de ondersteuning van de innovaties in het medisch onderwijs en biedt docentprofessionaliseringscursussen die aansluiten bij deze nieuwe ontwikkelingen (Zwierstra et al., 2006). Er wordt niet uitsluitend getraind in Groningen, de trainingen worden ook op locatie gegeven in de met het UMCG geaffilieerde ziekenhuizen. De cursusleiders reizen dus naar de cursisten toe. De programma’s zijn meer vraag- dan aanbodgericht. Voor het ‘Teach the Teacher’ programma, waarin drie modules worden aangeboden, houdt dit in dat affiliatieziekenhuizen aangeven welke modules uit het aanbod wanneer gewenst zijn. Daarnaast worden de cursussen niet alleen gegeven door onderwijsdeskundigen, juist ook artsen worden opgeleid tot cursusleiders. Doordat de arts de ‘taal’ van de cursist spreekt verhoogt dit de geloofwaardigheid van de cursus (Zwierstra et al., 2006).

In het kaderbesluit van 9 februari 2004 beschrijft het CCMS de eisen waaraan, specifiek een opleider binnen het medisch onderwijs moet voldoen. Onder Hoofdstuk C: De erkenning tot opleider en opleidingsinrichting, paragraaf I-A ‘Eisen voor erkenning van de medisch specialist tot opleider’ (zie bijlage 2.) staat genoemd dat de opleider moet beschikken over didactische kwaliteiten. Onder paragraaf I-B ‘Verplichtingen van de opleider’ wordt expliciet genoemd dat de opleider met betrekking tot bij- en nascholing aan de volgende eisen moeten

voldoen:

- zij houden hun kennis en inzicht als medisch specialist op peil door het regelmatig deelnemen aan geaccrediteerde bij- en nascholingsactiviteiten
- zij volgen systematisch geaccrediteerde bijscholing met didactische aspecten ten behoeve van de opleiding (CCMS, 2004).

Met het Kaderbesluit van de CCMS (2004) wordt de medisch specialist verplicht tot het deelnemen aan bijscholing in didactische vaardigheden. Het 'Teach the Teacher' programma kan een uitkomst bieden.

Parallel hieraan deed zich een andere ontwikkeling voor. In 2004 werd de werkgroep Docentprofessionalisering Medisch Onderwijs samengesteld. In juli 2007 heeft deze werkgroep een concept rapport uitgegeven genaamd: 'Competenties docenten in het medisch onderwijs'. Het document is tot stand gekomen op verzoek van de Onderwijs Commissie Geneeskunde van het Discipline-overleg Medische Wetenschappen die de initiële opleiding tot arts regardeert (OCG-DMW). Deze werkgroep is opgericht met als doel de oriëntatie op een landelijk toepasbaar docentkwalificatiesysteem voor het medisch onderwijs. In het voorjaar van 2005 werd een advies uitgebracht en in het najaar werd dit advies formeel overgenomen door de OCG-DMW. In dit advies werd benadrukt dat de werkgroep gecontinueerd diende te worden en het voorstel verder uitgewerkt moest worden. De werkgroep werd vanaf dat moment samengevoegd met de reeds bestaande werkgroep van de Nederlandse Vereniging voor Medisch Onderwijs (NVMO). Eén van de aanleidingen voor het beschrijven van docentcompetenties speciaal voor het medisch onderwijs is de specifieke context waarin dit onderwijs plaats vindt. Meestal hebben de werkzame docenten in het medisch onderwijs naast de onderwijskundige taken ook nog klinische- en onderzoekstaken. Daarnaast is er een ontwikkeling gaande in het universitair onderwijs die werkt aan wederzijdse erkenning van bestaande of nog te ontwikkelen basiskwalificaties onderwijs (Werkgroep docentprofessionalisering, 2007). De werkgroep geeft in

het document aan dat de medische vervolgopleidingen bezig zijn met een moderniseringsproces en dat in dit kader ook wordt gewerkt aan de beschrijving van competenties van opleiders. Afstemming van de competenties voor de medische vervolgopleidingen zal dus nog moeten plaatsvinden wanneer het moderniseringsproces wordt afgerond (Werkgroep docentprofessionalisering, 2007). In het najaar van 2007 is de eerste versie van de competentiebeschrijving beschikbaar gekomen en voorgelegd aan een grote groep deskundigen afkomstig uit de medisch- onderwijskundige wereld. Ondertussen is er in mei 2008 een nieuw document van deze werkgroep verschenen, waarin de reacties op het document, vanuit de medisch-onderwijskundige wereld, zijn meegenomen (Werkgroep docentprofessionalisering, 2008). Het 'Teach the Teacher' programma anticipeert op de ontwikkelingen rond de competentiebeschrijving voor de opleider. Door de medisch specialist te scholen en te ondersteunen in het geven van onderwijs, en daarmee een start te maken met het verwerven van de competenties voor het geven van onderwijs. De eerder genoemde CCMS initiatieven en de docentprofessionaliserings- activiteiten bleken naadloos op elkaar aan te sluiten.

Daarnaast is er binnen het onderwijs een verandering gaande naar andere vormen van onderwijs. "In de huidige nieuwe onderwijsvormen ligt de nadruk op actief, betekenisgericht, toepassingsgericht, zelfstandig en samenwerkend leren van studenten" (Dolmans, Stalmeijer & Wolfhagen, 2006, p. 267). Deze nieuwe onderwijsvormen moeten ervoor zorgen dat kennis beter beklijft, beter wordt begrepen en beter wordt toegepast. Daarnaast zullen deze onderwijsvormen moeten zorgen voor een betere voorbereiding op een 'leven lang leren'. De nieuwe onderwijsvormen stellen andere eisen aan docenten. De docent moet goed kunnen uitleggen en informatie over kunnen dragen maar daarnaast ook goede feedback kunnen geven, een goede coach zijn en kunnen begeleiden tijdens praktijkleren. Om docenten op een goede wijze voor te bereiden op deze nieuwe vormen van leren wordt er steeds meer aandacht besteed aan het opleiden en scholen van docenten (Dolmans et al., 2006). Het 'Teach the Teacher'

programma kan ondersteuning bieden bij deze nieuwe ontwikkelingen in het onderwijs. Samenvattend kan uit bovenstaande worden geconstateerd dat er binnen het medisch onderwijs een aantal ontwikkelingen gaande zijn op het gebied van de didactische professionalisering van opleiders, van de inhoud van de opleidingen (meer competentiegericht), en van de kwalificatie van docenten en opleiders. Deze masterthesis vormt een kleine bouwsteen te midden van deze ontwikkelingen.

#### **1.4 Context van de instelling waar het onderzoek wordt uitgevoerd**

Het onderzoek vindt plaats in het kader van de afstudeeropdracht voor de masteropleiding Onderwijskunde aan de Rijksuniversiteit Groningen. Het onderzoek wordt uitgevoerd binnen het Ontwikkelplatform van het Wenckebach Instituut van het Universitair Medisch Centrum Groningen (UMCG).

Het UMCG is één van de grootste ziekenhuizen in Nederland en met ongeveer 9000 personeelsleden één van de grootste werkgevers in het noorden van het land. In het UMCG zijn dagelijks zo'n 1000 patiënten opgenomen. Het UMCG heeft drie kerntaken gedefinieerd: zorg, onderwijs en onderzoek. Het Onderwijsinstituut en het Wenckebach Instituut van het UMCG richten zich daarbij op de tweede kerntaak van het UMCG, door zorg te dragen voor de ontwikkeling en opleiding van alle professionals in de zorg. Het Onderwijsinstituut is verantwoordelijk voor de opleiding tot arts, tandarts en bewegingswetenschapper, het Wenckebach Instituut verzorgt en ondersteunt in dit kader verpleegkundige beroepsopleidingen, vervolgoopleidingen en bij- en nascholingen voor zorgprofessionals in het UMCG en de regio. Tevens ondersteunt het instituut de professionele ontwikkeling van ziekenhuismedewerkers en hun leidinggevenden (UMCG, geraadpleegd: 27 mei 2008). Het Ontwikkelplatform van het Wenckebach Instituut laat

afdelingen van het UMCG (en andere opleidingsziekenhuizen in de regio) en studenten en docenten uit het hoger onderwijs optimaal van elkaar profiteren. Het Ontwikkelplatform inventariseert onderzoeksvragen waar studenten mee aan de slag kunnen. Artsen en andere professionals in de medische sector lopen in het dagelijkse werk tegen vragen aan waar studenten in het hoger onderwijs van niet-medische disciplines antwoord op kunnen geven. Binnen het Ontwikkelplatform voeren studenten vanuit verschillende disciplines onderzoek uit om deze onderzoeksvragen te beantwoorden.

#### **1.5 Aanleiding en probleemstelling van het onderzoek**

Het Wenckebach Instituut organiseert sinds 2002 het zogenaamde 'Teach the Teacher' programma. Het programma richt zich op alle medisch specialisten met opleidingstaken. De inhoud van het programma is gericht op het versterken en verbeteren van het didactisch handelen van de medisch specialisten. Binnen het 'Teach the Teacher' programma worden drie modules aangeboden. Onder paragraaf 1.6 zal verder worden ingegaan op de inhoud van de modules.

Om inzicht te krijgen in het effect van het 'Teach the Teacher' programma module (1): didactische basisprincipes, heeft de opdrachtgever het Ontwikkelplatform van het Wenckebach Instituut benaderd om een onderzoeksvraag te formuleren en bij deze onderzoeksvraag een student te selecteren die deze vraag kan beantwoorden.

In de medische wereld is al langer aandacht voor 'evidence based' werken. Sackett, Rosenberg, Gray, Haynes & Richardson (2000) geven de volgende definitie voor 'evidence based medicine': 'Evidence based medicine' is op een nauwgezette, expliciete en verstandige manier gebruik maken van huidige bevindingen die uit onderzoek voortkomen bij het maken van beslissingen over de zorg



voor individuele patiënten. De praktijk van het bedrijven van 'evidence based medicine' houdt in dat de individuele klinische expertise van de arts wordt geïntegreerd met het op dat moment best beschikbare externe klinische bewijs dat voortkomt uit systematisch onderzoek. Voortvloeiend uit de aandacht voor 'evidence based medicine' volgde ook de belangstelling voor het werken met bewezen methoden in het medisch onderwijs. Harden, Grant, Buckley en Hart (2000) melden het volgende over 'best evidence medical education' (BEME). BEME is het implementeren van onderwijsmethoden en benaderingen die gebaseerd zijn op bewezen onderzoeksresultaten. Met dit onderzoek zal een bijdrage worden geleverd aan de uitvoering van evidence based onderwijs. " 'Evidence based onderwijs' gaat er vanuit dat onderwijsbeleid en –praktijk gebaseerd zijn op de beste zekerheid over wat werkt (de werking moet bewezen zijn). Dit betekent dat specifieke onderwijsinterventies en onderwijsstrategieën wetenschappelijk geëvalueerd moeten worden, voordat deze worden aanbevolen of op brede schaal ingevoerd worden" (UMCG, geraadpleegd 27 juli 2008).

Naast het werken met 'evidence-based' methoden wordt er in het medisch onderwijs steeds meer aandacht besteed aan het onderwerp: 'docentprofessionalisering'. De laatste jaren heeft de faculteit der Medische Wetenschappen van het UMCG een goede naam opgebouwd met betrekking tot het onderwijs aan studenten Geneeskunde en Tandheelkunde. Voor een groot deel is dit te danken aan de inzet van vele docenten die verantwoordelijk zijn voor de opleiding van de medische studenten. Het UMCG wil deze reputatie graag behouden door grote waarde te hechten aan de systematische ontwikkeling van de didactische competenties van docenten. Het UMCG wil in de komende jaren meer ruimte voor onderwijstaken reserveren, op het gebied van onderwijs de carrièrepianing bevorderen, en door het instellen van onderwijskwalificaties eisen stellen aan didactische basiskennis- en vaardigheden van (nieuwe) docenten (UMCG, 2006). Mede onder invloed van het 'Teach the Teacher' programma dat van start ging in 2002 ging er steeds meer aandacht uit naar de professionalisering van docenten, daarnaast droeg de

oprichting van het Centrum voor docentprofessionalisering (CDP) bij aan deze ontwikkeling. Met het oog op docentprofessionalisering is het van belang dat de werking van het 'Teach the Teacher' programma wordt bewezen om op die manier een 'BEME' gefundeerde bijdrage te kunnen leveren aan de didactische scholing en daarmee de professionalisering van docenten. In het vervolgonderwijs en de bij- en nascholing van medisch specialisten vervult het UMCG een regionale (en deels landelijke) functie. Dit geldt specifiek voor de Onderwijs- en Opleidings Regio Noord- en Oost Nederland (OOR N&O), het UMCG heeft in deze regio een coördinerende en regierol op het gebied van zorg, onderwijs en opleidingen. Het UMCG vervult daarmee een voorbeeldrol voor de noordelijke opleidingsregio. Deze voorbeeldfunctie gaat ook op wanneer men praat over het continueren van het medisch onderwijs, ook wel 'medisch opleidingscontinuüm' genoemd. Om te zorgen dat de kwaliteit op een hoog niveau blijft is het belangrijk dat artsen zich constant blijven ontwikkelen, zowel in de initiële opleiding als in de vervolgopleiding (UMCG, 2006). In dit kader geven Dolmans, Stalmeijer en Wolfhagen (2006) aan dat er toenemende belangstelling bestaat voor het evalueren van docenten. Er wordt verondersteld dat het evalueren en het geven van feedback aanzet tot reflectie en de professionele ontwikkeling van de docent stimuleert, mogelijk leidt dit tot de verhoging van de kwaliteit van onderwijs en uiteindelijk van het leren door studenten.

Medisch specialisten leiden artsen in opleiding (aios) op tot specialisten. Het gaat hierbij voor het grootste deel om informeel leren op de werkplek. De inhoud van deze grotendeels informele opleidingen is afhankelijk van het medisch specialisme waarin de arts in opleiding wordt geschoold. Voor elk van deze meer dan 30 onderscheiden specialismen heeft de desbetreffende beroepsvereniging opleidingsdoelen en opleidingseisen vastgesteld. De voor de opleiding verantwoordelijke specialisten zijn veelal alleen vakinhoudelijk geschoold, zelden op didactisch gebied. Zoals al eerder werd genoemd is het doel van het 'Teach the Teacher' programma om te zorgen voor voldoende didactische competenties bij specialisten om

hun opleidingstaak naar behoren te kunnen uitvoeren. Deze competenties zijn onderdeel van de eisen die aan opleiders worden gesteld (zie bijlage 2.). Deze eisen werden eerder besproken in paragraaf 1.3. Medisch specialisten nemen tot nu toe op basis van vrijwilligheid deel aan het 'Teach the Teacher' programma. Zwierstra, Venekamp en Koppelman (2006) merken op dat deelnemers aan trainingen en cursussen tot nu toe vooral vanuit een interne motivatie worden gedreven. Zij verwachten dat met het toekennen van accreditatiepunten en met het stellen van explicietere eisen aan docenten op het gebied van didactische scholing door visiterende instanties (CCMS, QANU) de externe motivatie van deelnemers zal toenemen. Daarnaast zal de ontwikkeling in de verbetering van het loopbaanperspectief voor docerende artsen mogelijk kunnen bijdragen aan de externe motivatie van deelnemers aan cursussen en trainingen. Zoals uit de cursusevaluaties blijkt zijn de deelnemers zeer tevreden over de cursussen. Onduidelijk is echter in welke mate specialisten het geleerde toepassen in hun dagelijkse (opleidings)praktijk (UMCG, geraadpleegd 25 juli 2008). Het Wenckebach Instituut wil weten in hoeverre het 'Teach the Teacher' programma voor medisch specialisten effect heeft. Met het oog hierop willen zij een evaluatie uitvoeren om te bepalen of er sprake is van transfer van het geleerde in de cursus naar de werkvloer. Op dit moment beschikt het Wenckebach Instituut nog niet over een instrumentarium om deze transfer vast te stellen. Daarom wordt een onderzoek voorgesteld dat gericht is op de ontwikkeling en validering van een dergelijk instrument.

## 1.6 Doelstelling en operationalisatie van de onderzoeksvraag

De doelstelling van dit onderzoek richt zich op het verkrijgen van inzicht in instrumenten die het effect van het 'Teach the Teacher' programma module (1): 'didactische basisprincipes' betrouwbaar en valide kunnen meten. De onderzoeksvraag die voortvloeit uit deze doelstelling luidt:

*Is het mogelijk om een valide en betrouwbaar instrument te ontwikkelen, gericht op het meten van de door de medisch specialist geleerde vaardigheden in module (1) van het 'Teach the Teacher' programma?*

Uit deze onderzoeksvraag vloeien twee vraagstellingen voort.

De eerste vraagstelling luidt:

*In hoeverre zijn er instrumenten aanwezig om het effect van didactische basisvaardigheidstrainingen in het medisch onderwijs te bepalen en welke schalen uit deze instrumenten kunnen worden gebruikt voor het meten van het 'Teach the Teacher' programma module (1)?*

Met behulp van literatuuronderzoek wordt bepaald of er instrumenten bestaan om dit type training te evalueren. Mogelijk kan bij het ontwerpen van een meetinstrument gebruik worden gemaakt van kennis/onderdelen uit bestaande instrumenten.

De tweede vraagstelling luidt:

*Wat is de betrouwbaarheid en validiteit van het ontwikkelde instrument voor het bepalen van het effect van een training in didactische basisprincipes aan medisch specialisten?*

De verwachting is dat uit het literatuuronderzoek naar bruikbare (schalen in) bestaande instrumenten naar voren komt welk type instrumentarium voor het doel van deze studie het meest geschikt is. Het onderzoek naar de betrouwbaarheid en validiteit zal verschillen naar de aard van het instrumentarium dat wordt ontwikkeld.

De belangrijkste begrippen uit de onderzoeksvraag worden hieronder geoperationaliseerd. Baarda en de Goede (2001) geven aan dat met het operationaliseren van begrippen wordt bedoeld, het vertalen van een begrip in meetbare termen.

### **Valide**

Segers (2002) spreekt van een valide meting als de gemeten waarde van de variabele overeenkomt met de werkelijke waarde van die variabele bij het onderzoekselement.

Boendermaker (2003) spreekt van een valide meting wanneer met een bepaald instrument wordt gemeten wat de onderzoeker met het instrument beoogt te meten.

### **Betrouwbaar**

Boendermaker (2003) spreekt van betrouwbaarheid wanneer een meting reproduceerbaar is, dus bij herhaling van een meting is sprake van dezelfde uitslag als bij de eerste meting.

### **Vaardigheden**

Het Van Dale woordenboek der Nederlandse taal (1999) definieert het woord 'vaardigheid' als 'bekwaamheid' of het 'vaardig-zijn'. Het 'Teach the Teacher' programma richt zich specifiek op het aanleren van vaardigheden op het gebied van onderwijs, aan de medisch specialist. In module (1) van het 'Teach the Teacher' programma staat het aanleren van didactische basisprincipes centraal (Wenckebach Instituut, 2008).

### **Programma 'Teach the Teacher'**

Binnen het 'Teach the Teacher' programma zijn drie modules ontwikkeld: (1) training in didactische basisprincipes, (2) onderwijs tijdens de dagelijkse patiëntenzorg en (3) het begeleiden, toetsen en beoordelen van aios. Hieronder volgt een beschrijving van de drie cursussen:

- *Module (1) 'Teach the Teacher': didactische basisprincipes*

In deze cursus wordt inzicht verworven in de leerprocessen bij volwassenen op de gebieden van kennis, vaardigheden en professioneel gedrag. Er wordt geoefend met de uitvoering van korte onderwijsmomenten; het geven en ontvangen van feedback; het doceren van

vaardigheden en het uitvoeren van een korte praktijk beoordeling.

- *Module (2) 'Teach the Teacher': onderwijs in de dagelijkse praktijk*

In deze cursus wordt aandacht besteed aan allerlei vormen van onderwijs die zich in de dagelijkse praktijk aandienen. Hierbij draait het om cursorische onderwijsmomenten en onderwijs dat zich aandient tijdens het dagelijks werk in de patiëntenzorg. Vragen die in deze cursus worden behandeld zijn: Hoe kan ik het onderwijs op de polikliniek organiseren? Hoe maak ik van een patiëntenbespreking een effectief onderwijsmoment?

- *Module (3) 'Teach the Teacher': begeleiden, toetsen, beoordelen*

In deze cursus worden de functie en de mogelijkheden van het portfolio besproken. De cursisten krijgen inzicht in de samenhang tussen Korte Praktijk Beoordelingen (KPB), voortgangsgesprekken en beoordelingsgesprekken. Daarnaast verwerft men inzicht in de achtergronden en de principes van toetsen en beoordelen. Het grootste aandeel bestaat uit het oefenen met KPB-gesprekken, voortgangsgesprekken en beoordelingsgesprekken (Wenckebach Instituut, geraadpleegd 11 april 2008).

In dit onderzoek wordt getracht een instrument te ontwikkelen dat de transfer van de cursusinhoud naar de werkvloer kan gaan meten. Nijman (2004) vermeldt verschillende definities voor de term 'transfer van training'. De definities die in deze context van toepassing zijn, zijn die van Swinney (in Nijman, 2004, p. 10) "...that almost magical link between classroom performance and something which is supposed to happen in the real world". Daarnaast kan de definitie van Burke en Baldwin (in Nijman, 2004, p.10) ook in deze context dienen. Burke en Baldwin (in Nijman, 2004, p.10) definiëren 'transfer van training' als de mate waarin de cursisten de geleerde kennis, vaardigheden en gedragingen in de cursus, toepassen op de werkvloer (Nijman, 2004).

De laatste jaren is er steeds meer aandacht voor de effectiviteit en het rendement van trainingen. Ondanks de toegenomen aandacht voor dit onderwerp blijven de typische effectiviteitsvragen, zoals 'wat werkt' en 'waarom werkt het' nog vaak onbeantwoord en blijkt dat deze vraagstukken in de praktijk vaak weinig worden onderzocht (Sanders en Ardts, 2008). Nijman (2004) geeft aan dat het empirische bewijs dat aanwezig is vaak niet aantoont dat er werkelijk transfer optreedt van het geleerde in de cursus naar de werkvloer.

De meeste trainingen worden vaak alleen geëvalueerd op het eerste niveau van het model van Kirkpatrick, cliënttevredenheid (Sanders en Ardts, 2008; Nijman, 2004; Rubak, Mortensen, Ringsted, Mallings, 2008). Kirkpatrick heeft een classificatieschema ontwikkeld voor het evalueren van trainingen op vier niveaus. Het eerste niveau richt zich op het evalueren van de reactie van de cursisten, het meten van de perceptie van de deelnemers omtrent de cursus of training die is gegeven. Zijn de deelnemers tevreden met de cursus? In dit stadium worden alleen de reacties van de deelnemers gemeten, het meet niet of de deelnemers iets hebben geleerd in de cursus. Het tweede niveau van het model gaat in op de evaluatie van wat er is geleerd. Evaluatie vindt plaats om te bepalen of de deelnemers al of niet de inhoud van de cursus tot zich hebben genomen. Welke principes, feiten en technieken hebben de cursisten begrepen en zijn opgenomen door cursisten? Niveau (3) evalueert het gedrag van de cursisten. Er wordt geëvalueerd om te bepalen in welke mate de cursist de nieuw aangeleerde vaardigheden en/of kennis in zijn baan gebruikt. Voor veel mensen is dit het kernpunt: Kunnen ze iets met het geleerde in praktijk? Echter een moeilijke evaluatie waarbij goed moet worden geanalyseerd wanneer er wordt geëvalueerd, hoe vaak er wordt geëvalueerd en hoe er wordt geëvalueerd. Het laatste niveau, niveau (4), evalueert het uiteindelijke resultaat van de cursus. Het evalueert de organisatorische impact van de training op de werkkracht van het bedrijf. Het welslagen van de opleiding gemeten in voor managers begrijpelijke termen: geld, opbrengst. Hierbij moet worden gedacht aan toenames in verkoop, in de productie, verbeterde kwaliteit, afgenomen kosten, minder ongevallen

enzovoort. Voor organisaties zijn dit namelijk de punten waarom wordt geïnvesteerd in opleidingen. Deze evaluatie is echter erg moeilijk omdat er meer zaken van invloed zijn dan enkel de opleiding (Businessballs, z.j., Kirkpatrick, 1996, Jellema, 2003).

Bij de te ontwikkelen vragenlijst zal worden nagegaan wat er van de cursusinhoud terecht komt op de werkvloer. Er zal worden getracht een instrument te ontwikkelen dat evalueert op *Niveau (3): Evaluatie van gedrag*. Wat past de cursist in de praktijk toe van het geleerde in de cursus?

Sanders en Ardts (2008) verbazen zich over het feit dat er meestal alleen op niveau (1) en (2) van het evaluatiemodel van Kirkpatrick wordt geëvalueerd. Bedrijven investeren vaak veel geld in opleidingen en trainingen van medewerkers en de verwachtingen in de verbetering van kennis en vaardigheden of een toename van productiviteit zijn vaak hooggespannen.

De redenen die Sanders en Ardts (2008) noemen voor het feit dat cursussen vaak niet worden geëvalueerd zijn:

- Vaak komen 'waarom' vragen aan de orde als het al goed is misgegaan, bij leerinterventies treden deze vragen dus pas op wanneer er echt aanwijzingen zijn dat een training of programma echt beter kan of als er duidelijk twijfels bestaan over de werking van de training.
- Vaak wordt er een beroep gedaan op onderzoeksmatige kennis, de meeste HRD-professionals hebben weinig affiniteit met het uitvoeren van dit soort onderzoek.
- Vaak worden de leerdoelen van te voren niet goed vastgesteld en vaak ook niet op een manier dat ze meetbaar zijn, voor het vaststellen van effectiviteit zijn meetbare doelen nodig (Sanders en Ardts, 2008).

Vanuit professioneel oogpunt is het zeker van belang om trainingen, cursussen en programma's te evalueren. Door de focus te leggen op het verklaren van de effectiviteit van een training is het op langere termijn mogelijk het mechanisme te achterhalen waardoor bepaalde trainingen of interventies meer of minder effectief zijn. Dit komt de effectiviteit van toekomstige leerinterventies ten goede en daarnaast wordt op deze manier de professionaliteit van die HRD'ers en HRM'ers vergroot waar de evaluatie van

interventies tot het takenpakket behoort. (Sander & Ardts, 2008).

Onderzoek naar de effectiviteit van leerinterventies richt zich op drie verschillende vlakken: kenmerken van de deelnemer, kenmerken van de interventie (training) en kenmerken van de context waarin de deelnemer het geleerde in praktijk brengt, de werkomgeving. Er is nog een vierde gebied van onderzoek, namelijk de interactie tussen bovenstaande vlakken (Sander & Ardts, 2008, Nijman, 2004). In de volgende paragraaf zal nog kort worden ingegaan op effectstudies van soortgelijke trainingen als het 'Teach the Teacher' programma.

### 1.7 Effectstudies naar soortgelijke trainingen als 'Teach the Teacher'

In deze paragraaf zal kort worden ingegaan op wat er in de literatuur bekend is over effectstudies van soortgelijke trainingen als het 'Teach the Teacher' programma.

Notzer en Abramovitz (2008) onderzochten in welke mate een korte workshop bijdraagt aan het verbeteren van de klinische instructie op lange termijn en welke dimensies van de klinische instructie specifiek werden verbeterd. Het programma besteedde vooral aandacht aan professioneel gedrag in de instructeur- student- patiënt relaties. Daarnaast werd er aandacht besteed aan het oefenen met gesimuleerde klinische instructiemomenten. De studie maakte gebruik van een pre- en posttest met een vergelijkingsgroep. De doceervaardigheden werden een jaar na de workshop beoordeeld door studenten die feedback gaven op de uitvoering van de vaardigheden. Met behulp van een vragenlijst werd op de volgende dimensies gemeten: 'overall assessment of the instructor', 'presentation of theoretical material', 'contribution to clinical training', 'instructor-student relationship', 'tutor availability to students'. De resultaten van het onderzoek laten zien dat de doceervaardigheden significant verbeterden op de vijf dimensies van klinische instructie bij

de groep die deelnam aan de workshop, bij de controlegroep bleven de scores onveranderd.

De workshop vertoont overlap met het 'Teach the teacher' programma module (1) op de punten: 'feedback geven', 'demonstreren van enthousiasme' daarnaast wordt er ook geoefend met 'klinische instructiemomenten'. Ook in het 'Teach the Teacher' programma is er uitgebreid aandacht voor het oefenen van het geven van een onderwijsmoment. Er werd een specifiek evaluatie-instrument ontwikkeld dat aansluit bij de inhoud van de cursus. Er wordt in het artikel niets gemeld over de betrouwbaarheid en validiteit van dit instrument.

Malling, Bested, Skjelsager, Ostergaard en Ringsted (2007) rapporteerden over het effect van een training gericht op het beoordelen assistenten in opleiding tot anesthesioloog. De doelen van het onderzoek zijn het bepalen van de korte- en lange termijn effecten van deze cursus. Onderwerpen die aan de orde kwamen in de cursus waren: formatieve- en summatieve evaluatie, feedback geven en ontvangen, beoordelingsgesprek voeren. Voor de pre- en posttest werden dezelfde vragenlijsten gebruikt. Er werd een significante verbetering gevonden van kennis over het beoordelen van leerlingen gelijk na het volgen van de cursus en ook bij de tweede meting (follow-up) werd een toename in kennis geconstateerd. De conclusie van dit onderzoek luidt dat de kennis van deelnemers over beoordeling van leerlingen duidelijk is toegenomen. De data werd verzameld door de medisch specialisten die deelnamen aan de cursus via de mail te benaderen en te vragen om antwoord te geven op vijf open vragen die gebaseerd waren op de inhoud van de cursus. Daarnaast werd aan de cursisten gevraagd om te beoordelen hoe vaak de cursist de geleerde vaardigheden in praktijk bracht. Op het geven van feedback werd de grootste vordering geboekt. Hierbij moet worden gemeld dat de medisch specialisten zichzelf beoordeelden. De inhoud van deze cursus komt voor een groot deel overeen met het 'Teach the Teacher' programma module (1). In deze studie werd gemeten op niveau (2): 'evaluatie van het leren' en niveau (3): 'verandering in gedrag' van het

model van Kirkpatrick. In de vragenlijst werden vragen gesteld over de kennis die in de cursus aan de orde waren geweest en of deze kennis door de medisch specialisten kon worden gereproduceerd. Het instrument is dus specifiek ontwikkeld op de aangeboden cursusinhoud. Het is natuurlijk de vraag of de artsen het geleerde in de cursus werkelijk toepassen op de werkvloer omdat er gebruik werd gemaakt van zelfevaluatie. De cursus richt zich voornamelijk op het onderwerp: 'beoordelen van leerlingen'. Het 'Teach the Teacher' programma module (1) heeft ook aandacht voor aspecten als: 'het scheppen van een veilig leerklimaat' en de 'organisatie van een onderwijsmoment'.

Busari, Scherpbier, Van der Vleuten en Essed (2006) rapporteren over het effect van een tweedaagse workshop op de doceervaardigheden van artsen in opleiding. De studie werd uitgevoerd op de afdeling: Pediatrie, Obstetrie en Gynaecologie van de UMC's van Leiden en Amsterdam. De experimentele groep onderging een tweedaagse workshop gericht op doceervaardigheden. De dataverzameling bestond uit twee delen:

Er werd gemeten of de arts in opleiding tevreden was met de workshop, dit gebeurde door een vragenlijst over de satisfactie met de structuur van de workshop en de kwaliteit van de instructie en de inhoud van de verschillende onderwerpen: helderheid en organisatie, presentatiestijl, groepsinteractie en de inhoud van de onderwerpen. Daarnaast werden de artsen in opleiding door de medisch studenten beoordeeld om de effectiviteit van de klinische instructie te meten. De resultaten laten zien dat er een significante verbetering optrad in de experimentele groep. De scores op de post testen van beide groepen op 'overall teaching ability' waren hoger in de experimentele groep ten opzichte van de controlegroep maar het verschil was niet significant. De workshop bestond uit de volgende onderwerpen: effectief doceren, zelfkennis en in hoeverre cursisten in staat zijn om te doceren, geven van feedback, voorkennis te activeren, problemen op te lossen en tijd goed te managen. Ook in deze cursus komen onderwerpen aan de orde die ook aandacht krijgen in het 'Teach the Teacher' programma,

zoals 'feedback-vaardigheden' en het 'activeren van voorkennis'. In de voorgaande studies werd een meetinstrument gebruikt wat specifiek aansloot bij de cursusinhoud. Het is de vraag of daar in deze studie rekening mee is gehouden. Daarnaast moet gemeld worden dat het een cursus betrof die werd aangeboden aan artsen in opleiding. Het eerste deel van de evaluatie richtte zich vooral op het evalueren van niveau 1 van het Kirkpatrick model, zijn de cursisten tevreden met de cursus? In het tweede deel werden de aiOS beoordeeld op de doceervaardigheden door medische studenten hier wordt op niveau 3 gemeten. Bringen de aiOS het geleerde in praktijk?, ook in deze evaluatie zal sprake zijn van evaluatie op dit niveau.

Godfrey, Dennick en Welsh (2004) rapporteren over een studie die werd uitgevoerd onder artsen van verschillende specialismen. Er werd een vragenlijst bestaande uit achttien items, die verwezen naar doceervaardigheden, afgenomen onder de participanten van het onderzoek. In het onderzoek wordt niet genoemd hoe de items tot stand zijn gekomen. Er werd gewerkt met een experimentele groep die deelnam aan de cursus en een controlegroep die op de wachtlijst stond voor deze cursus. De cursus bestond uit een workshop die praktische handvatten bood voor het geven van onderwijs, bijvoorbeeld het aanbieden van korte onderwijs momenten en daaropvolgend reflectie met participanten van de cursus. De cursusedagen zijn verdeeld over een aantal weken zodat de cursist het geleerde gelijk kan oefenen in praktijk. Gelijk na de deelname aan de cursus werd een meting verricht met behulp van vragenlijsten, acht à tien maanden later werd er nog een follow-up meting uitgevoerd. Voor de controlegroep werd de meting met dezelfde tijdsintervallen afgenomen. Respondenten werd gevraagd om zelf een beoordeling te geven in de mate waarin de cursist zichzelf in staat achtte om de doceervaardigheid uit te voeren en daarnaast werd gevraagd om aan te geven hoe vaak de vaardigheid werd uitgevoerd in praktijk, hoe zelfverzekerd de specialist is in het doceren en het effect van het eigen doceren aan te geven. De meerderheid van de participanten aan de cursus rapporteerde positieve verbeteringen in

doceervaardigheid, als groep werd er op zestien van de achttien vaardigheden significant beter gescoord. De controlegroep liet een verbetering zien op één vaardigheid. Opvallend is dat de toename in frequentie waarin de vaardigheid werd toegepast maar voor vier van de achttien vaardigheden significant verbeterde. Wanneer op individueel niveau werd gekeken bij de deelnemers aan de cursus werd er een verbetering geconstateerd op vier of meerdere vaardigheden, de 25 % deelnemers aan de top verbeterden met zeven of meer vaardigheden. In de controlegroep werd per individu een verbetering gezien van één vaardigheid of meer en de top 25% cursisten verbeterden met twee of meer vaardigheden. Bij deze studie moet worden opgemerkt dat er binnen dit onderzoek gebruik werd gemaakt van zelfevaluatie, de vraag is of dit voor een objectieve meting zorgt.

Rubak, Mortensen, Ringsted en Malling (2008) voerden een studie uit naar korte- en lange termijn effecten van een drie- daagse 'Train the Trainer' cursus. De onderwerpen die aan de orde kwamen betroffen: 'guiding and instructing trainees'; 'delivering clinical supervision'; 'giving and receiving feedback'; 'structuring a clinical training session'; 'career counseling'; 'improving the learning climate'; 'planning and administrating education'; 'and the role of clinical trainer'. Er werd gebruik gemaakt van een pre- en een post test op lange termijn. En er werd gewerkt met een interventie- en controle groep. De metingen werden verricht met een schriftelijk toets (toename in kennis te meten), de metingen betreffende een verandering in gedrag van de medisch specialist en een verandering met betrekking tot het leerklimaat werden geëvalueerd door vragenlijsten af te nemen. De schriftelijk toets bevatte zes vragen en betroffen de belangrijkste onderwerpen die in de cursus aan bod kwamen. De vragenlijst over doceergedrag en leerklimaat werd geconstrueerd op basis van de cursusinhoud en aanbevelingen vanuit de literatuur. De vragenlijst bestond uit 32 items. De vragenlijst werd voor de cursus ingevuld en na zes maanden. De resultaten van de studie laten zien dat er een significante toename optreedt in kennis bij de interventiegroep. In de

controlegroep trad er een significante vermindering op van kennis. De data met betrekking tot de vragenlijst werd opgedeeld in aios en medisch specialisten. De cursisten beoordeelden zichzelf met een vragenlijst. De interventiegroep verbeterde significant (na zes maanden) op het geven en ontvangen van feedback en het geven van supervisie. Daarnaast werd er een verandering geconstateerd in doceergedrag dat leidde tot een meer gestructureerde en meer leerling-gerichte benadering in vergelijking tot de controlegroep. Er werden geen verschillen gevonden tussen aios en medisch specialisten. De interventiegroep scoorde ook significant hoger op het bewustzijn van potentiële leersituaties voor leerlingen. Hierbij wordt vermeld dat er al bij de voormeting een significant verschil werd geconstateerd op dit gebied. Daarnaast werd aan de aios in de controle- en interventiegroep een beoordeling gevraagd in hoeverre het leerklimaat was veranderd. Bij de interventiegroep trad een significante verbetering op in vergelijking tot de controlegroep. De onderzoekers geven zelf al aan dat er mogelijk een vertekend beeld kan optreden omdat er gebruik wordt gemaakt van twee groepen uit twee verschillende ziekenhuizen. Er wordt wel gebruik gemaakt van een voormeting zodat het beginpunt kon worden vastgesteld maar het kan zijn dat er andere factoren op de achtergrond meespelen in de ziekenhuizen zelf die de resultaten mogelijk beïnvloeden. Dit wordt deels ontkracht omdat het verschil tussen interventie- en controlegroep uiteindelijk significant blijkt te zijn. Er wordt elke keer gebruik gemaakt van dezelfde vragen bij het schriftelijk examen dit kan mogelijk tot een leereffect leiden bij beide groepen. Ook dit wordt weer recht getrokken omdat er op zowel de vragenlijst en het schriftelijk examen een verschil optreedt tussen de groepen. Daarnaast wordt gemeld dat de interventiegroep al significant hoger scoorde op voorgaande ervaring met cursussen op het gebied van doceren en leren. Mogelijk gaat de interventiegroep al anders om met de informatie en het geleerde in de cursus en scoren de groep daarom alleen al hoger en niet alleen vanwege de interventie. Verder is er in deze studie opnieuw gebruik gemaakt van een meting door de medisch specialist

of aios zelf. Om de meting betrouwbaarder te maken zou er in een volgend onderzoek mogelijk gebruik gemaakt kunnen worden van een meting door aios of co-assistenten. In dit onderzoek is wel gebruik gemaakt van een meting door de aios/co-assistent en een meting door de medisch specialist zelf.

## 1.8 Samenvatting hoofdstuk 1 en beschrijving van de opzet van de masterthesis

In dit hoofdstuk werd de basis geschetst voor deze masterthesis. De laatste ontwikkelingen met betrekking tot het medisch onderwijs werden uitgebreid beschreven. Er is steeds meer aandacht voor kwaliteitszorg in het medisch onderwijs. Daarnaast wordt steeds meer aandacht besteed aan docentprofessionalisering binnen de medische context. Vanuit de aandacht voor 'evidence based medicine' ontstond ook de aandacht voor 'evidence based' methoden in het medisch onderwijs het zogeheten 'best evidence medical education' (BEME). Het competentiegerichte onderwijs is in opkomst, binnen de medische context wordt nu gewerkt met het CanMeds-model waarin duidelijk aandacht wordt besteed aan de 'Scholar-rol' van de medisch specialist. Sanders en Ardts (2008) merken op dat er de laatste jaren steeds meer aandacht is voor de effectiviteit en het rendement van trainingen. Ondanks de toegenomen aandacht voor dit onderwerp blijven de typische effectiviteitsvragen, zoals 'wat werkt' en 'waarom werkt het' nog vaak onbeantwoord en blijkt dat deze vraagstukken in de praktijk vaak weinig worden onderzocht (Sanders en Ardts, 2008). Nijman (2004) geeft aan dat het empirische bewijs dat aanwezig is vaak niet aantoont dat er werkelijk transfer optreedt van het geleerde in de cursus naar de werkvloer. De meeste trainingen worden vaak alleen geëvalueerd op het eerste niveau van het model van Kirkpatrick, cliënttevredenheid (Sanders en Ardts, 2008; Nijman, 2004; Rubak, Mortensen, Ringsted, Malling, 2008). In dit onderzoek zal worden nagegaan of er instrumenten aanwezig zijn die het effect van het 'Teach the Teacher' programma module (1) kunnen meten. Wanneer dit niet

het geval is zal gestart worden met de ontwikkeling van een nieuw instrument. Er zal getracht worden een instrument te ontwikkelen dat nagaat wat er van de cursusinhoud terecht komt op de werkvloer. In dit onderzoek zal getracht worden een instrument te ontwikkelen dat evalueert op *Niveau (3): Evaluatie van gedrag* van het Kirkpatrick model. Wat past de cursist in de praktijk toe van het geleerde in de cursus?

De opzet van deze masterthesis zal er als volgt gaan uitzien. In het tweede hoofdstuk van deze scriptie wordt het theoretisch kader voor dit onderzoek beschreven. Vanuit dit theoretisch kader wordt een conceptueel kader opgesteld dat de rode draad van deze scriptie zal vormen. De onderzoeksmethoden van het onderzoek worden in hoofdstuk (3) besproken. Het onderzoek valt in twee delen uiteen, enerzijds een literatuuronderzoek naar bestaande instrumenten en een uitgebreide analyse van de psychometrische kwaliteiten van de geselecteerde instrumenten anderzijds het statistische onderzoek gericht op de ontwikkeling van het instrument en het bepalen van de validiteit en betrouwbaarheid. Hoofdstuk (4) beschrijft de resultaten van het onderzoek. Tot slot geeft hoofdstuk (5) de conclusies weer van het onderzoek gevolgd door de discussiepunten.



## 2 Theoretisch kader

### 2.1 Inleiding

In dit hoofdstuk zal de theoretische fundering voor het onderzoek plaatsvinden. In de eerste twee paragrafen, 2.2 en 2.3, worden de onderwijskundige- en psychologische theorieën besproken die zullen dienen om het theoretisch kader van deze masterthesis te schetsen. De paragrafen 2.2 en 2.3 beschrijven de onderwijskundige modellen die gebruikt zullen worden voor het conceptueel kader. De modellen van Gagné en van Van Gelder worden geanalyseerd en met elkaar vergeleken. In de paragrafen 2.4 en 2.5 worden de twee onderwijskundige modellen aangevuld met theorie vanuit de psychologie. Geschetst wordt waarom het ‘*creëren van veiligheid*’ belangrijk is binnen de medische context. De sociaal-cognitieve leertheorie van Bandura wordt in verband gebracht met het *rolmodel-leren* tijdens de opleiding tot arts. Paragraaf 2.6 besteedt beknopt aandacht aan een aantal grondprincipes die van belang zijn bij het leren van een volwassene. In de afsluitende paragraaf worden de genoemde theorieën bij elkaar gebracht en wordt een conceptueel kader geschetst van kernpunten die aandacht behoeven bij de ontwikkeling van het meetinstrument.

### 2.2 Instructional events van Gagné

In het model dat in de eerste plaats bedoeld is voor leraren in het basisonderwijs beschrijft Gagné een reeks van stappen die de docent in een lesuur neemt. De negen stappen vormen samen een voorschrift voor de docent die lessen systematisch wil voorbereiden en uitvoeren. De docent en de leerlingen samen kunnen worden opgevat als het instructiesysteem. De ‘Instructional Events’ van Gagné

vormen dan, enerzijds een model daarvan, maar anderzijds een model dat een voorschrijvend karakter heeft, in de eerste plaats voor de docent, maar desgewenst ook voor de ontwerper. Het model van de ‘Instructional events’ van Gagné geeft een beschrijving van hoe onderwijzen en leren tijdens de les plaatsvinden en schrijft voor hoe ze zouden moeten plaatsvinden (Warries & Pieters, 1992).

Het model van Gagné is ontstaan tijdens een samenwerking met Leslie Briggs binnen de Florida State University. Het model is een onderdeel van de meer omvattende Gagné/Briggs theorie die een beschrijving geeft van de vijf leersoorten, de leercondities en de leerhiërarchie volgens Gagné et al. (1988) (Warries & Pieters, 1992).

#### ***De vijf leersoorten***

Gagné heeft het menselijke leren geclassificeerd in vijf domeinen:

- *verbale informatie*
- *houdingen*
- *intellectuele vaardigheden*
- *motorische vaardigheden*
- *cognitieve strategieën*

Gagné gaat ervan uit dat *leren* plaatsvindt wanneer een individu een bepaalde vaardigheid verwerft om een bepaalde gedraging uit te voeren. Omdat de geleerde vaardigheid op zichzelf niet observeerbaar is, zal aan het gedrag van de lerende te zien zijn of een bepaalde vaardigheid is verworven. Gagné geeft aan dat verschillende geleerde vaardigheden resulteren in verschillende uitkomsten. Wanneer op deze uitkomsten wordt geanticipeerd en deze worden gepland, worden deze aangeduid als de instructiedoelen. De vaardigheid die wordt verworven als een persoon *verbale informatie* aanleert noemt Gagné ‘stating information’, het verklaren van informatie. De vaardigheid die wordt verworven als een individu een bepaalde *houding* aanneemt hangt samen met de keuze om je op een bepaalde manier te gedragen. Wanneer een individu een bepaald concept heeft geleerd is

de persoon in staat om elk nieuw voorbeeld van het concept correct te identificeren of te classificeren, daarmee laat de persoon zien dat hij een *intellectuele vaardigheid* heeft verworven. Wanneer een *motorische vaardigheid* is verworven, is de persoon in staat om de subvaardigheden van de motorische vaardigheid in de goede volgorde en op een goede en correcte manier uit te voeren. Ten slotte, wanneer een individu een nieuwe oplossing voor een probleem heeft gevonden of een persoonlijk systeem heeft uitgedacht om informatie te onthouden laat hij zien dat hij zichzelf een *cognitieve strategie* eigen kan maken (Aronson & Briggs, 1983).

### **De leercondities**

Gagné geeft aan dat aan bepaalde startcondities moet zijn voldaan om het mogelijk te maken dat bepaalde uitkomsten optreden. Gagné maakt onderscheid in interne en externe condities.

De interne condities staan voor kennis die de leerling al eerder heeft verworven en die essentieel en ondersteunend kunnen zijn voor het verdere leren. De externe condities, buiten de leerling om, helpen bij het activeren en ondersteunen van interne leerprocessen. (Warries & Pieters, 1992) "Volgens Gagné is de essentiële interne conditie voor het leren van elke intellectuele vaardigheid, dat de noodzakelijke voorkennis al is geleerd. Een belangrijke externe conditie is het stimuleren van de herinnering aan die voorkennis" (Warries & Pieters, 1992, p.172). Voorbeelden hiervan zijn:

- De voorwaarden voor het leren van een hogere-orde regel zijn dat de regels die eerder geleerd zijn weer worden opgehaald.
- Voordat een leerling een begrip kan leren moet hij eerst in staat zijn om te discrimineren tussen kritische kenmerken van begrippen.

### **De leerhiërarchie**

Vanuit het domein van de intellectuele vaardigheden blijkt dat de noodzakelijke voorwaardelijke kenniselementen en de relatie tot elkaar een leerhiërarchie vormen. Het uiteindelijke doel bevindt zich bovenin de leerhiërarchie, naar beneden bevinden zich de noodzakelijke voorkenniselementen. De hiërarchie in intellectuele vaardigheden is als volgt: associaties of stimulus-respons verbindingen; discriminaties (van attributen); begrippen of concepten; regels of principes; hogere- orde regels. (Warries & Pieters, 1992).

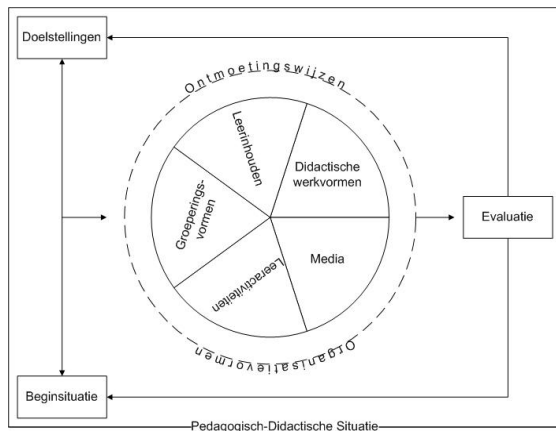
### **De negen instructional events van Gagné**

De negen stappen in het model van Gagné geven een beschrijving van de acties die een docent moet ondernemen om ervoor te zorgen dat leerlingen leren. De eerste drie activiteiten vormen een voorbereiding of inleiding tot de instructie. Dan volgen vier stappen voor de eigenlijke instructie en de laatste twee stappen van het model vormen de afsluiting van de instructie. Hieronder zijn de negen stappen van Gagné beschreven:

1. Zorgen dat de leerlingen aandacht hebben voor de les
2. Uitleggen wat de leerdoelen van de les zijn
3. Noodzakelijke voorkennis in herinnering brengen
4. De informatie (de leerstof) aanbieden
5. De leerlingen helpen bij het leren
6. Leerlingen het gevraagde gedrag laten tonen
7. Feedback geven over de performance
8. Het eindresultaat vaststellen van het leren
9. Herhaling en wijzen op toepassingen

In bijlage (3) is een schema opgenomen, dit schema geeft een beschrijving van de inhoud van de 'Instructional Events' van Gagné.

## 2.3 Model 'Didactische Analyse' van Van Gelder



**Figuur 2:** Didactische analyse Van Gelder

In het model van 'Didactische analyse' worden verschillende overeenkomsten geconstateerd met het model van Gagné. Aan het einde van deze paragraaf worden de overeenkomsten van dit model uiteengezet. Om inzicht te krijgen in het model van 'Didactische analyse' zijn een aantal vragen geformuleerd, zogenaamde 'didactische sleutelvragen'. Met deze vragen kunnen de begrippen en samenhangen van het model worden verduidelijkt. De didactische sleutelvragen verwijzen naar een aantal aspecten die relevant zijn bij het doceren. Deze aspecten kunnen fungeren als steunpunten voor de lesvoorbereiding, lesrealisatie en lesevaluatie.

Sleutelvragen die helpen bij het analyseren van de didactische situatie zijn:

1. *Wat moeten de leerlingen bereiken?*

Met het antwoord op deze vraag wordt duidelijk welke doelen bereikt dienen te worden aan het einde van het onderwijsmoment. De formulering van deze vraag is van groot belang, het maakt nogal verschil of de leerlingen na één les het register van een atlas moeten kunnen hanteren of pas na drie lessen.

2. *Hoe kan ik aansluiten bij de leerlingen?*

Om een doel te kunnen bereiken moet eerst de beginsituatie van de leerlingen bekend zijn. Zijn de benodigde vaardigheden bij de leerling aanwezig om het lesdoel te kunnen bereiken?

3. *Hoe kan ik het onderwijs geven?*

In het concreet verzorgen van een pedagogisch-didactische situatie zijn twee belangrijke elementen te onderscheiden. Enerzijds de ontmoetingswijzen tussen de leerlingen, de leerkracht en de leerinhouden en anderzijds de organisatie en maatregelen die moeten worden genomen om tot het onderwijsmoment te komen.

Deze didactische sleutelvraag wordt in het model van Van Gelder in verschillende vragen opgesplitst om tot een goed verloop van de ontmoetingswijzen en organisatievormen te komen:

a. *Hoe dien ik de leerstof te kiezen en te ordenen?*

Bij leerstof moet worden gedacht aan datgene wat aan de leerlingen wordt aangeboden, om te kunnen leren, om te kunnen oefenen, om een oordeel te kunnen vormen.

b. *Welke leeractiviteiten dien ik de leerlingen toe te laten passen?*

Onder leeractiviteiten moet worden verstaan de manier waarop de leerlingen zich de stof eigen maken (luisteren, oefenen, samenwerken). De leeractiviteiten worden, naast de mogelijkheden van leerlingen en het doel dat wordt nagestreefd met de les, ook bepaald door de leerstof, de keuze van didactische werkvormen en de keuze van de onderwijs- en leermiddelen.

c. *Wat doe ik als leerkracht om de leerlingen te helpen?*

De didactische werkvormen zijn de wijzen waarop de docent de leerlingen in contact brengt met de leerstof, bijvoorbeeld: vertellen, vragen stellen, groepswork enz. De keuze voor een bepaalde werkvorm hangt samen met de doelstellingen, met de leerstof die wordt aangeboden en de leeractiviteiten van de leerlingen

d. *Welke onderwijs- en leermiddelen dien ik te gebruiken?*

De keuze van onderwijs- en leermiddelen houdt nauw verband met de leerstof die wordt aangeboden.

e. *Hoe organiseer ik de leerlingengroep (groeperingsvormen)?*

Hier wordt bedoeld de manier waarop de leerlingen het beste kunnen worden gegroepeerd met het oog op het doel van de les, bijvoorbeeld: klassikaal in een grote groep of in kleine groepjes.

4. *Met welk resultaat heb ik mijn onderwijs gegeven?*

Aan het einde van een les wordt nagegaan welke resultaten zijn geboekt. Uitgaande van de beginsituatie wordt vastgesteld in hoeverre de lesdoelen zijn bereikt door middel van de reeks onderwijsleersituaties.

De didactische sleutelvragen zijn in het model van Van Gelder verwerkt. Uit het schema blijkt, dat er een nauwe samenhang bestaat tussen de diverse componenten van het schema

(Van Gelder, Oudkerk Pool, Peters & Sixma, 1979).

Na de analyse van bovenstaande onderwijskundige modellen wordt duidelijk dat de modellen overlap vertonen op verschillende punten. In beide modellen is aandacht voor het opstellen van de doelen van een onderwijsmoment, wat moeten de leerlingen aan het einde van de les bereikt hebben? (Van Gelder didactische sleutelvraag 1, Gagné event 2). Daarnaast is er aandacht voor het bepalen van de beginsituatie van de leerlingen, waar zijn de leerlingen en hoe kan op een goede manier bij dit niveau worden aangesloten? (Van Gelder didactische sleutelvraag 2, Gagné event 3). De modellen hebben allebei aandacht voor het aanbieden van de leerstof, op welke manier dit moet gebeuren en met welke hulpmiddelen

(Van Gelder didactische sleutelvraag 3, Gagné event 4 en 5). Het laatste onderdeel waarop deze modellen overeenkomst vertonen is de evaluatie van de resultaten, er kan gesproken worden van een soort 'loop', gekeken wordt of de doelen die werden opgesteld zijn bereikt aan het einde van het onderwijsmoment (Van Gelder didactische sleutelvraag 4, Gagné event 8). In het conceptueel kader dat aan het einde van deze paragraaf zal worden beschreven wordt hier dieper op ingegaan.

## 2.4 Sociaal-cognitieve leertheorie van Bandura

*“Do as I say, not as I do”, seldom works. What we do is likely to have more impact on learners than what we tell them to do”* (citaat van Westberg & Jason, 1993, in: Harden & Crosby, 2000, p. 338).

Wanneer mensen voor het eerst een sport gaan beoefenen weet het individu vaak al welke stappen ongeveer moeten worden ondernomen voordat het spel voor de eerste keer wordt beoefend. Dit komt voort uit het feit dat mensen het spel al een keer hebben gezien, uitgevoerd door andere individuen. Albert Bandura (1977) herkende hoe belangrijk dit proces was, het zogenaamde 'observatieel leren', ook wel 'modeling' genoemd. 'Modeling' houdt in dat mensen leren door anderen te observeren en een idee vormen van hoe bepaalde nieuwe gedragingen moeten worden uitgevoerd. Op latere tijdstippen dient deze gecodeerde informatie als een leidraad voor actie. Mensen leren iets zonder dat de gedragingen zelf uitgevoerd worden en zonder dat ze daarna gelijk worden beloond of worden afgestraft. Het voordeel van deze manier van leren is dat mensen iets leren van de ervaring van andere mensen, zonder dat individuen het wiel opnieuw moeten uitvinden wanneer iemand een nieuwe handeling verricht (Bandura, 1977).

Bandura voerde verschillende experimenten uit en constateerde dat er bepaalde stappen worden genomen in het proces van 'modeling':

1. **Aandacht.** Als mensen op het punt staan om iets nieuws te gaan leren en dus iets willen leren van de observatie moet er aandacht zijn voor de componenten van het gedrag dat gemodelleerd gaat worden. Verschillende factoren beïnvloeden de aandacht. Als een persoon moe, chagrijnig, onder invloed is van een drug, ziek, nerveus of juist hyperactief leert het individu minder goed (ook wel de sensorische capaciteiten van een persoon). Als andere stimuli in de ruimte de persoon afleiden leert men ook minder goed. Naast de eigenschappen van de persoon zelf kunnen de karakteristieken van het rolmodel de aandacht beïnvloeden. Als het rolmodel bijvoorbeeld aantrekkelijk, ambitieus en zeer competent is hebben de toehoorders meer aandacht voor degene die iets verteld. Ook de manier van interactie tussen het rolmodel en het individu kunnen de aandacht beïnvloeden.

2. **Retentie.** Ten tweede moet de persoon zich herinneren wat het rolmodel vertelde, het individu moet de informatie op kunnen halen uit het geheugen. Personen slaan symbolische beelden of verbale beschrijvingen van de vaardigheden die het model uitvoert op in het geheugen. Als men deze informatie opslaat, kan men het op een later tijdstip ophalen en reproduceren in het eigen gedrag.

3. **Reproductie van gedrag.** In dit stadium zet men de beelden (visueel) en beschrijvingen (verbaal) die de persoon heeft opgeslagen in het geheugen om in gedragingen. Het individu moet de mogelijkheid/vaardigheid bezitten om het gedrag te reproduceren. De reproductie verbetert wanneer de persoon de vaardigheid veel oefent, en ook wanneer men mentaal de vaardigheid al vooroefent zonder deze echt uit te voeren. Het individu oefent net zo lang zelf totdat er een match met het gedrag van het rolmodel ontstaat.

4. **Motivatie.** Daarnaast moet het individu de motivatie bezitten om het gedrag werkelijk te uit te voeren. Er moet een reden zijn om het gedrag van het rolmodel te gaan imiteren. (Boeree, z.j., Gibson, 2004).

Naast deze belangrijke bevinding constateerde Bandura ook dat mensen over een bepaalde zelf-controle beschikken, mensen kunnen zelf bepalen welke gedragingen ze wel kopiëren en welke niet. Deze zelf-controle wordt geoefend door cognitieve denkprocessen, men noemt dit proces ook wel 'perceived self-efficacy', dit 'perceived self-efficacy' verwijst naar, in hoeverre een persoon gelooft of hij in staat is om een bepaalde gedraging uit te voeren in een bepaalde situatie. Deze twee centrale thema's in het werk van Bandura: 'observatieel leren' en 'self efficacy', zijn samengebracht onder de sociaal-cognitieve leertheorie. De sociaal-cognitieve leertheorie gaat ervan uit dat de interactie tussen de omgeving en de innerlijke persoon, de individuele acties van de persoon bepalen. De resultaten van deze acties vormen uiteindelijk de toekomstige gedachten, gevoelens en het denken (Bandura, 1977). Ook in de studie tot arts is er sprake van 'observatieel leren' en 'self-efficacy'. De leerling observeert tijdens de co-assistentenschappen en bepaalt zelf welke vaardigheden van de begeleidende arts worden overgenomen en welke niet.

Harden en Crosby voerden in 2000 een onderzoek uit naar de rollen van een docent in het medisch onderwijs. In het figuur dat voortkwam uit de resultaten van het onderzoek wordt ook aandacht besteed aan de arts als rolmodel tijdens de dagelijkse patiëntenzorg en daarnaast aan de arts als rolmodel in de rol van docent. Harden en Crosby (2000) geven in het artikel aan dat medische studenten voor een groot deel worden beïnvloed door de docenten die de leerling tegenkomt tijdens de studie tot arts. Het beïnvloedt de medisch studenten in de keuzes die ze maken ten opzichte van de toekomstige carrière als arts, in de professionele houding die ze zelf aannemen en in de mate hoe belangrijk ze bepaalde vakken vinden. Verder wordt in de bevindingen van de literatuurstudie gemeld dat men wereldwijd erkend dat 'observatieel leren' een van de meest krachtige doceertechnieken is in de vorming, het coachen, doceren en ondersteunen van toekomstige artsen. Uit bovenstaande noten blijkt hoe belangrijk het is dat een arts zich bewust is van zijn rolmodel zowel als arts als docent.

## 2.5 Het creëren van een veilig leerklimaat in het medisch onderwijs

De Haan, Boendermaker en Heij (2002) beschrijven de basis van veiligheid tijdens het medisch onderwijs die noodzakelijk is voor een volledige ontwikkeling van de arts in opleiding. Voordat de co-assistent en de arts in opleiding zichzelf verder kunnen ontplooien moet men eerst voldoen aan de behoefte van veiligheid. Althouse, Stritter en Steiner (1999) voerden een studie uit naar invloedrijke rolmodellen in het medisch onderwijs. Medische studenten identificeerden de artsen die in hun ogen een grote invloed hadden gehad op de ontwikkeling tot arts. Deze rolmodellen werden geïnterviewd en deelden de mening dat artsen een leerklimaat dienen te scheppen waar men fouten tolereert en waarin men accepteert niet gelijk alles te weten, dit alles binnen de grenzen van het ontwikkelingsproces en het streven om uiteindelijk wel competent te zijn. Buchel en Edwards (2005) rapporteren over karakteristieken van de effectieve klinische docent. Artsen in opleiding geven in dit onderzoek aan dat een effectieve docent beschikbaar moet zijn voor zijn studenten en in staat moet zijn om een veilige, oordeelvrije en angstvrije leeromgeving te creëren. Sutkin, Wagner, Harris en Schiffer (2008) sluiten hierbij aan en melden dat een positieve relatie met studenten en een steunende leeromgeving van belang zijn in het onderwijs aan medische studenten. Een goede sfeer beïnvloedt het leren, dit verwijst niet alleen naar de fysieke omgeving en de methoden die men gebruikt, maar ook naar de persoonlijkheid van de docerende arts en het algemene klimaat van de instelling waar men onderwijs geeft.

## 2.6 Leren van volwassenen

Bolhuis (1988) geeft de volgende definitie voor het begrip 'leren': "het realiseren van een relatief duurzame wijziging in de handelingscompetentie van een persoon" (Bolhuis, 1988, p. 70). Knowles, Holton en Swanson (2005) nemen de volgende definitie over van Boyd (in Knowles, 2005, p.

10) 'leren is de activiteit of het proces waarbij gedragsverandering, kennis, vaardigheden en houdingen worden verworven' (Knowles et al., 2005, p.10). Volwassenen leren anders dan kinderen. Het belangrijkste verschil is dat het leren van volwassenen minder blanco is dan het leren van kinderen. "De mogelijkheden tot leren zijn in grote mate gevormd door eerder leren. Dat eerdere en het huidige leren worden sterk beïnvloed door vroegere en actuele sociale factoren" (Bolhuis, 1988, p. 75).

Thijssen verrichtte als onderdeel van een studie naar bijscholing van oudere werknemers een onderzoek naar een aantal grondprincipes van zes internationaal bekende auteurs op het gebied van het leren van volwassenen (Cramer & Van der Kamp, 1990). Uit dit onderzoek kwamen een zestal grondprincipes voort die belangrijk zijn bij het leren van volwassenen:

1. *Actieve verwerking*: met het toenemen van de leeftijd en levenservaring van een persoon is het belangrijk dat een volwassene nieuwe leerstof actief verwerkt, dit heeft een positieve invloed op het leerresultaat
2. *Systematische feedback*: bij het leren van volwassenen doet zich het gevaar voor dat volwassenen met bepaalde preconcepties kijken naar nieuwe leerstof daardoor kunnen vertekeningen en misverstanden ontstaan. Specifieke en concrete feedback tijdens het leerproces is zeer belangrijk
3. *Herkenbaar begrippenkader*: in dit stadium gaat het erom dat de in het verleden opgedane ervaringen van de volwassene worden gebruikt om de kern van de leerstof te plaatsen. Tijdens het verloop van de cursus moeten begrippen en ideeën met elkaar in verband worden gebracht en worden gerelateerd aan voorkennis
4. *Directe toepasbaarheid*: het is belangrijk dat het geleerde in de cursus relevant is voor de cursist, dat het na één of twee bijeenkomsten kan worden toegepast in de praktijk.
5. *Aangepaste sociale context*: Voor volwassenen met levenservaring is het van belang dat het leren van nieuwe vaardigheden of kennis plaats vindt vanuit een sociale, vriendelijke sfeer. In een training is het van

belang om gelijkwaardigheid na te streven zowel tussen de cursist en de docent als tussen de cursisten onderling.

6. *Aangepaste logistieke context*. Voor het leren van volwassenen is het tot slot belangrijk dat volwassenen het gevoel hebben dat zij de leersituatie aankunnen, dit houdt in dat de leersituatie te overzien is en te beheersen/beïnvloeden (programma-overzicht, flexibele planning) (Cramer & Van der Kamp, 1990).

## 2.7 Samenvatting hoofdstuk 2

In dit hoofdstuk werden de theorieën besproken die belangrijk zijn in het kader van dit onderzoek. *Wat betekenen deze theorieën nu voor dit onderzoek?* De theorieën werden al voor een groot deel met elkaar vergeleken en voor een deel in verband gebracht met de cursusinhoud. In deze afsluitende paragraaf worden de theorieën nogmaals in het licht van het onderzoek naast elkaar gelegd en zal bovenstaande vraag worden beantwoord.

De onderwijskundige modellen bevatten een aantal belangrijke punten die dienen te worden meegenomen bij de keuze of ontwikkeling van het meetinstrument. Het model van Gagné is in de eerste plaats bedoeld voor leraren in het basisonderwijs. De negen stappen van het model vormen samen een voorschrift voor de docent die lessen systematisch wil voorbereiden en uitvoeren. Het model kan een leidraad vormen voor de docent bij de uitvoering van het onderwijsmoment maar het model kan ook als leidraad dienen voor een ontwerper van lessen. De negen stappen laten een duidelijke opbouw zien van een onderwijsmoment. De eerste drie stappen vormen de voorbereiding en inleiding tot het onderwijsmoment, de vier stappen die daarop volgen vormen de eigenlijke instructie en de laatste twee stappen vormen de afsluiting van de les. Wat we kunnen leren van Gagné is dat er sprake moet zijn van een duidelijke opbouw in het onderwijsmoment. Ook het model van Van Gelder geeft aan dat er sprake moet zijn van een duidelijke opbouw in

het lesmoment. De didactische sleutelvragen verwijzen naar een aantal aspecten die relevant zijn bij het doceren. Deze aspecten kunnen fungeren als steunpunten voor de lesvoorbereiding, lesrealisatie en lesevaluatie. Wanneer de verschillende stappen in de twee modellen nader worden geanalyseerd kan het volgende worden opgemerkt. Het is belangrijk dat de leerdoelen van het onderwijsmoment worden vastgesteld en gecommuniceerd. Daarnaast is het van belang om aan te sluiten bij de voorkennis van de leerling. De modellen hebben allebei aandacht voor het aanbieden van de leerstof, op welke manier dit moet gebeuren en met welke hulpmiddelen. Tot slot is het belangrijk om het onderwijsmoment te evalueren.

Wanneer bovenstaande onderwijskundige modellen vervolgens worden vergeleken met de theorie over het *leren van volwassenen* kunnen de volgende opmerkingen worden geplaatst. *'Actieve verwerking'* heeft te maken met de organisatie van het onderwijsmoment. Laat ik de leerlingen een groepsopdracht uitvoeren? Laat ik leerlingen een presentatie houden? Gebruik ik instrumenten tijdens het onderwijsmoment en laat ik leerlingen demonstreren hoe je de instrumenten gebruikt? *Actieve verwerking* staat dus in verband met de organisatie van het onderwijsmoment. In bovenstaande onderwijskundige modellen wordt hier uitgebreid aandacht aan besteed. In het model van Gagné wordt genoemd dat er sprake moet zijn van *'feedback op de performance'* van een leerling. In de theorie van het volwassen leren komt dit terug als het geven van *'systematische feedback'*. Specifieke en concrete feedback is zeer belangrijk voor het leerproces. Het volgende punt wat belangrijk wordt gevonden bij het leren van volwassenen is het scheppen van een *'herkenbaar begrippenkader'*. Gagné en Van Gelder geven beiden aan dat het belangrijk is dat de docent op de hoogte is van de voorkennis van een leerling, de beginsituatie, en deze voorkennis ophaalt voordat de nieuwe stof wordt aangeboden.

De psychologische theorie die wordt betrokken in het theoretisch kader is gekozen omdat deze belangrijk is voor de setting waarin het leren van een arts in opleiding plaatsvindt. De sociaal-cognitieve leertheorie van Bandura

schetst het 'rolmodel leren'. Daarnaast worden nog enkele onderzoeken aangehaald die het belang schetsen van het scheppen van een veilige setting in de medische context. In de theorie van het 'volwassen leren' wordt het scheppen van een gelijkwaardige sfeer, 'aangepaste sociale context', ook als een van de belangrijke punten genoemd tijdens een onderwijsmoment met volwassenen.

Hieronder zullen bovengenoemde aandachtspunten uitgebreid worden besproken en wordt aangegeven waarom deze punten belangrijk zijn voor het leren van volwassenen in het medisch onderwijs.

***In een onderwijsmoment moet sprake zijn van een duidelijke opbouw (introductie onderwijsmoment, lesrealisatie, afsluiting)***

In een onderwijsmoment moet sprake zijn van een duidelijke opbouw. Een duidelijke opbouw is van belang voor de docent zelf, het vormt een leidraad bij het aanbieden van de leerstof en zorgt ervoor dat er geen essentiële stappen worden overgeslagen. Het hanteren van een duidelijke opbouw van het onderwijsmoment kan ook helpen bij het voorbereiden en organiseren van het onderwijsmoment. Glaser (1962) heeft een basismodel beschreven bestaande uit vier stappen: het bepalen van de onderwijsdoelstellingen, het bepalen van de beginsituatie van de lerenden, het realiseren van de onderwijsleersituatie, en het bepalen van de bereikte leerresultaten. Daaraan koppelt Glaser een feedbackmechanisme, waarbij de docent aan de hand van informatie over elk van de fasen kan besluiten tot bijstelling van één van de voorgaande fasen. Andere auteurs hebben meer specifieke instructiemodellen ontwikkeld, waarvan het model van Gagné en Briggs (1979) een van de meest gebruikte is. Zoals beschreven in paragraaf 2.2 onderscheiden zij negen 'instructional events', waarin met name de fasen van het aanbieden van instructie, en de terugkoppeling over het leereffect verder zijn uitgewerkt. Gagné en Briggs koppelen dit aan verschillende onderwijsdoelstellingen, en geven aan dat – afhankelijk van de doelstelling die nagestreefd wordt – de instructiestrategie zal verschillen. Belangrijk is echter, zo

benadrukken zij, dat deze onderling op elkaar zijn afgestemd.

Dat heeft vergaande implicaties voor de didactische kwaliteiten van docenten in het medisch onderwijs. Veelal is hier sprake van verschillende doelstellingen die gelijktijdig aan bod komen. Een voorbeeld hiervan is een onderwijssituatie waarin een lerende wordt geleerd hoe te hechten. Daarbij gaat het niet alleen om motorische vaardigheden, maar komen ook allerlei cognitieve doelstellingen aan de orde. Bovendien spelen daarbij attitudes eveneens een belangrijke rol, zoals de houding van de arts tegenover de patiënt. Aangezien het medisch onderwijs voor een belangrijk deel tijdens de dagelijkse patiëntenzorg plaatsvindt en dit gepaard gaat met een bepaalde tijdsdruk voor onderwijsmomenten is het van belang dat de arts van te voren duidelijk weet wat hij wil vertellen tijdens een onderwijsmoment. Met behulp van een duidelijke structuur in het lesmoment kan binnen zeer korte tijd een onderwijsmoment worden gegeven.

***Het is belangrijk dat de leerdoelen van het onderwijsmoment worden vastgesteld en worden gecommuniceerd naar de leerlingen toe***

Het is belangrijk dat voordat er met een onderwijsmoment wordt aangevangen, leerdoelen worden opgesteld en dat deze doelen ook duidelijk naar het publiek worden gecommuniceerd. Doelen dienen om richting te geven aan de inhoud van het onderwijsmoment en bij het structureren van de leerstof. Bepaald moet worden welke kennis moet worden overgebracht om de doelen te bereiken. De doelen geven richting aan de organisatie van het onderwijsmoment. Kyriacou (1991) geeft aan dat het stellen van doelen een essentieel onderdeel is van de planning van een les. Er zijn vier hoofdzaken in de voorbereiding van een les: (1) keuze maken over de onderwijsdoelen van de les, (2) selectie van werkvormen en schrijven van een plan voor de les (inclusief tijdsplanning en keuze hulpmiddelen), (3) eigenlijke voorbereiding van alle hulpmiddelen die gebruikt gaan worden (spullen klaarzetten, lokaal inrichten, weten hoe digitale hulpmiddelen bediend moeten worden), (4) beslissing over hoe je de les gaat evalueren en hoe de leerresultaten van de



leerlingen worden geëvalueerd. De docent wordt geacht de doelen te specificeren in termen van kennis, begrijpen van de leerstof, vaardigheden en houdingen. Ook Gagné geeft aan dat doelen specifiek moeten worden opgesteld in de vorm van gewenste eindgedragingen. Doelen bepalen hoe de les eruit zal zien. Afhankelijk van of de stof herkend moet worden, uit het hoofd geleerd moet worden of toegepast moet worden zal bepalen welke werkhouding van de leerlingen wordt gevraagd en welke werkvormen worden geselecteerd. Leerdoelen helpen de leerling bij wat er van hem wordt verwacht. Gagné en Van Gelder geven beide aan dat het opstellen van doelen belangrijk is om uiteindelijk goed te kunnen evalueren. Worden de vooraf gestelde eindgedragingen uiteindelijk bereikt? Doelen helpen dus bij het toets- en evaluatieproces. Het leerplanontwikkelingsmodel van Tyler bestaat uit drie componenten waarvan een wisselwerking uitgaat: 'objectives', 'learning experiences' en 'evaluation'. Het model beoogt inzicht te schetsen in het onderwijsstelsel van de docent zelf. Elk van de drie componenten zal invloed uitoefenen op de twee andere componenten. Het model moet docenten helpen om de drie genoemde elementen goed te beschrijven en de onderlinge verbanden van deze componenten steeds in ogenschouw te nemen. De onderwijsactiviteiten dienen zo te worden ingericht dat de doelstellingen worden beantwoord, dat er wordt toegewerkt naar de evaluatie die vervolgens weer is afgestemd op de eerder geformuleerde doelstellingen (Warries en Pieters, 1992). Voor docenten in het medisch onderwijs is het een eenvoudig model om kort op het eigen onderwijs te kunnen reflecteren. Vanwege het drukke programma van de arts is een vaak gehoord excuus voor het evalueren van onderwijs dat daar geen tijd voor is. Met het model van Tyler kan in korte tijd worden geanalyseerd wat de doelen van een onderwijsmoment zijn, welke activiteiten daarbij horen om de doelen te bereiken en of de doelen uiteindelijk worden bereikt. Wanneer het 'opstellen van doelen' in verband wordt gebracht met de theorie van het 'leren van volwassenen', dient bij het opstellen van doelen rekening te worden gehouden met de zogenaamde 'aangepaste logistieke context', wat wil de

volwassene leren in een training? De volwassene dient te worden betrokken bij het opstellen van de leerdoelen zodat deze het gevoel heeft de situatie aan te kunnen en de leersituatie te overzien is. Wanneer de leerdoelen met de docent worden besproken committeren de docent en de leerling zich aan het onderwijsprogramma (Kallenberg, Grijsparde, Ter Braak, Van Horzen, 2003). Door de leerdoelen krijgt de student ook een idee van wat er van de docent verwacht mag worden. Volwassenen kunnen bepalen wat voor hen op dat moment belangrijk is om te leren. Met welk doel kom ik naar deze training? Op deze manier heeft de volwassene het gevoel dat hij betrokken wordt bij de leersituatie. De volwassene kan ook bepalen of de stof op dat moment relevant is voor zijn opleiding en bepalen om het onderwijs op een ander moment te volgen. Leerdoelen zijn richtinggevend voor het eigen leerproces, de leerdoelen geven aan wat er van de leerling wordt verwacht. Bovendien is leren het meest succesvol wanneer studenten hun eigen doelen bepalen en nastreven (Kallenberg et al. , 2003). In het medisch onderwijs is het van groot belang dat goed wordt gecommuniceerd wat de doelen zijn van onderwijsmomenten. Omdat er gewerkt wordt tijdens de dagelijkse patiëntenzorg en er sprake is van patiëntcontact moet van te voren zo duidelijk mogelijk worden omschreven wat het doel is van het onderwijs zodat het niet ten koste gaat van de zorg aan de patiënt. Tijd is kostbaar in de medische wereld daarom dient een doel binnen het medisch onderwijs zo snel mogelijk bereikt te worden en dient er efficiënt te worden omgesprongen met tijd. Vaak moeten er in korte tijd verschillende doelen worden bereikt, niet alleen cognitieve doelen maar ook moeten er motorische vaardigheden worden opgedaan tijdens het leerproces. De implicaties voor de docerend arts zijn dat hij moet zorgen dat leerdoelen geprioriteerd worden en dat er een overwogen keuze gemaakt moet worden over de programmadoelen die uiteindelijk bereikt dienen te worden.

### ***Aansluiten bij de voorkennis van de leerling***

Aansluiten bij de voorkennis van een leerling is zeer belangrijk tijdens een onderwijsmoment. Er is sprake van

een cumulatief proces in het onderwijs. De student maakt gebruik van en bouwt voort op kennis en vaardigheden waarover hij al beschikt. Een volwassene gaat niet blanco een leersituatie in, kennis en handelingsstrategieën die de volwassene al eerder heeft opgedaan worden meegenomen in de oplossing van het probleem (Kallenberg et al.). Kallenberg (2003) maakt een onderscheid in de cognitieve- en affectieve beginsituatie van een leerling. De cognitieve beginsituatie verwijst naar de voorkennis en studievaardigheden van de leerling. De affectieve beginsituatie verwijst naar zelfvertrouwen en motivatie van de leerling. Gagné geeft aan dat voordat nieuwe informatie wordt aangeboden eerst de voorkennis moet worden opgehaald die nodig is om de nieuwe informatie te begrijpen. Het komt ook uitgebreid aan de orde in de Gagné/Briggs- theorie bij het onderdeel 'leercondities'. Er moet aan bepaalde startcondities zijn voldaan om het mogelijk te maken dat bepaalde uitkomsten optreden. Van Gelder beschrijft dit als het aansluiten bij de beginsituatie van leerlingen. Zijn de benodigde vaardigheden bij de leerling aanwezig om het lesdoel te kunnen bereiken? Er is een klein nuanceverschil op te merken. Gagné pleit voor herhaling van die voorkennis, dus daadwerkelijk actief voorkennis ophalen voordat gestart wordt met het aanbieden van nieuwe informatie. Van Gelder is daar iets minder specifiek in en houdt het op 'op de hoogte zijn van de beginsituatie van de leerling' en hierbij aansluiten. In de theorie van het leren door volwassenen wordt ook aandacht besteed aan de 'voorkennis van leerlingen'. Hier wordt het als 'herkenbaar begrippenkader' omschreven. In het verleden opgedane ervaringen van de volwassene moeten worden gebruikt om de kern van de leerstof te plaatsen. Begrippen en ideeën moeten met elkaar in verband worden gebracht en worden gerelateerd aan voorkennis. Een docent dient goed op de hoogte zijn van de beginsituatie van de leerling. In een groep heeft de docent te maken met niveaoverschillen tussen leerlingen. Ook hier dient rekening mee te worden gehouden tijdens een onderwijsmoment. Iedere leerling heeft een eigen hoeveelheid voorkennis, de aard van die voorkennis varieert ook per leerling. De hoeveelheid en aard van de voorkennis worden bepaald door de fase in de

opleiding, de gevolgde vakken, soort vooropleiding, de interesse en belangstelling voor het onderwerp, stage- en werkervaring, het land van herkomst en de cultuur (Kallenberg et al., 2003). Inschatten en toetsen van die voorkennis is dan ook een zeer belangrijk aspect waarbij bij de ontwikkeling van onderwijsmomenten rekening dient te worden gehouden. In het medisch onderwijs dienen in zeer korte tijd verschillende onderwijsdoelen te worden bereikt. Om te zorgen dat leerlingen in korte tijd de doelen bereiken is het goed om regelmatig de voorkennis te testen om te kijken of leerlingen vorderingen maken, of de kennis op pijl is en waar mogelijk aan te kunnen geven dat de voorkennis dient te worden uitgebreid om nieuwe doelen bereiken. Ook is het belangrijk om de affectieve beginsituatie na te gaan zoals deze door Kallenberg (2003) werd benoemd. Het zelfvertrouwen en de motivatie van de leerling kunnen voor een groot deel de cognitieve beginsituatie bepalen. In het medisch onderwijs kan een docent bepalen in welke fase van de opleiding de leerling zich bevindt, welke co-assistentschappen de leerling heeft doorlopen en op basis daarvan besluiten welke basiskennis dus aanwezig hoort te zijn. Aansluiten bij de voorkennis van een leerling heeft een positief effect op het leerresultaat dat uiteindelijk wordt bereikt. Het inschatten van de beginsituatie is makkelijker bij een aiOS dan bij een co-assistent. Co-assistenten zijn vaak voor een zeer korte periode in opleiding bij een arts. De docerend arts dient een strategie te ontwikkelen om in een korte tijd te bepalen wat de beginsituatie van een leerling is.

### ***Bepaling van de leerstof die wordt aangeboden en aan de hand daarvan bepaling van de organisatie van het onderwijsmoment***

Hiervoor is al genoemd dat de lesdoelen de inhoud van een onderwijsmoment voor een groot deel bepalen, het bepaalt welke leerstof wordt aangeboden, hoe het lesmoment eruit gaat zien en tot slot welke werkvormen en hulpmiddelen er gebruikt worden. De leerstof moet relevant zijn en dus nauw aansluiten bij de leerdoelen en de beginsituatie van de leerlingen. De twee aspecten behoren tot de voorbereiding van het onderwijsmoment volgens Kyriacou (1991). Er zijn vier hoofdzaken in de

voorbereiding van een les: (1) keuze maken over de onderwijsdoelen van de les, (2) selectie van werkvormen en schrijven van een plan voor de les (inclusief tijdsplanning en keuze hulpmiddelen), (3) eigenlijke voorbereiding van alle hulpmiddelen die gebruikt gaan worden (spullen klaarzetten, lokaal inrichten, weten hoe digitale hulpmiddelen bediend moeten worden), (4) beslissing over hoe je de les gaat evalueren en hoe de leerresultaten van de leerlingen worden geëvalueerd. Stap (2) en (3) beschrijven in feite de organisatie die voorafgaand aan het onderwijsmoment plaatsvindt.

Bij Gagné komen deze stappen terug in de volgende 'instructional events':

4. *De informatie (de leerstof) aanbieden*
5. *De leerlingen helpen bij het leren*
6. *Leerlingen het gevraagde gedrag laten tonen.*

Bij Van Gelder krijgt dit onderdeel vorm in de volgende didactische sleutelvragen:

- a. *Hoe dien ik de leerstof te kiezen en te ordenen?*
- b. *Welke leeractiviteiten dien ik de leerlingen toe te laten passen?*
- c. *Wat doe ik als leerkracht om de leerlingen te helpen?*
- d. *Welke onderwijs- en leermiddelen dien ik te gebruiken?*
- e. *Hoe organiseer ik de leerlingengroep (groeperingsvormen)?*

De nuances die hier kunnen worden aangebracht zijn dat Gagné de nadruk voornamelijk legt op hoe zorg ik dat de aangeboden informatie landt bij de leerling, hoe breng ik die informatie over? Bij Van Gelder ligt de nadruk meer op de organisatie van het onderwijsmoment. Welke activiteiten kies ik? Welke hulpmiddelen gebruik ik? Hoe groepeer ik mijn leerlingen? Van Gelder gaat dieper in op dit onderdeel. Omdat het bij het volwassen leren belangrijk is dat er sprake is van 'actieve verwerking', met het oog op het toenemen van de leeftijd en levenservaring van een persoon, is het belangrijk dat een volwassene nieuwe leerstof actief verwerkt. Dit heeft een positieve invloed op het leerresultaat. Wat betekent bovenstaande informatie nu voor het medisch onderwijs? Voor een docent in het medisch onderwijs is het van groot van belang om goed

over de leerstof na te denken en vervolgens over de organisatie van het onderwijsmoment, op die manier zal leerstof sneller en beter worden opgenomen. Ook 'directe toepasbaarheid' kan hier een rol spelen. Het is belangrijk dat het geleerde in de cursus relevant is voor de cursist, dat het na één of twee bijeenkomsten kan worden toegepast in de praktijk. Een goede voorbereiding is het halve werk. Bovendien is er tijdens de dagelijkse patiëntenzorg vaak weinig tijd voor het geven van onderwijs daarom dient de beschikbare tijd voor onderwijs op een goede manier te worden overdacht en te worden ingericht.

### ***Evaluatie van het onderwijsmoment***

De evaluatie van onderwijs is van groot belang voor het bepalen van de leerresultaten van de leerlingen. Daarnaast vormt de evaluatie van het onderwijsmoment vaak het beginpunt voor het volgende onderwijsmoment. Kyriacou (1991) noemt zes doelen voor het evalueren van een onderwijsmoment: (1) feedback over handelen docent (zijn er onduidelijkheden bij leerlingen en dienen deze in een volgend onderwijsmoment op een andere manier te worden uitgelegd?), (2) leerlingen van educatieve feedback voorzien (welke eisen worden gesteld en op welke manier bereik ik die eisen?), (3) om leerlingen te motiveren, (4) regelmatige feedback om vorderingen over een langere periode te bepalen, (5) beoordeling van de prestaties op een van te voren bepaald moment, bijvoorbeeld een toets aan het einde van een lesblok (6) om te bepalen of leerlingen klaar zijn voor de volgende leeractiviteit. Over het algemeen kan gesteld worden dat het evalueren van onderwijs als algemeen doel heeft het verbeteren van de kwaliteit van het onderwijs. Gagné omschrijft het evalueren van het onderwijsmoment als: 'het eindresultaat vaststellen van het leren'. Is het leerdoel bereikt en kan met zekerheid vastgesteld worden dat het gewenste eindgedrag vertoond kan worden? Bij Van Gelder wordt het grotendeels op dezelfde manier geïnterpreteerd. Wat zijn de resultaten van het onderwijsmoment en hoeverre zijn de lesdoelen bereikt. Qua inhoud zijn er geen verschillen op te merken op dit aandachtspunt. Opvallend is dat de twee modellen niet aangeven dat het de uitkomsten van de evaluatie

implicaties kunnen hebben voor de invulling van een volgend onderwijsmoment zoals Kyriacou (1991) dit wel noemt in zijn doelen. Binnen de beschrijving over het '*leren van volwassenen*' in voorgaand hoofdstuk wordt niet expliciet aandacht besteed aan het evalueren in het volwassen leren. Hiervoor werd aangehaald dat het belangrijk is om de volwassene te betrekken bij het leerproces en het opstellen van eigen doelen. Natuurlijk is het belangrijk om aan het einde van bijvoorbeeld een onderwijsmoment tijdens de dagelijkse patiëntenzorg vast te stellen of de lesdoelen of persoonlijke leerdoelen zijn behaald. Is het gelukt om alle onderdelen van de anamnese zelfstandig af te nemen en ben ik in staat om het volgende keer helemaal zonder begeleiding uit te voeren? Een evaluatiemoment is nodig om de vorderingen van een leerling vast te leggen, om te toetsen en aan de hand daarvan weer nieuwe doelen op te kunnen stellen voor het volgende onderwijsmoment. De doelen die door Kyriacou (1991) werden gesteld kunnen een leidraad vormen voor de evaluatie tijdens het medisch onderwijs. Het evalueren van het didactisch handelen van de docent vormt in dit onderzoek een belangrijke factor. Hoe reflecteer ik op mijn didactisch handelen en hoe pas ik mijn didactisch handelen tijdens een volgend onderwijsmoment aan zodat de onderwijsdoelen en daarmee de leerresultaten worden bereikt?

### ***Het geven van systematische feedback***

Specifieke en concrete feedback is van zeer groot belang voor het leerproces. Opbouwende feedback zorgt ervoor dat een leerling gemotiveerd blijft en daarnaast zorgt het voor verdere progressie van de leerontwikkeling. Door feedback te geven kunnen problemen worden geïdentificeerd, goed werk wordt beloond en zorgt voor een grotere motivatie, voor problemen in de ontwikkeling wordt een oplossing aangereikt en aan de probleemoplossing kan in praktijk verder worden gewerkt. Daarnaast laat een docent op die manier zien dat het leerproces zorgvuldig wordt gevolgd en ook voor de docent zelf belangrijk is (Kyriacou, 1991). Op die manier wordt ook bijgedragen aan het scheppen van een veilig leerklimaat waar later in deze samenvatting verder op in

wordt gegaan. Gagné heeft in zijn model aandacht voor het geven van feedback. Van Gelder laat dit niet expliciet terugkomen in zijn model. Feedback kan bij het model van Van Gelder impliciet onder het onderdeel: *Wat doe ik als leerkracht om de leerlingen te helpen?* kunnen vallen. De docent kan de leerling motiveren door feedback te geven in de trant van: *'je bent op de goede weg als je die punten nog aanpast is het een zeer goed werkstuk'*. In de theorie van het volwassen leren wordt het als een van de aandachtspunten genoemd, '*systematische feedback*'. Specifieke en concrete feedback tijdens het leerproces van de volwassene is zeer belangrijk. Bij het leren van volwassenen ligt het gevaar bij het feit dat volwassenen met bepaalde preconcepties kijken naar nieuwe leerstof daardoor kunnen vertekeningen en misverstanden ontstaan. De volwassene heeft meer voorkennis dan een kind. Feedback kan ook dienen om bepaalde aangeleerde gedragingen te doorbreken. Voor de medische context is het van belang dat feedback opbouwend is, dit met het oog op het scheppen van een veilig leerklimaat. Feedback dient daarnaast concreet te zijn zodat de leerling gelijk met de feedback in praktijk aan de slag kan. In het medisch onderwijs is er sprake van onderwijs tijdens de dagelijkse patiëntenzorg, de leerling moet dus handvatten krijgen om feedback op het handelen gelijk toe te kunnen passen op de werkvloer. In een medische setting is het van belang dat er regelmatig feedback wordt gegeven op het functioneren van een arts in opleiding. In het medisch onderwijs kunnen fouten cruciale gevolgen hebben.

### ***Rolmodel leren***

Het zogenaamde model leren vormt een bewust en onbewust proces in de ontwikkeling van leerlingen. Vaak nemen we tijdens het onderwijs onbewust gedrag over van docenten en begeleiders. Daarentegen wordt gedrag soms juist heel bewust gekopieerd. Door de leerling wordt de keuze gemaakt om het gedrag over te nemen van een positief rolmodel maar ook bewust om gedrag niet over te nemen van een negatief rolmodel. Bandura beschrijft de theorie van het zogenaamde rolmodel leren. Deze theorie van Bandura wordt gebruikt in dit onderzoek omdat vanuit de literatuur is gebleken dat er in het medisch onderwijs

vaak sprake is van rolmodel leren. In paragraaf 2.4 werden al enkele auteurs genoemd die het rolmodel leren in het medisch onderwijs bestudeerden. Vooral tijdens de opleiding van de arts zal door de leerling worden geconstateerd dat hij bepaalde artsen wel copieert en andere niet ('perceived self-efficacy'). De sociaal-cognitieve leertheorie gaat ervan uit dat de interactie tussen de omgeving en de innerlijke persoon, de individuele acties van de persoon bepalen. De leerling observeert tijdens de co-assistentieschappen en bepaald zelf welke vaardigheden van de begeleidend arts worden overgenomen en welke niet. De arts dient zich daarom constant bewust te zijn van zijn rolmodelfunctie zowel als arts als docent.

### ***Het scheppen van een veilige setting***

Het scheppen van een veilige setting, ook wel veilig leerklimaat genoemd, is van belang voor het bevorderen van het leerproces van de leerling. Het leerklimaat dat door de docent wordt geschapen kan een grote invloed hebben op de motivatie van de leerlingen en de houding ten opzichte van de leerstof die op dat moment wordt aangeboden. Het type leerklimaat dat het leren het beste faciliteert is volgens Kyriacou (1991) een docent die doelbewust lesgeeft, taakgeoriënteerd te werk gaat, ontspannen is, beschikt over een warme persoonlijkheid, ondersteunend en iemand die orde kan houden tijdens het onderwijs. In essentie komt het neer op het scheppen en onderhouden van positieve houdingen en motivatie van de leerlingen ten opzichte van het onderwijsmoment. Binnen het medisch onderwijs moet eerst worden voldaan aan het creëren van een veilige setting om de arts in opleiding tot volledige ontwikkeling te laten komen. Ook in het volwassen leren wordt aandacht besteed aan het creëren van een veilige setting tijdens het onderwijs, hier spreekt men van de '*aangepaste sociale context*'. Voor volwassenen met levenservaring is het van belang dat het leren van nieuwe vaardigheden of kennis gebeurt vanuit een sociale, vriendelijke sfeer. In een training is het van belang om gelijkwaardigheid na te streven zowel tussen de cursist als docent en tussen de cursisten onderling. Voor het medisch onderwijs heeft dit grote gevolgen. De docerend arts dient

niet alleen een veilige sfeer te scheppen voor de patiënt in kwestie maar ook een veilige sfeer voor de leerling die hij begeleidt tijdens de zorg aan de patiënt. De leerling dient op een goede manier begeleidt te worden en in een veilige setting de opleiding te genieten en zich op deze manier te ontwikkelen tot een volwaardig arts die is voorbereid op het nemen van grote beslissingen.

Het conceptueel kader dat aan het einde van dit hoofdstuk werd geschetst zal de rode draad vormen voor dit onderzoek. Op basis van de aandachtspunten die voortvloeiden uit de theorieën die werden besproken zullen beslissingen worden genomen in de volgende fasen van dit onderzoek. In hoofdstuk 2 werd het theoretisch kader voor het onderzoek beschreven daar vloeide een conceptueel kader uit voort. In hoofdstuk 3 zullen de onderzoeksmethoden van deze masterthesis worden beschreven.



### 3 Onderzoeksmethoden

#### 3.1 Inleiding

Het doel van deze studie is tweeledig. De studie heeft allereerst tot doel instrumenten te identificeren die de 'opleiderskwaliteiten van artsen' meten en daarmee mogelijk in aanmerking kunnen komen om het effect van het 'Teach the Teacher' programma module (1) te bepalen. Het tweede doel richt zich op het ontwikkelen van een betrouwbaar en valide instrument om het effect van de eerste module van het 'Teach the Teacher' programma te meten. In dit hoofdstuk wordt beschreven welke onderzoeksmethoden er gebruikt zijn om deze twee vraagstellingen te beantwoorden. In paragraaf 3.2 en 3.3 zal per vraagstelling een uiteenzetting worden gegeven van de methode van dataverzameling en de onderzoeksmethoden die zijn toegepast.

#### 3.2 Onderzoeksmethoden met betrekking tot vraagstelling 1

Om na te gaan in hoeverre er bestaande instrumenten zijn die gebruikt kunnen worden om het effect van didactische basisvaardigheidstrainingen in het medisch onderwijs te bepalen, is gebruik gemaakt van literatuuronderzoek in verschillende elektronische bestanden van de Rijksuniversiteit Groningen (ERIC, PiCarta, Medline en PubMed). Tijdens het onderzoek zijn experts op het gebied van het medisch onderwijs geraadpleegd om de literatuurstudie met extra informatie aan te vullen. In de eerste periode van het onderzoek heeft een exploratief literatuuronderzoek plaatsgevonden naar meetinstrumenten die de didactische vaardigheden van een docent meten. In dit oriënterende onderzoek werd duidelijk welke zoektermen er gehanteerd dienden te worden voor de uitvoering van het literatuuronderzoek. Op

basis hiervan zijn de volgende *criteria* gesteld aan de artikelen die uiteindelijk in de analyse werden opgenomen:

- het instrument moest van toepassing zijn op het medisch onderwijs of universitair onderwijs;
- er moest een voorbeeld aanwezig zijn van de items die in de lijst zijn opgenomen;
- er moest worden gerapporteerd over de betrouwbaarheid en/of de validiteit van het instrument.

Gezocht is in de volgende elektronische databestanden: ERIC, PiCarta, Medline (via EBSCOhost) en PubMed. In PiCarta en Medline is gezocht op de volgende sleutelwoorden: 'evaluation clinical teaching', 'instrument evaluation teaching', 'evaluation clinical faculty', 'standardized instruments evaluation teaching', 'instrument evaluation clinical teaching', 'instruments designed for evaluation clinical teaching', 'instruments designed for evaluation teaching', 'instrument evaluation teaching higher education', 'validation instrument for teaching quality', 'reliability instrument evaluation teaching', 'validity instrument evaluation teaching', 'evalueren docenten', 'evalueren van medisch docenten', 'evalueren van medisch onderwijs', 'evalueren van docenten in het hoger onderwijs', 'evalueren van docenten', 'evalueren didactische vaardigheden van docenten'.

De zoekactie in *Picarta* leverde negen artikelen op die van toepassing waren op het medisch- of universitair onderwijs. Uiteindelijk bleven er vier artikelen over die geschikt waren voor het doel van deze studie. In *Medline* (via *EBSCOhost*) werd op dezelfde zoektermen gezocht als in *Picarta*, hetgeen zeven verwijzingen opleverde. Veel artikelen stamden uit de zeventiger jaren en waren niet meer te traceren in de databases. Uiteindelijk bleef er één artikel over dat echter onvoldoende voldeed aan de gestelde psychometrische eisen, en om die reden niet in de verdere analyse is opgenomen.

In *ERIC* werden de eerder gebruikte zoektermen versmald, en werd de zoekactie beperkt tot publicaties verschenen in de periode van 1966-2008. Er is op de volgende zoektermen gezocht: 'evaluation instrument clinical

*teaching*, *'questionnaire evaluation clinical teaching course*, *'instrument clinical teaching effectiveness*, *'standardized instruments evaluation clinical teaching*, *'validation instrument for teaching quality*, *'reliability instrument evaluation teaching*, *'reliability instrument evaluation clinical teaching*'. Uit de resultaten kwamen uiteindelijk vijf artikelen voort die betrekking hadden op het universitair- of medisch onderwijs, maar vier van deze artikelen bleken niet via het digitale systeem te traceren. Uiteindelijk werd er één artikel opgenomen in de literatuurstudie.

In *Pubmed* werden de zoektermen nog verder versmald, waarbij de zoekactie werd beperkt tot publicaties in de volgende categorieën: Humans, Evaluation studies, English, Dutch, all adult: 19+ years. In de tijdspanne 1990 tot heden werd op de volgende zoektermen gezocht: *'questionnaire evaluation of teachers*, *'evaluation of teaching physicians*, *'evaluation didactic skills*, *'evaluation medical teachers*, *'evaluation of clinical teachers*'. Hieruit kwamen geen verwijzingen voort die betrekking hadden op het universitair- of medisch onderwijs, en voldeden aan de gestelde psychometrische criteria.

Vanuit de vier databases zijn uiteindelijk vijf artikelen geselecteerd voor verdere analyse. Door toepassing van de 'sneeuwbalmethode' en het raadplegen van experts op het gebied van het medisch onderwijs zijn daarnaast nog zes instrumenten gevonden die aan de eisen voldeden. Deze instrumenten zijn beoordeeld aan de hand van de algemeen didactische principes (zie hoofdstuk 2). Daarbij zijn ook observaties betrokken van de cursus, met het doel te bepalen of alle relevante didactische principes in het 'Teach the Teacher' programma module (1) in voldoende mate in het instrumentarium waren vertegenwoordigd.

### 3.3 Onderzoeksmethoden met betrekking tot vraagstelling 2

Om na te gaan of het mogelijk is om een betrouwbaar en valide instrument te ontwikkelen voor het bepalen van het

effect van een training in didactische basisprincipes aan medisch specialisten, is aan de hand van algemeen didactische principes bepaald waarop het instrument betrekking zou moeten hebben nagegaan in hoeverre bestaande schalen en items gebruikt konden worden voor het te ontwikkelen instrument. Daarbij werden ook twee observaties betrokken van het 'Teach the Teacher' programma module (1), om te bepalen wat de medisch specialisten leren in de cursus. Met behulp van deze gegevens werden gedragingen opgesteld die werden omgevormd tot items. Op basis van bovenstaande informatie werd de vragenlijst ontwikkeld.

Voor de dataverzameling met betrekking tot de betrouwbaarheid en validiteit van het instrument werden er medisch specialisten benaderd die deelnamen aan de cursussen in Amsterdam en Emmen. De medisch specialisten is gevraagd om de 'leerlingen' (dat wil zeggen de aios en co-assistenten die zij op de werkplek opleiden), enthousiast te maken voor het onderzoek en de aios de vragenlijst te laten invullen. Omdat deze vorm van dataverzameling weinig respons opleverde is in een later stadium besloten om nog een tweede vorm van dataverzameling te starten. In de tweede vorm van dataverzameling zijn alle aios die binnen het UMCG werkzaam zijn benaderd om de lijst in te vullen.

	Amsterdam		Emmen	
	med.spec.	aios/co	med.spec.	aios/co
<b>Benaderd</b>	11	13	9	4
<b>Respons</b>	6	4	2	2
<b>Totale respons</b>	<b>medisch specialisten: 8</b>		<b>aios/co-assistenten: 6</b>	

Tabel 3.1: Aantal respondenten dataverzameling 1

Respondenten aios UMCG	
Benaderd	Respons
338	38

Tabel 3.2: Aantal respondenten dataverzameling 2



### 3.3.1 Onderzoeksmethode

Deze paragraaf bespreekt aan de hand van verschillende onderwerpen het proces met betrekking tot het bepalen van de betrouwbaarheid en validiteit. Naar aanleiding van de resultaten van vraagstelling (1) werd een keuze gemaakt voor de vorm van het instrument en werd een selectie gemaakt van items, op basis van het conceptueel kader in hoofdstuk (2), die richtinggevend konden zijn bij de formulering van de items. Toen duidelijk was aan welke eisen het instrument moest voldoen en een keuze was gemaakt voor het soort instrumentarium werd gestart met de ontwikkeling van het meetinstrument.

### 3.3.2 Validiteit

Bij de ontwikkeling van een nieuw meetinstrument is het belangrijk om uitspraken te kunnen doen over de validiteit van het instrument. In de literatuur wordt ook wel gesproken van de geldigheid van een meting. Segers (2002) spreekt van drie verschillende vormen van validiteit:

- inhoudsvaliditeit (face-validiteit)
- criteriumvaliditeit (predictieve validiteit)
- constructvaliditeit (begripsvaliditeit)

In deze studie richtte de analyse zich op de inhouds- en constructvaliditeit. In eerste instantie zou ook de criteriumvaliditeit worden bepaald. In de eerste onderzoeksofzet zou een nulmeting worden verricht voorafgaande aan de cursus. Uit praktisch oogpunt bleek dit in een later onderzoeksstadium niet haalbaar. De nulmeting kon niet plaatsvinden, zodoende konden er geen uitspraken worden gedaan over de predictieve validiteit.

#### *Inhoudsvaliditeit*

In dit onderzoek werd de inhoudsvaliditeit bepaald door analyses op verschillende gebieden. De eerste analyse richtte zich op de cursusinhoud van het 'Teach the Teacher' programma. Met behulp van observaties tijdens de cursus werd een goed beeld gevormd van de aangeboden stof. In de eerste fase van het onderzoek werd twee keer geobserveerd tijdens module (1) van het 'Teach the Teacher' programma. De observaties tijdens de cursus

vormden de basis voor de ontwikkeling van het meetinstrument. Belangrijk was ook de deelname aan twee cursussen, die door verschillende cursusleiders werden verzorgd. Op deze manier werd bepaald of de inhoud van de cursus elke keer dat deze wordt gegeven dezelfde inhoud heeft. Naast de observaties werd ook het lesplan van de cursus onder de loep genomen. Na het volgen van de cursus werd een lijst ontwikkeld van geobserveerde vaardigheden. De geobserveerde vaardigheden werden omgezet in stellingen voor de vragenlijst. De items die in bestaande instrumenten werden geselecteerd op basis van het conceptuele kader fungeerden als leidraad bij de ontwikkeling en formulering van de items. De inhoudsvaliditeit van de vragenlijst werd verder bepaald door het consulteren van twee docenten binnen het 'Teach the Teacher' programma. In eerste instantie werd vooral gekeken of de items de inhoud van de cursus dekten. Daarna werd ook gediscussieerd over de formulering van items. Na elk gesprek volgden aanpassingen in de lijst. De vragenlijst is beoordeeld door enkele medestudenten van het Ontwikkelplatform met het oog op de formulering van items en onduidelijkheid over woordkeuze. Het oordeel van de vier medestudenten was dat de items voldoende duidelijk en helder waren. Met bovenstaande analyses werd de inhoudsvaliditeit vastgesteld. Na goedkeuring van de items werd de vragenlijst vervolgens digitaal gemaakt. De digitale lijst is opgenomen in bijlage (4).

#### *Constructvaliditeit*

Ter bepaling van de constructvaliditeit is met behulp van een factoranalyse de factorstructuur van de vragenlijst vastgesteld. Daarnaast werd met behulp van een correlatieanalyse de samenhang tussen de items bepaald. Voordat de factoranalyse werd uitgevoerd zijn een aantal vooranalyses uitgevoerd. Deze analyses worden in de volgende paragraaf besproken.

#### *Analyse voorafgaand aan de factoranalyse*

Voordat met de uitvoering van de factoranalyse werd gestart, werd onderzocht of de gegevens normaal zijn verdeeld. Met behulp van de beschrijvende statistieken in

het programma Statistical Package for Social Sciences (SPSS) werd per vraag geanalyseerd of de antwoorden voldoende gespreid waren. Uit de analyse blijkt dat er geen grote uitschieters zijn. Met het oog op de kleine steekproefgrootte is de spreiding van de antwoorden acceptabel voor dit onderzoek. In totaal werden er 44 vragenlijsten ingevuld en teruggestuurd. Geconstateerd kan worden dat de specialismen die geëvalueerd werden over het algemeen vrij evenredig verdeeld waren zodat we kunnen concluderen dat de lijst inzetbaar is op verschillende afdelingen. Daarnaast zijn er nog een aantal voorwaarden waaraan moet worden voldaan voordat met een factoranalyse mag worden gestart. De variabelen moeten in ieder geval op interval niveau zijn gemeten, dit is in deze studie het geval. Tot slot dient de onderzoeksgroep niet te homogeen te zijn. Over het algemeen zijn de respondenten afkomstig uit verschillende specialismen, de respondenten waren afkomstig uit 21 verschillende specialismen.

Met het uitvoeren van een factoranalyse met Varimax rotatie met een Kaiser's criterium van 1,00 werd bepaald uit welke constructen het didactisch handelen van medische specialisten is opgebouwd. Al snel bleek dat het niet mogelijk was om een normale factoranalyse uit te voeren vanwege het feit dat er minder respondenten (44) zijn dan items in de vragenlijst (50). Daarnaast is er sprake van het feit dat in de ongeroteerde factor analyse de 'missing values' in paren worden verwijderd, 'pairwise deletion', dit is de tweede reden dat de de factoranalyse niet kon worden uitgevoerd. De optie 'kan ik niet beoordelen' is in deze analyse vervangen door 'missing value'. Er werden geen resultaten gegenereerd. Vervolgens werd besloten om met behulp van gemiddelden de 'missing values' te compenseren ('replace with mean'). Voor elke 'missing value' wordt dan het gemiddelde op het betreffende item berekend en het gemiddelde wordt ingevuld op de lege plaatsen. In bijlage (5) is een tabel toegevoegd waarin de 'missing values' per vraag worden genoemd.

Omdat er veel 'missing values' werden geconstateerd in deze tabel werden de items en respondenten verder

geanalyseerd. Bij meer dan zes missende waarden wordt besloten om de items te verwijderen.

Vervolgens werd geanalyseerd of er ook sprake was van veel 'missing values' onder de respondenten. Respondenten die meer dan 15 'missing values' hadden werden uit het databestand verwijderd. Mogelijk zouden de vele missende waarden lijden tot een vertekend beeld in de data. Er werd een nieuw databestand zonder deze respondenten en de factoranalyse werd uitgevoerd.

Met de overige gegevens werden de eerste analyses uitgevoerd. Hoofdstuk 4 bespreekt de resultaten van de vooranalyse. Op basis hiervan werd besloten opnieuw een aantal items te verwijderen; ze vertoonden ambiguïteit.

Tot slot werd in de voorbereidende fase een betrouwbaarheidsanalyse uitgevoerd om te kijken in verband met eventueel te verwijderen items die de Cronbach's alpha voor de schalen voor een groot deel beïnvloeden.

Er is een principale componentenanalyse met Varimax uitgevoerd, waarbij het aantal factoren is gebaseerd op Kaiser's-criterium voor de Eigenwaarde van 1.00 (vgl. Kaiser, 1960 in: Bernaards, 2000). Het artikel McGrath et al. (2005) hanteert het criterium dat items worden geselecteerd wanneer deze een factorlading hebben van 0,40 of meer. Dit is als criterium gehanteerd om te bepalen of items aan een bepaalde factor worden toegeschreven.

### 3.3.3 Betrouwbaarheid

Een instrument wordt betrouwbaar geacht wanneer het dezelfde uitkomsten genereert bij herhaalde metingen onder dezelfde condities (Maslowski, 2001; Robson, 2007). Segers (2002) noemt verschillende methoden om de instrument-betrouwbaarheid te onderzoeken:

- test-hertest betrouwbaarheid
- alternatieve vormen van test-hertest betrouwbaarheid
- split-half betrouwbaarheid

- berekenen van de homogeniteitscoëfficiënt alpha (Cronbach's coëfficiënt alpha)

In deze studie werd gekozen om de homogeniteit van de schalen van het instrument te bepalen. Deze meting is alleen toepasbaar op samengestelde variabelen, er moet dus sprake zijn van meerdere indicatoren, bijvoorbeeld items (Segers, 2002). Omdat in dit onderzoek sprake is van een gestandaardiseerde meting richt het onderzoek op het bepalen van de interne consistentie en homogeniteit van de ontwikkelde schalen.

De interne consistentie binnen de schalen werd berekend door een betrouwbaarheidsanalyse uit te voeren met behulp van Cronbach's alpha. Nunnally (1978) adviseert een alpha waarde van 0,70 als de acceptabele grens voor Cronbach's alpha (Nunnally, 1978, in: DeVellis, 2003).

### 3.4 Samenvatting hoofdstuk 3

In dit hoofdstuk werden de methoden van onderzoek voor deze masterthesis beschreven. De validiteit en betrouwbaarheid van het te ontwikkelen instrument werden met behulp van de voorgaande onderzoeksmethoden vastgesteld. Deze studie dient een tweeledig doel. Met behulp van een uitgebreide literatuurstudie werd bepaald of er instrumenten aanwezig zijn die het effect van het 'Teach the Teacher' programma module (1) kunnen meten. De instrumenten die geschikt bleken voor opname in de literatuurstudie werden verder geanalyseerd aan de hand van het conceptuele kader dat in hoofdstuk 2 werd opgesteld. Vanuit deze analyse bleek dat er items geschikt waren bij het richting geven aan de ontwikkeling van het instrument. Uiteindelijk werd bepaald of er een betrouwbaar en valide instrument kon worden ontwikkeld om het effect van het 'Teach the Teacher' programma module (1) te kunnen bepalen. Het eerste stadium van het onderzoek richtte zich op de ontwikkeling van het instrument en het bepalen van de inhoudsvaliditeit. Aan de hand van observaties, de bestudering van het cursusmateriaal en de items die voortkwamen uit de literatuurstudie werd de inhoudsvaliditeit vastgesteld en

het instrument ontwikkeld. Het tweede deel van het valideringsonderzoek richtte zich op het bepalen van de constructvaliditeit. De constructvaliditeit werd bepaald met behulp van een factor- en correlatieanalyse. De homogeniteit van de schalen werd bepaald aan de hand van Cronbach's alpha. In het volgende hoofdstuk worden de onderzoeksresultaten van dit onderzoek besproken.



## 4 Onderzoeksresultaten

### 4.1 Inleiding

Het doel van deze studie is de ontwikkeling van een meetinstrument om het effect van het 'Teach the Teacher' programma module (1) te meten. Tijdens de constructiefase werden items opgesteld op basis van de observaties tijdens de cursus en de literatuurstudie die werd verricht naar aanleiding van vraagstelling (1). Op basis van het conceptuele kader dat hoofdstuk 2 beschreef, bleek dat er geen geschikte instrumenten bestaan die volledig zijn te benutten voor het meten van het effect van het 'Teach the Teacher' programma module (1). Voor de uiteindelijke analyse werden elf instrumenten geschikt bevonden die op basis van drie criteria (zie paragraaf 4.2) geschikt bleken voor opname in de literatuurstudie. Paragraaf 4.2 beschrijft en bediscussieert de instrumenten uitgebreid. De bestaande items die op basis van het conceptuele kader werden geselecteerd vormden een gids bij de ontwikkeling van het instrument, voornamelijk voor de formulering van de nieuwe items. Het beginstadium van het proces bestond grotendeels uit de ontwikkeling van het instrument en het bepalen van de inhoudsvaliditeit. Het laatste stadium bestond voornamelijk uit het bepalen van de constructvaliditeit en de verfijning van het instrument. Paragraaf 4.3 beschrijft het valideringsproces uitgebreid, waardoor uiteindelijk de algemene onderzoeksvraag kan worden beantwoord.

### 4.2 Onderzoeksresultaten vraagstelling 1

In het eerste stadium van het onderzoek diende eerst te worden bepaald of er bestaande instrumenten aanwezig waren voor het bepalen van het effect van het 'Teach the

Teacher' programma module (1). Er diende een antwoord te worden gegeven op de volgende vraagstelling:

*In hoeverre zijn er instrumenten aanwezig om het effect van didactische basisvaardigheidstrainingen in het medisch onderwijs te bepalen en welke schalen uit deze instrumenten kunnen worden gebruikt voor het meten van het 'Teach the Teacher' programma module (1)?*

Op basis van drie criteria, die hierna worden genoemd, werden elf instrumenten geschikt bevonden voor opname in de literatuurstudie. Opvallend is dat vier van de elf geselecteerde instrumenten zijn gebaseerd op het SFDP (Stanford Faculty Development Program) - raamwerk. Deze instrumenten vertoonden dan ook veel overlap en vaak werden dezelfde items geselecteerd zodat de selectie van items niet veel nieuwe informatie opleverde. Eén van de instrumenten was afkomstig uit een andere sector, vanuit de Tandheelkunde. Ook dit instrument bleek uiteindelijk gebaseerd te zijn op het SFDP-raamwerk. Ook uit het onderzoek in de onderwijskundige databases kwamen geen nieuwe instrumenten voort die gebruikt konden worden voor deze specifieke medische context. Ook juist vanwege deze specifieke context, en de manier van onderwijs geven, tijdens de dagelijkse patiëntenzorg, zijn de instrumenten uit het algemene onderwijs vaak niet geschikt voor een meting in het medisch onderwijs. In paragraaf 4.2 zal vraagstelling 1 worden beantwoord. De instrumenten worden bekritiseerd en uiteindelijk wordt besproken welke items op basis van het conceptueel kader geschikt zijn als richtinggevend bij de te ontwikkelen lijst. Aan het einde van de resultatensectie van vraagstelling 1 wordt in schema weergegeven welke items uiteindelijk zijn geselecteerd. De paragraaf wordt afgesloten met een aantal aandachtspunten, deze zijn gegenereerd vanuit de analyse van de instrumenten. Met deze aandachtspunten wordt rekening gehouden bij de ontwikkeling van de nieuwe lijst. Daarnaast heeft er een uitgebreide analyse plaatsgevonden naar de psychometrische kwaliteiten van de geselecteerde

elf instrumenten. De resultaten van deze analyse zijn in bijlage (6) opgenomen.

Hieronder zullen de resultaten van het literatuuronderzoek worden beschreven. De literatuurstudie is met steeds nauwere zoektermen verricht. Gezocht werd naar instrumenten die de 'opleiderskwaliteiten van artsen' meten. De instrumenten moesten aan de volgende drie criteria voldoen om in de uiteindelijke analyse te worden opgenomen:

- *Het instrument moet van toepassing zijn op het medisch- of universitair onderwijs*
- *Het instrument moet een voorbeeld bevatten van de items die in de lijst zijn opgenomen zodat bepaald kan worden welke schalen/items bruikbaar zijn voor het te ontwikkelen instrument (voorbeeld vragenlijst)*
- *Het instrument moet aan een betrouwbaarheids- dan wel valideringsstudie zijn onderworpen*

De volgende instrumenten werden in eerste instantie wel geselecteerd naar aanleiding van de analyse op titel en abstract, bij nadere analyse bleken de instrumenten niet aan bovenstaande criteria te voldoen. Mogelijk voldoen de instrumenten later wel aan deze criteria daarom worden de instrumenten kort genoemd.

*Johnson en Chen (2006)* rapporteren over een studie die de kwaliteit van doceren van artsen in opleiding en stafleden Obstetrie en Gynaecologie met elkaar vergelijkt. De onderzoekers rapporteren wel over de vergelijking van de scores van de aios en stafleden op de vaardigheden in de lijst. Daarentegen wordt niets vermeld over de betrouwbaarheid en de validiteit van de vragenlijst die wordt gebruikt om de stafleden en aios te evalueren. Omdat er in deze studie geen melding wordt gemaakt van de validiteit en/of betrouwbaarheid van de lijst wordt deze niet opgenomen voor nadere analyse. *Donnelly en Woolliscroft (1989)* verrichten een studie met behulp van een beoordeling op een aantal descriptieve items. Deze items werden gebruikt om na te gaan of de beoordelingen van studenten betrouwbaar zijn voor het evalueren van docenten in het medisch onderwijs. In de studie is niet duidelijk uitgewerkt hoe is gekomen tot de selectie van de

items die in de vragenlijst werden gebruikt. Er worden geen aanwijzingen gevonden voor de theoretische fundering van de items die uiteindelijk worden opgenomen in de lijst. Er wordt vooral gekeken of er verschillen in meting zijn betreft de beoordeling van aios en medisch specialisten en niet zozeer of de lijst valide is. Deze studie is dan ook niet geschikt voor opname. *Hamdy, Williams, Tekian, Benjamin, El Shazali, Bandaranayake (2001)* hebben een studie verricht naar het gebruik van het instrument 'VITALS'. Het instrument werd ingezet om het doceren tijdens de co-assistentenschappen te evalueren. Daarnaast werden de resultaten van het onderzoek gebruikt om feedback te geven aan faculteitsleden. Wederom wordt er niets gemeld over de betrouwbaarheid en/of validiteit van de negen indicatoren die gebruikt worden voor de evaluatie. Ook 'VITALS' is niet geschikt voor opname. *Flikweert, Miedema en Briët (2007)* rapporteren over een onderzoek dat plaatsvond in het Deventer ziekenhuis met het doel om de medisch specialist te evalueren op zijn opleiderskwaliteiten. In het artikel van *Flikweert, Miedema en Briët (2007)* worden er geen uitspraken gedaan over de betrouwbaarheid of validiteit van de lijst. In het artikel van *Abu-Zidan en Premadasa (2002)* worden geen uitspraken gedaan over de validiteit en betrouwbaarheid van de lijst die gebruikt wordt in het onderzoek naar de evaluatie van het patroon in instructievaardigheden bij tutoren van de afdeling Chirurgie. Ook deze lijst kan niet worden opgenomen in de analyse. *Dolmans, Wolfhagen, Gerver, De Grave, Scherpbier (2004)* verrichten een beschrijvende studie in Maastricht. Er werd een instrument ontwikkeld om artsen feedback te kunnen geven op de rol van supervisor tijdens patiëntcontacten met studenten. De sterke punten van dit instrument zijn dat het sterk theoretisch gefundeerd is; het is ontwikkeld in samenwerking met artsen die studenten supervisie geven tijdens patiëntcontacten en er is steun vanuit de medische staf voor het instrument. Er is nog geen studie gedaan naar de validiteit en betrouwbaarheid van het instrument. In Canada werd door *Cohen, MacRae en Jamieson (1996)* een studie uitgevoerd naar de effectiviteit van doceren onder chirurgen. De studie duurde negen jaar en registreerde elke jaar de gemiddelde doceereffectiviteits-score (Teaching

Effectiveness Score) van 43 chirurgen. Met deze studie werd een betrouwbaar systeem ontwikkeld voor het evalueren van doceren. In deze studie wordt alleen gemeld dat er gebruik werd gemaakt van een vragenlijst met 28 items. Na verdere literatuurstudie blijkt dat niet te traceren is welke vragenlijst in dit onderzoek is gebruikt.

Hierna volgt een beschrijving van de instrumenten die wel zijn opgenomen in de uiteindelijke analyse. De instrumenten voldoen aan de criteria die aan het begin van deze paragraaf werden genoemd. Voorbeelden van de vragenlijsten zijn opgenomen in bijlage (7).

#### 4.2.1 Analyse instrumenten

Op basis van bovenstaande drie criteria werd besloten om elf instrumenten op te nemen in de analyse hieronder wordt een beschrijving gegeven van de geselecteerde instrumenten. Met behulp van het conceptuele kader dat in hoofdstuk 2 werd beschreven werd bepaald of er instrumenten bestonden die het effect van het 'Teach the Teacher' programma module (1) konden meten. Het instrument moest over items beschikken die de volgende kernpunten bevatte:

- *In een onderwijsmoment moet sprake zijn van een duidelijke opbouw (introductie/inleiding onderwijsmoment, eigenlijke instructie/lesrealisatie, afsluiting/lesevaluatie)*
- *Het is belangrijk dat de leerdoelen van het onderwijsmoment worden vastgesteld en worden gecommuniceerd naar de leerlingen toe*
- *Aansluiten bij de voorkennis van de leerling*
- *Bepaling van de leerstof die wordt aangeboden en aan de hand daarvan bepaling van de organisatie van het onderwijsmoment*
- *Evaluatie van het onderwijsmoment*
- *Het geven van systematische feedback*
- *Rolmodel leren*
- *Het scheppen van een veilige setting*

#### **CCTEI (Cleveland Clinic's Clinical Teaching Effectiveness Instrument)**

Copeland en Hewson rapporteren in 1999 voor het eerst over de CCTEI (Cleveland Clinic's Clinical Teaching Effectiveness Instrument). De CCTEI werd ontwikkeld met de eisen dat het sterk theoretisch gefundeerd moest zijn en gebaseerd op de klinische onderwijskundige literatuur. Verder moest de CCTEI een generiek instrument zijn zodat het op alle afdelingen te implementeren was. Het doel van de implementatie was om artsen met een opleidingstaak op systematische wijze te voorzien van feedback en om op deze manier het doceren te kunnen verbeteren. Het was dan ook van groot belang dat het instrument breed gedragen werd binnen het instituut en bruikbaar zou zijn binnen de instelling (Copeland en Hewson, 1999; Copeland en Hewson, 2000). Copeland en Hewson (1999) geven aan dat het instrument gebruiksvriendelijk is. Het is daarnaast een bruikbaar instrument omdat er een toename wordt gesignaleerd in artsen die hulp zoeken bij vragen rond het doceren. Het instrument kan gebruikt worden op verschillende afdelingen. De dataverzameling in dit onderzoek heeft ook plaatsgevonden op verschillende afdelingen. Geadviseerd wordt om bij de dataverzameling gebruik te maken van meerdere bronnen, bijvoorbeeld door de arts ook te laten beoordelen door collega's. Verder wordt in de discussie gemeld dat er mogelijk sprake kan zijn van een 'ceiling effect' door het hoge gemiddelde en de scheve verdeling van de data. Daardoor kan er bijvoorbeeld geen onderscheid worden gemaakt tussen de groep die goed presteert en de docenten die uitstekend presteren. In dit onderzoek maakt dat niet uit omdat men wil analyseren welke groep nog slecht presteert zodat men kan identificeren welke groep extra scholing behoeft. Boendermaker en Remmelts (2008) leveren de volgende kritiek op de CCTEI. Er wordt gemeld dat de CCTEI in toenemende mate wordt gebruikt voor de beoordeling van de kwaliteit van docenten. Het vermoeden heerst dat er sprake is van een storende factor bij de afname van de lijst.

Er wordt gebruik gemaakt van een vijfpunts-Likert schaal waarbij zowel om een kwantitatieve- als kwalitatieve beoordeling wordt gevraagd (1= never/poor, 5=always/superb). Boendermaker en Remmelts (2008) hebben een pilot-studie verricht om te kijken of er inderdaad een verschil optrad in de interpretatie van antwoordmogelijkheden. In de pilot-studie werd gebruik gemaakt van een aangepaste lijst waarbij een scheiding werd gehanteerd in antwoordcategorieën (tussen frequentie en de kwaliteit van opleidergedrag). Aan de hand van de aangepaste lijst werden er zes klinische docenten beoordeeld door negen aios. Uit het onderzoek blijkt dat er inderdaad een verschil is interpretatie. Wanneer alle scores bij elkaar worden genomen is er in 292 van de 509 scores (totaal) sprake van overeenstemming in de twee aangepaste versies. In 43% van de gevallen was er sprake van een discrepantie tussen de scores (frequentie en kwaliteit worden dan niet hetzelfde gescoord). Voor vijf van de vijftien vragen was dit verschil significant ( $p < ,01$ ). Hier moet gemeld worden dat er gebruik is gemaakt van een zeer kleine groep beoordelaars. Dolmans et al. (2004) geven daarnaast aan dat het instrument niet sterk theoretisch is gefundeerd.

Van der Hem-Stokroos, van der Vleuten, Daelmans, Haarman en Scherpbier (2005) hebben de CCTEI getest op de betrouwbaarheid en validiteit in de Europese context. Daarnaast analyseerden de onderzoekers of het mogelijk was om de kwaliteit van de opleiders te evalueren op groepsniveau in plaats van op individueel niveau. Het onderzoek werd uitgevoerd in Nederland, alvorens het onderzoek van start ging werd de vragenlijst vertaald naar het Nederlands. In deze studie werden de aios en medisch specialist beoordeeld als opleider, de aios en medisch specialisten werden beoordeeld door co-assistenten. In de originele studie werden alleen de medisch specialisten als opleider beoordeeld. In de periode tussen augustus 2002 en april 2003 evalueerden de co-assistenten 51 opleiders. Het aantal co-assistenten per opleider varieerde van één tot dertien co-assistenten. Uiteindelijk bleven er van deze groep opleiders, 37 opleiders over waarvan de data analyseerbaar bleek. De kleine steekproef zorgt ervoor dat

de resultaten eigenlijk niet generaliseerbaar zijn. De bevindingen voor de betrouwbaarheid waren minder gunstig dan de resultaten genoemd in de studie van Hewson & Copeland (2000). Er waren minstens zeven of meer metingen nodig voor een betrouwbare score van één individuele opleider. Om betrouwbare resultaten op groepsniveau te meten waren er vijftien opleiders of meer nodig bij één meting per opleider. De uitkomsten tussen de aios en medisch specialist verschilden niet significant. In de studie van Hewson & Copeland (2000) werden respectievelijk voor één, zes en zeven metingen per opleider generaliseerbaarheidscoëfficiënten gemeten van 0,74, 0,94 en 0,95, in deze studie lagen de uitkomsten een stuk lager 0,40, 0,80 en 0,83 voor hetzelfde aantal metingen.

Voor de interpretatie van individuele scores werd een betrouwbaarheid van 0,80 bereikt met zeven metingen of meer. Om betrouwbare metingen te bereiken op groepsniveau, waren er vijftien opleiders of meer nodig met één meting per opleider. De resultaten laten zien dat de CCTEI gebruikt kan worden om het onderwijs aan co-assistenten te evalueren daarnaast kan het instrument gebruikt worden om de kwaliteit van het lesgeven te evalueren op groepsniveau. (Van der Hem-Stokroos, et al., 2005).

Er zijn een aantal punten van kritiek op de CCTEI te melden. Een pluspunt van dit instrument is dat de data werd verzameld op verschillende afdelingen zodat het ook op verschillende afdelingen ingevoerd kan worden. Boendermaker en Remmelts (2008) voeren als kritiek op dat het instrument gebruik maakt van een kwalitatieve en kwantitatieve scoringsschaal. Het is moeilijk om de schaal goed te scoren. Dolmans et al. (2004) geven aan dat het instrument niet sterk theoretisch is gefundeerd. Gemeld moet worden dat de studie naar de betrouwbaarheid en validiteit in de Europese context werd uitgevoerd binnen het VU medisch centrum te Amsterdam en dat de metingen *alleen* zijn uitgevoerd tijdens het co-assistentenschap Chirurgie. Mogelijk heeft dit invloed op de uitkomsten van dit onderzoek. De bevindingen in dit onderzoek lieten andere resultaten zien dan in de studie van Hewson en Copeland (2000) het is belangrijk om hier



rekening mee te houden bij het gebruik in een andere setting.

### **SFDP 26 (Stanford Faculty Development Program)**

De SFDP is ontwikkeld aan de Stanford University School of Medicine. Dolmans et al. (2004) melden dat de constructvaliditeit van het instrument uitvoerig is onderzocht en dat het instrument valide is gebleken. De SFDP 26 heeft dezelfde zwakte als de CCTEI, het is niet sterk theoretisch gefundeerd. Zowel de CCTEI als de SFDP 26 zijn gebaseerd op componenten van effectief klinisch doceren en niet op de huidige geldende algemene theorieën van contextgebonden leeromgevingen of vormen van facilitatie van bevordering van transfer van kennis naar de werkvloer. Deze karakteristiek is zeer belangrijk want het echte doel van het verzamelen van informatie met deze vragenlijsten is het leveren van een bijdrage aan theorieontwikkeling van effectief leren in deze authentieke omgevingen (Dolmans et al., 2004). Het doel van de studie van Litzelman (1998) naar de SFDP 26 is het instrument te analyseren met behulp van een factor-analyse en het instrument aan te passen tot een praktische lengte voor gebruik in de klinische setting.

De beperkingen van deze studie zijn dat de dataverzameling alleen plaats vond binnen de co-assistentenschappen bij de Interne Geneeskunde en ook alleen binnen deze universiteit.

De respons op het onderzoek was hoog. In deze studie werd alleen gebruik gemaakt van metingen door studenten. De dataverzameling heeft alleen plaatsgevonden bij studenten die op de interne verpleegafdeling co-assistentenschappen liepen. In een volgende studie zal onderzoek uitgevoerd moeten worden naar studenten die op de polikliniek co-assistentenschappen lopen.

### **MTEF-28 (Mayo Teaching Evaluation Form)**

Beckman, Lee, Rohren & Pankratz (2003) rapporteren over de ontwikkeling van de MTEF-28. De MTEF-28 werd ontwikkeld om een beoordeling te kunnen geven over de kwaliteit van doceren in de Mayo Clinic. In deze studie werd gebruik gemaakt van beoordelingen door collega's.

Deze vragenlijst is speciaal ontwikkeld voor het beoordelen van het evalueren van docenten op de behandelafdeling. Het instrument is vrij makkelijk in gebruik, met 28 items. Omdat het instrument alleen binnen één vakgebied is afgenomen zal uit verdere studie moet blijken of de resultaten ook geldig zijn in andere settings. Er is geen validiteitsstudie verricht. Dit is het enige instrument in deze analyse dat alleen gebruik maakt van beoordelingen door collega's. Nadeel is dat de personen niet anoniem waren voor elkaar. De observatoren zagen de medisch specialist en andersom, dit kan de metingen beïnvloeden. Er gaat heel veel tijd in het evaluatieproces zitten voor artsen. Voor sommige items wordt geen significante inter-rater betrouwbaarheid bereikt, er worden verschillende redenen genoemd voor dit feit in de discussie van het artikel waaronder: een verschillende interpretatie van de items, sommige artsen geven hoge scores en anderen zijn heel strikt in het scoren en geven van lagere scores.

### **Instrument voor de evaluatie van opleiderskwaliteiten van stafleden, Academisch medisch centrum/ Universiteit van Amsterdam**

In 2005 werd in het Amsterdams Medisch Centrum (AMC) de start gemaakt met de beoordeling van opleiders en stafleden door artsen in opleiding. Lombarts, Bucx, Rupp, Keijzers, Kokke en Schlack (2007) ontwikkelden een instrument voor de evaluatie van de opleiderskwaliteiten van de stafleden van de afdeling Anesthesiologie van het AMC. Voorafgaand aan de ontwikkeling van het meetinstrument werd een literatuurstudie uitgevoerd naar de beschikbaarheid van bestaande evaluatie-instrumenten. Naar aanleiding van deze literatuurstudie werd besloten een eigen evaluatieformulier te ontwikkelen uitgaande van de SFDP-26 (Stanford Faculty Development Program) vragenlijst (Lombarts Bucx, Rupp, Keijzers, Kokke en Schlack, 2007).

Het instrument is alleen gebruikt op de afdeling Anesthesiologie van het AMC te Amsterdam. De geëvalueerde opleiders kregen van aios hoge scores toebedeeld waarbij aiossen in de beginfase van de

opleiding over het algemeen positiever over de stafleden oordeelden dan aiossen die langer in opleiding waren. Mogelijk is het zo dat naarmate er een kleinere afstand in kennis en kunde is tussen de aios en zijn opleider er mogelijk een kritische houding wordt aangenomen (Lombarts et al. 2007). Het is moeilijk te bepalen of het instrument ook geschikt en bruikbaar is in andere settingen. Er worden geen uitspraken gedaan over de predictieve validiteit. Ook hier wordt geadviseerd om gebruik te maken van meerdere informatiebronnen voor de evaluatie.

### **ECDT (Evaluation of clinical dental teachers)**

Deze studie werd door McGrath, Wai Kit Yeung, Comfort en McMillan (2005) opgezet met het doel de psychometrische kwaliteiten van een vragenlijst te meten voor het evalueren van de doceerkwaliteiten van tandartsen. Het is de enige studie die afkomstig is vanuit een andere sector, dit instrument wordt gebruikt voor het evalueren van tandartsen. Al snel blijkt dat ook dit instrument is gebaseerd op het SFDP- raamwerk en dus niet volledig afkomstig is uit een andere sector. In deze studie werden verschillende vormen van validiteit beoordeeld. Daarnaast werd een hoge mate van betrouwbaarheid berekend. Er is sprake van een grote steekproef met minstens tien beoordelingen per docent. De vraag is of het instrument bruikbaar is voor evaluatie van de doceervaardigheden van artsen. Het instrument zal voor gebruik eerst gevalideerd moeten worden voor het onderwijs tijdens de dagelijkse patiëntenzorg in het ziekenhuis.

### **MedIQ (Medical Instructional Quality)**

De MedIQ is een instrument dat unieke aspecten meet van het klinische instructieproces. In het onderzoek van James, Kreiter, Shipengrover en Crosson (2002) werd onderzoek gedaan naar de constructvaliditeit van de MedIQ. James en Osborne (1999) rapporteren ook over de MedIQ. De resultaten van dit onderzoek laten zien dat de betrouwbaarheid voor elke schaal hoog is, er is sprake van een gemiddelde score voor de constructvaliditeit, en een zwakke criteriumvaliditeit.

De kritiek die wordt geleverd op de MedIQ is dat er met het oog op de dataverzameling wordt gepleit voor afname van de vragenlijst in andere settingen. Hier is het beperkt gebleven tot drie verschillende co-assistentschappen.

### **An encounter card to evaluate teachers in clerkship**

De auteurs Oppenheimer, Keely en Marks (2006) rapporteren over een onderzoek met als doel het evalueren van de doceereffectiviteit. In het onderzoek werden zowel aios als medisch specialisten beoordeeld. Het onderzoek is alleen uitgevoerd op de afdeling Obstetrie en Gynaecologie. In de resultaten van het onderzoek wordt niets gemeld over de validiteit van het instrument. Het is een beknopt instrument met tien sleutelaspecten dus het is eenvoudig af te nemen.

### **SETOC (Student evaluation of teaching in outpatient clinics)**

Het doel van de studie van Zuberi, Bordage, Norman (2007) is om specifiek een instrument te ontwikkelen dat kan worden gebruikt om het doceren op de poliklinieken te beoordelen. De conclusie van het onderzoek luidt dat er betrouwbare en valide scores van metingen door studenten worden verkregen. Deze scores kunnen specifieke feedback genereren voor individuele feedback aan faculteitsleden met zwakke of uitstekende vaardigheden daarnaast kan het tekortkomingen van de faculteitsleden op groepsniveau identificeren. Deze vragenlijst is speciaal ontwikkeld voor de poliklinische setting. Het formulier bevat een overzichtelijk aantal categorieën, waardoor het vrij snel is af te nemen. Er werd een hoge inter-rater betrouwbaarheid en validiteit bereikt. De lijst is in verschillende settingen gevalideerd. Daarnaast is het instrument betrouwbaar gebleken om individuele feedback te geven aan stafleden of zij zwakke of sterke doceervaardigheden hebben en kan het instrument feedback geven op groepsniveau over waar de zwakke punten in het systeem zich bevinden. De lijst is getest in een westerse universiteit (Canada) en in een universiteit in het Midden-Oosten (Pakistan). De lijst is alleen getest voor de evaluatie van het doceren op de polikliniek. Het

instrument zal opnieuw gevalideerd moeten worden wil het bruikbaar zijn voor evaluatie op de verpleegafdeling.

### **GRS (Global Rating Scale, University of Michigan)**

Binnen de universiteit van Michigan hebben Williams, Litzelman, Babbott, Lubitz en Hofer, (2002) een instrument ontworpen om de kwaliteit van klinisch doceren te meten. Het doel van deze studie is om een instrument te ontwikkelen wat makkelijk af te nemen is in plaats van de instrumenten die meerdere items bevatten en vaak veel tijd in beslag nemen. Litzelman en Shea (2005) melden dat korte instrumenten het voordeel hebben dat ze relatief gemakkelijk zijn af te nemen en dat de verwerking betrekkelijk weinig tijd en kosten met zich meebrengt maar vaak alleen de bekwaamheid van klinische docenten meet. De korte instrumenten kunnen geen specifieke feedback genereren om individuen inzicht te geven in de verbeterpunten met betrekking tot het doceren.

Wanneer het instrument in praktijk wordt gebruikt moeten er dus, met het oog op de betrouwbaarheid, in ieder geval zeven of meer metingen verricht worden bij een docent. Het instrument mist precisie om kleine verschillen tussen docenten te bepalen. Met het instrument kan wel een globale schatting gemaakt worden tussen docenten, bijvoorbeeld: docenten die niet zo goed presteren, docenten die gemiddeld presteren en docenten die goed presteren. Er kan in ieder geval een uitspraak worden gegenereerd over de docenten die heel slecht presteren. De GRS dient gebruikt te worden naast andere evaluatiemethoden bijvoorbeeld: het evalueren van colleges en lezingen. In het artikel worden geen uitspraken gedaan over de validiteit van het instrument. Het instrument is absoluut niet geschikt om het effect van de 'Teach the Teacher' cursus te meten. Eén item is niet specifiek genoeg om te meten of de medisch specialist de cursusinhoud toepast op de werkvloer.

### **WICT (Wisconsin Inventory of Clinical Teaching)**

Het doel van de studie van Hewson en Jensen (1990) is om een instrument te ontwikkelen dat kan worden gebruikt als

onderdeel van een programma voor het verbeteren van de docererkwaliteiten van docenten aan de universiteit van de Wisconsin. De conclusie van het onderzoek geeft aan dat er valide en betrouwbare resultaten zijn gevonden. Deze vragenlijst is speciaal ontwikkeld voor gebruik in de faculteit van Wisconsin. In de discussie wordt dan ook aangegeven dat er verder onderzoek zal moeten plaatsvinden wanneer het in andere settings en universiteiten wordt ingezet. Daarentegen vertonen de items van de WICT ook grote overlap met de inhoud van andere instrumenten die ook betrouwbaar en valide zijn gebleken. Deze instrumenten werden ook in verschillende settings gebruikt. Dit pleit voor betrouwbaarheid en validiteit van de WICT in andere settings. Het formulier bevat vrij veel items, mogelijk neemt het invullen van de lijst veel tijd van de arts in beslag.

### **CTAF (Clinical Teaching Assessment Form)**

In het voorjaar van 1977 besloot de afdeling Obstetrie en Gynaecologie van de Universiteit van Washington dat het klinisch doceren van stafleden en artsen in opleiding geëvalueerd diende te worden. Het doel van Irby en Rakestraw (1981) was een evaluatiesysteem te ontwikkelen voor het geven van diagnostische feedback aan stafleden en aios en op basis van deze evaluatie goede administratieve besluiten te kunnen nemen (promotie, salarisverhoging). Het instrument is alleen gebruikt op de afdeling Obstetrie en Gynaecologie en ontwikkeld voor het gebruik op de afdeling en niet op de polikliniek. De vraag is of het instrument dezelfde betrouwbaarheid oplevert wanneer het op andere afdelingen wordt toegepast en wanneer het op de polikliniek wordt ingezet.

Ramsbottom-Lucier, Gillmore, Irby en Ramsey (1994) rapporteren met het doel om de meetkarakteristieken te beoordelen van de CTAF voor gebruik op de polikliniek. Met deze studie werd beoordeeld of de scores op de polikliniek afweken van de scores die dezelfde stafleden kregen op de interne behandelafdeling. Het formulier werd betrouwbaar gevonden voor gebruik op de polikliniek. Er werden in totaal 29 artsen geëvalueerd die op de poli en op

de afdeling doceerden binnen de University of Washington. Van 1985-86 tot 1989-90 werden de evaluatieformulieren ingevuld. Artsen in opleiding genereerden in totaal 639 formulieren. Wanneer er meer dan tien metingen per staflid aanwezig waren lag de betrouwbaarheid in de range van .58 tot .81. De metingen die op de poli werden verricht lagen significant hoger (5.01 versus 5.35,  $p < .05$ ). Hier moet wel worden opgemerkt dat in de eerste studie beoordeeld werd door studenten en in deze studie werd er beoordeeld door artsen in opleiding. Mogelijk treed er ook daarom een verschil op in meting. De CTAF is een vrij beknopt instrument daarom is het gemakkelijk af te nemen. Er wordt niets gemeld over de predictieve validiteit van het instrument.

Onderstaand overzicht geeft de items weer die in aanmerking komen om in de te ontwikkelen lijst te worden opgenomen.

### Selectie van items uit instrumenten op basis van conceptueel kader

---

**Criterion 1:** *In een onderwijsmoment moet sprake zijn van een duidelijke opbouw (introdunctie/inleiding onderwijsmoment, eigenlijke instructie/lesrealisatie, afsluiting/lesevaluatie)*

---

In de instrumenten werden geen items gevonden die aansluiten bij dit criterium.

---

**Criterion 2:** *Het is belangrijk dat de leerdoelen van het onderwijsmoment worden vastgesteld en worden gecommuniceerd naar de leerlingen toe*

---

- *Clearly specifies what I am expected to know and do during this training period (CCTEI)*
- *Stated goals clearly and concisely (SFDP 26)*
- *Stated relevance of goals to learners (SFDP 26)*
- *Prioritized goals (SFDP 26)*
- *Repeated goals periodically (SFDP 26)*
- *Stated goals and expectations of the team (MTEF-28)*
- *Stated relevance of goals to learners (MTEF-28)*

- *Prioritized goals (MTEF-28)*
- *Repeated goals periodically (MTEF-28)*
- *leerdoelen duidelijk formuleert (IEOS)*
- *relevante leerdoelen opstelt (IEOS)*
- *leerdoelen prioriteert (IEOS)*
- *regelmatig terugkomt op gestelde leerdoelen (IEOS)*
- *Stated goals clearly and concisely (ECDT)*
- *Stated relevance of goals to me (ECDT)*
- *Prioritised goals (ECDT)*
- *Repeated goals periodically (ECDT)*
- *orientation of learner to expectations (encounter card)*
- *setting of expectations in the clinics (responsibilities, assignments, grading etc.) (SETOC)*

---

### Criterion 3: Aansluiten bij de voorkennis van de leerling

---

- *Instructed me at a level consistent with my knowledge and skill (MedIQ)*
- *adjusts teaching to my needs (experience, competence, interest, etc.) (CCTEI)*
- *Adjusts amount of supervision to level of competence of residents (WICT)*
- *Adjusts instructional approach to different levels of resident training (WICT)*

---

### Criterion 4: Organisatie van het onderwijsmoment

---

- *Asks questions that promote learning (clarifications, probes, Socratic questions, reflective questions, etc.) (CCTEI)*
- *Called attention to time (SFDP 26)*
- *Avoided digressions (SFDP 26)*
- *Discouraged external interruptions (SFDP 26)*
- *Presented well organized material (SFDP 26)*
- *Explained relationships in materials (SFDP 26)*
- *Used whiteboard or other visual aids (SFDP 26)*
- *started and finished rounds on time (MTEF-28)*
- *Discouraged external interruptions (MTEF-28)*
- *zich bij besprekingen aan het leerdoel houdt; hij/zij kijkt niet onnodig af van het onderwerp (IEOS)*

- zich goed voorbereid op onderwijspresentaties en besprekingen (**IEOS**)
- *paid attention to time (ECDT)*
- *Avoided digression (ECDT)*
- *Discouraged external interruptions (ECDT)*
- *was well organised (ECDT)*
- *Explained relationships of dental knowledge to clinical situations (ECDT)*
- *Used aids in learning (ECDT)*
- *Asked questions to enhance my learning (MedIQ)*
- *'organised' (encounter card)*
- *probing your understanding by asking questions (why, what if, suppose, what else) (SETOC)*
- *'is clear and organized' (CTAF)*

---

**Criterion 5: Evaluatie van het onderwijsmoment**

- *Evaluated learners' knowledge of factual medical information (SFDP 26)*
- *Evaluated learners' ability to analyze or synthesize medical knowledge (SFDP 26)*
- *Evaluated learners' ability to apply medical knowledge to specific patients (SFDP 26)*
- *Evaluated learners' medical skills as they apply to specific patients (SFDP 26)*
- *Asked learner to discuss differential diagnosis on most patients (MTEF-28)*
- *Asked learners to discuss alternative management options on most patients (MTEF-28)*
- *Asked learners to demonstrate physical exam skills on rounds (MTEF-28)*
- *Evaluated learners' knowledge of factual medical information (MTEF-28)*
- *Evaluated my knowledge of factual information (ECDT)*
- *Evaluated my ability to analyse or synthesise clinical learning (ECDT)*
- *Evaluated my ability to apply dental knowledge to specific patients (ECDT)*

- *Evaluated my clinical skills ability as they applied to specific patients (ECDT)*
- *'assessment of learner's knowledge' (encounter card)*
- *'formal evaluation of performance' (encounter card)*
- *Promotes reflection and clinical practice (WICT)*
- *Questions residents to probe their knowledge and judgements (WICT)*

---

**Criterion 6: Het geven van systematische feedback**

- *Gave negative (corrective) feedback to learners (SFDP 26)*
- *Explained to learners why he/she was correct or incorrect (SFDP 26)*
- *Offered learners suggestions for improvement (SFDP 26)*
- *Gave feedback frequently (SFDP 26)*
- *Gave learners regular, useful feedback on their performance (MTEF-28)*
- *Gave negative (corrective) feedback to learners (MTEF-28)*
- *explained to learners why they were correct or incorrect (MTEF-28)*
- *Offered learners suggestions for improvement (MTEF-28)*
- *regelmatig positieve feedback geeft aan assistenten (IEOS)*
- *corrigerende feedback geeft aan assistenten (IEOS)*
- *corrigerende feedback ook toelicht (IEOS)*
- *suggesties voor verbetering geeft (IEOS)*
- *Gave me negative (corrective) feedback (ECDT)*
- *Explained why I am correct or not (ECDT)*
- *Offered me suggestions for improvement (ECDT)*
- *Gave me feedback frequently (ECDT)*
- *'provided feedback', 'helpful feedback' (encounter card)*
- *identification of your strengths (SETOC)*
- *identification of areas needing improvement (SETOC)*
- *'provides direction and feedback' (CTAF)*

- Provides constructive criticism for inappropriate behaviours (WICT)
- Gives positive feedback for appropriate behaviours (WICT)
- Give timely feedback to residents (WICT)

---

**Criterion 7: Rolmodel leren**

---

- was a good role model of a caring doctor (MTEF-28)
- een voorbeeldfiguur is voor mij als toekomstig anesthesioloog (IEOS)

---

**Criterion 8: Het scheppen van een veilige setting**

---

- Establishes a good learning environment (approachable, nonthreatening, enthusiastic, etc.) (CCTEI)
- listened to learners (SFDP 26)
- encouraged learners to participate actively in the discussion(SFDP 26)
- Expressed respect for learners (SFDP 26)
- Encouraged learners to bring up problems (SFDP 26)
- made learners feel comfortable asking questions on rounds (MTEF-28)
- allowed learners to present without frequent interruptions (MTEF-28)
- never ordered tests without telling the learner (MTEF-28)
- expressed respect for learners (MTEF-28)
- treated his or her residents kindly (MTEF-28)
- showed enthusiasm for his/her work/learners (MTEF-28)
- assistenten aanmoedigt actief te participeren in discussies (IEOS)
- assistenten stimuleert om problemen in te brengen (IEOS)
- goed luistert naar assistenten (IEOS)
- respectvol is naar assistenten (IEOS)
- goed toegankelijk is in de dienst (IEOS)
- goed toegankelijk is voor overleg (op de pijn-poli) (IEOS)
- listened to me (ECDT)

- encouraged me to participate actively in discussion (ECDT)
- showed me respect (ECDT)
- Encouraged me to bring up problems (ECDT)
- Listened well to me (MedIQ)
- Created an environment in which I felt comfortable accepting challenges, even at the risk of making mistakes (MedIQ)
- 'enthusiasm' (encounter card)
- approach (encounter card)
- friendly environment (encounter card)
- is enthusiastic and stimulating (CTAF)
- actively involves students (CTAF)
- Is available for additional teaching when requested by residents (WICT)
- Shows sensitivity to resident's feelings (WICT)
- Is available, approachable and receptive to resident's personal problems (WICT)
- Inspires confidence in residents to solve clinical problems (WICT)
- Is enthusiastic about challenges of practise (WICT)
- Demonstrates enthusiasm for teaching in clinic (WICT)
- Engages in didactic teaching when resident's responses indicate a need (WICT)
- Demonstrates a flexible, open-minded approach to alternative suggestions by residents (WICT)

Omdat uit bovenstaande analyse is gebleken dat op basis van het conceptueel kader geen enkel instrument volledig te gebruiken is voor het meten van het effect van het 'Teach the Teacher' programma module (1) zal worden gestart met de ontwikkeling van een nieuw meetinstrument.

Uit bovenstaande analyse zullen de volgende aandachtspunten worden meegenomen waarmee rekening dient te worden gehouden bij de ontwikkeling van het nieuwe instrument. Het merendeel van de instrumenten is alleen getest binnen een beperkt aantal specialismen. De CCTEI werd getest op verschillende afdelingen en daarmee werd een generiek instrument ontwikkeld dat op

verschillende afdelingen inzetbaar was. In andere artikelen wordt in de discussie gemeld dat er nog vervolgonderzoek moet plaatsvinden om er zeker van te zijn dat het instrument op meerdere afdelingen toepasbaar is, dit is het geval bij de SFDP-26, MTEF-28, de lijst voor de afdeling Anesthesiologie van het AMC, MedIQ, 'an encounter card to evaluate teachers in clerkship' en de CTAF. Het te ontwikkelen instrument moet dus in verschillende settingen af te nemen zijn. De CCTEI brengt echter weer een ander probleem met zich mee, de lijst maakt gebruik van een schaal waar zowel om een kwalitatieve als kwantitatieve beoordeling wordt gevraagd. Dit zorgt voor onduidelijkheid bij het interpreteren van de items zo melden Boendermaker en Remmelts (2008). In het te ontwikkelen instrument zal in ieder geval gebruik moeten worden gemaakt van een eenduidige schaal. Naast de dataverzameling op verschillende afdelingen zal ook rekening moeten worden gehouden met instrumenten die alleen informatie hebben verzameld op de poliklinische afdeling of op de interne behandelafdeling. De SETOC is bijvoorbeeld specifiek ontwikkeld voor de evaluatie op de polikliniek. De CTAF is wel uitgebreid getest voor zowel gebruik op de polikliniek als gebruik op de interne behandelafdeling en kan dus in beide contexten worden toegepast. Het nieuwe instrument moet zowel toepasbaar zijn op de poli als op de interne behandelafdeling. In de studies die hier beschreven zijn wordt voornamelijk gebruik gemaakt van vragenlijsten. Ook in deze studie zal gebruik worden gemaakt van een vragenlijst. Er wordt voor gekozen om de medisch specialist en aios zo min mogelijk te belasten met betrekking tot de tijdsinvestering in dit onderzoek, daarom wordt gekozen voor een vragenlijst. In deze analyse zijn ook alleen vragenlijsten aangetroffen om de docent te evalueren. Daarom wordt er ook vanuit gegaan dat de vragenlijst op dit moment het breedst toegepaste middel is bij het evalueren van het medisch onderwijs. In een volgend onderzoek kan mogelijk gebruik worden gemaakt van zowel kwalitatief als kwantitatief onderzoek om te zorgen dat de betrouwbaarheid van de waarnemingen wordt verhoogd. De vragenlijst kan dan bijvoorbeeld worden gecombineerd met observaties of

interviews met betrokkenen. Ook kan er gebruik worden gemaakt van meerdere informatiebronnen. De beoordeling kan plaatsvinden door bijvoorbeeld de arts zelf, door de aios en door collega's van de arts. Beckman et al. (2003) geven aan dat het evalueren door collega's nog een extra waarde kan vervullen. Collega's leren mogelijk ook van collega artsen door een andere stijl van doceren waar te nemen en hiermee mogelijk nieuwe doceervaardigheden kunnen overnemen voor het eigen doceren. Snell, Tallett, Haist, Hays, Norcini, Prince, Rothman en Rowe (2000) geven aan dat bij de evaluatie, waar mogelijk, gebruik dient te worden gemaakt van meerdere informatiebronnen, de evaluatie moet subjectieve en objectieve methoden bevatten en zowel kwantitatieve als kwalitatieve informatie. Wanneer de studies nog verder worden geanalyseerd op de contexten van het onderzoek kan er worden geconstateerd dat er verschillen optreden wanneer de lijst gebruikt wordt in een ander werelddeel. De CCTEI werd bijvoorbeeld ook gevalideerd voor gebruik in de Europese context. Hieruit bleek dat de scores een stuk minder betrouwbaar waren dan in de Amerikaanse context. In de studie naar de SETOC werd de lijst in Canada en in Pakistan gevalideerd. Hier werden in beide onderzoeken wel betrouwbare scores bereikt. In een later stadium van het valideringsproces dient hiermee rekening te worden gehouden.

Verder wordt in de studie van Lombarts, Bucx, Rupp, Keijzers, Kokke en Schlack (2007) gemeld dat rekening dient te worden gehouden met de verschillen in scores tussen aios die in het begin van de opleiding verkeren en aios die verder in de opleiding zijn. Uit dit onderzoek blijkt dat er een verschil is in beoordelingen tussen deze twee groepen aios. Mogelijk treedt er ook een verschil op wanneer de medisch specialist wordt beoordeeld door een co-assistent of een aios. Dit verschil kan niet alleen optreden door een verschil in kennis, maar ook doordat een co-assistent een medisch specialist korter meemaakt dan een aios. Ramsbottom-Lucier, Gillmore, Irby en Ramsey (1994) melden dat verschillen in metingen mogelijk kunnen optreden doordat er een verschil is in de tijd die de leerling

met de arts heeft doorgebracht. In de analyses dient hiermee rekening te worden gehouden. Bij de afname van de nieuwe lijst moet worden overwogen welke keuze er wordt gemaakt met betrekking tot de beoordelaar. Lombarts et al. (2007) merken op dat het vragenlijstonderzoek onder aios de meest gebruikte evaluatiemethode is en dat de aios bij uitstek geschikt is om feedback te geven op de prestaties de opleider en dat deze groep suggesties kan geven voor verbetering. Tevens melden Lombarts et al. (2007) dat er gebruik dient te worden gemaakt van verschillende evaluatiemethoden. “Voor het krijgen van een compleet en betrouwbaar beeld van iemands opleiderskwaliteiten of diens bredere professionele functioneren, adviseren deskundigen daarom gebruik te maken van een combinatie van evaluatiemethoden” (Lombarts et al., 2007, p. 2007). In de lijst van Lombarts et al. (2007) is ook een item opgenomen dat vraagt om een algemeen oordeel van het te evalueren staf lid. Dit item zien we ook terug bij de SFDP-26 en de CCTEI. Omdat het bij de ontwikkeling van het nieuwe instrument in eerste instantie gaat om het meten van het effect van de ‘Teach the Teacher’ cursus zal in dit ontwikkelingsstadium van de vragenlijst niet een dergelijk item worden opgenomen. De items in de lijst van Lombarts et al. (2007) zijn duidelijk en bondig omschreven, bij de te ontwikkelen lijst dient hier rekening mee te worden gehouden, de items moeten voor alle betrokkenen in een keer duidelijk zijn. Naast de Likertschaal is de optie ‘kan ik niet beoordelen’ opgenomen. Naar mijn mening is dit een goede optie wanneer er ook door co-assistenten beoordeeld zal worden. Dit is de enige Nederlandse versie van een vragenlijst voor het evalueren van opleiders, die in aanmerking kwam voor de uiteindelijke analyse. Omdat hier gebruikt wordt gemaakt van Nederlandse items treden er geen problemen op bij de vertaling van items. In de studie naar de ECDDT werd uitvoerig onderzoek gedaan naar verschillende vormen van validiteit daarnaast werd een hoge mate van betrouwbaarheid berekend en er werd gebruik gemaakt van een grote steekproef. Bij het onderzoek naar de psychometrische kwaliteiten van het nieuwe instrument kan een voorbeeld worden genomen aan de studie naar de ECDDT. Veel instrumenten richten zich

of op de betrouwbaarheid of op de validiteit, het is van belang dat er in de nieuwe studie zowel uitspraken worden gedaan over de validiteit en de betrouwbaarheid van het instrument. De GRS maakt gebruik van één item in het nieuwe instrument bij de nieuwe lijst dient in ieder geval gebruikt te worden gemaakt van meerdere items om een goede representatie van de cursusinhoud in de lijst te verwerken. Bij voorkeur dienen de medisch specialisten door meerdere leerlingen te worden beoordeeld om een betrouwbare score te verkrijgen.

Afsluitend kunnen nog enkele praktische punten worden genoemd. Omdat de lijst in ontwikkeling is dient er ruimte beschikbaar te zijn voor opmerkingen, de opmerkingen kunnen zorgen voor eventuele aanpassingen in de lijst. Om te zorgen dat de enquête zo min mogelijk tijd in beslag neemt zal de enquête digitaal worden gemaakt. Het instrument dient goed theoretisch te worden gefundeerd op de huidige onderwijskundige theorieën (Dolmans et al., 2004). Tenslotte zal het instrument in een later stadium tot een praktische lengte moeten worden gereduceerd om te zorgen dat er niet te veel tijd in beslag wordt genomen bij het invullen van de vragenlijst.

Geconstateerd kan worden dat er op basis van het conceptueel kader dat in hoofdstuk (2) werd geschetst geen volledige instrumenten geschikt zijn die het effect van het ‘Teach the Teacher’ programma module (1) kunnen meten. Op basis van de criteria die voortvloeiden uit het conceptueel kader zijn wel een aantal items geselecteerd die bruikbaar kunnen zijn bij de ontwikkeling van het nieuwe instrument. Deze items kunnen richting geven bij de formulering van items voor het nieuwe instrument. Met behulp van de literatuurstudie is een duidelijk beeld gevormd van de aandachtspunten waarbij de ontwikkeling van het nieuwe instrument rekening dient te worden gehouden. In de volgende paragraaf zal het nieuwe instrument worden ontwikkeld. Daarnaast wordt de validiteit en betrouwbaarheid van het instrument bepaald.



### 4.3 Ontwikkeling van een betrouwbaar en valide instrument

In deze paragraaf zal antwoord worden gegeven op de vraag in hoeverre het mogelijk is een betrouwbaar en valide instrument te ontwikkelen voor het bepalen van het effect van didactische basisvaardigheidstrainingen aan medisch specialisten. In deze paragraaf wordt antwoord gegeven op vraagstelling 2:

*Wat is de betrouwbaarheid en validiteit van het ontwikkelde instrument voor het bepalen van het effect van een training in didactische basisprincipes aan medisch specialisten?*

In de vorige paragraaf is geconcludeerd dat bestaande instrumenten onvoldoende geschikt zijn om het effect van het 'Teach the Teacher' programma module (1) te meten. Er zijn wel items uit bestaande instrumenten die gebruikt kunnen worden bij de ontwikkeling van het instrument. Deze items kunnen bijvoorbeeld een rol spelen bij de formulering van de nieuwe items. In paragraaf 4.3.1 wordt beschreven op welke wijze het instrument is ontwikkeld, daarnaast wordt ingegaan op het proces rond het bepalen van de inhoudsvaliditeit. In paragraaf 4.3.2 wordt vervolgens ingegaan op het bepalen van de constructvaliditeit van het instrument en de gegevens die voortkomen uit betrouwbaarheidsstudie van het instrument worden gepresenteerd.

#### 4.3.1 Inhoudsvaliditeit van het ontwikkelde instrumentarium

De inhoudsvaliditeit werd in het eerste stadium vastgesteld door twee maal te observeren tijdens het 'Teach the Teacher' programma module (1). Tijdens deze twee cursussen verschilden de docenten die de cursus gaven, op deze manier werd uitgesloten dat de cursusinhoud niet veranderde door de docent die de cursus verzorgde. Daarnaast werd er twee keer geobserveerd om ervoor te zorgen dat er bij de eerste observatie geen vaardigheden

over het hoofd werden gezien. Voor de formulering van de items werd gebruik gemaakt van de selectie van items die uit vraagstelling 1 voortkwamen.

Samenvattend kan het volgende worden gemeld over de bevindingen tijdens de cursus. Tijdens de eerste cursusdag is er aandacht voor de inleiding van de cursusinhoud, de doelen van de cursus worden gecommuniceerd en de structuur van de cursus wordt duidelijk uiteengezet. De cursusleiders zorgen dat de aandacht van de cursisten vast wordt gehouden (stellen van verschillende vragen aan verschillende cursisten, afwisseling in de cursusinhoud, gebruik van verschillende werkvormen). De cursusleiders stellen de cursisten op hun gemak, op die manier wordt mede gezorgd voor het creëren van een veilige setting. De beginsituatie van de cursisten wordt vastgesteld. Hoe ver zijn de cursisten al in het geven van onderwijs? Hebben de cursisten al eerder cursussen gevolgd met betrekking tot het geven van onderwijs? Wat willen de cursisten leren tijdens de cursus? Afwisselend wordt er theorie gegeven en wordt er ruimte ingebouwd om de theorie in praktijk te brengen en te oefenen in het geven van onderwijsmomenten. De arts wordt attent gemaakt op het feit dat deze altijd een rolmodel functie heeft zowel als arts, als docent en als mens. Er wordt geoefend in het geven van feedback. Er worden methoden aangeleerd over hoe feedback te geven en er wordt ook daadwerkelijk geoefend met het geven van feedback. Aandacht wordt besteed aan het opstellen van doelen en daarnaast wordt bepaald of de doelen ook bereikt worden door middel van evaluatie. De cursist krijgt handvatten hoe het onderwijs te organiseren. Er moet bijvoorbeeld sprake zijn van een goede voorbereiding en een duidelijke opbouw in het onderwijsmoment. Tijdens de tweede cursusdag ligt het hoofddaccent voornamelijk op het oefenen van de stof die tijdens de eerste cursusdag werd aangeboden. De cursisten blikken terug op de vorige dag. Zijn er al persoonlijke leerdoelen bereikt of moeten leerdoelen worden bijgesteld? De tweede cursusdag wordt nieuwe theorie aangeboden met betrekking tot de beoordeling van de leerling. Het is belangrijk dat er vooraf criteria worden

gesteld en op basis van deze criteria wordt getoetst, op die manier kan een goede en objectieve beoordeling plaatsvinden.

Vervolgens werden de bevindingen uit de observaties omgezet naar gedragingen. Deze omzetting is te vinden in bijlage (8). Deze lijst vormde het uitgangspunt voor de ontwikkeling van het nieuwe instrument en bestond uit 28 stellingen. Vervolgens werd deze lijst besproken met twee docenten afkomstig uit het 'Teach the Teacher' programma module (1). De lijst leverde nieuwe input op. De stellingen dienden te worden vereenvoudigd en verder uit elkaar te worden gehaald. Daarnaast werden de stellingen in categorieën ondergebracht om een heldere structuur neer te zetten in de ontwikkelingsfase. Bij de lijst werd een uitleg gevoegd om de stellingen te verduidelijken voor de docenten en de onderzoeker zelf. Tevens worden hier discussiepunten opgenomen die in het gesprek besproken dienden te worden. Deze lijst werd bediscussieerd en weer aangepast. De tweede lijst bestond uit 52 stellingen. Met deze uitgebreide lijst werden de gesprekken gestart rondom het bepalen van de inhoudvaliditeit.

Met betrekking tot de items over '*het opstellen van doelen*' werden er drie items verwijderd en samengevat in één item. Het is belangrijk dat de begeleider het leerproces aan de hand van persoonlijke leerdoelen monitort. Hoe vaak dit gebeurd en op welke manier daar kan de begeleider zelf een invulling aan geven. In het laatste stadium werd hier nog een item aan toegevoegd: *Mijn begeleider weet wat mijn persoonlijke leerdoelen zijn.*

Ook bij de items over het '*geven van feedback*' werden een aantal items verwijderd. Vanuit de cursus is er één gegeven zeer belangrijk, dat er feedback wordt gegeven en dat dit op een constructieve manier gebeurd. Er werden in totaal zes items verwijderd die vervolgens werden vervangen door één samenvattend item. Op welke manier er feedback wordt gegeven is niet van belang zo lang de feedback maar constructief en de gegeven feedback wordt geconcretiseerd zodat de arts in opleiding zich verder kan ontwikkelen met de feedback die hij krijgt.

Met betrekking tot de stellingen over het onderwerp '*doceren*' werden vijf items verwijderd. Deze stellingen hadden betrekking op het stellen van verschillende soorten vragen (kennisvragen, begripsvragen, toepassingsvragen, synthesevragen en evaluatievragen). Deze stellingen kunnen samengevat worden in een vraag: *Mijn begeleider stelt verschillende soorten vragen (kennisvragen, begripsvragen, toepassingsvragen enz.).* De twee items die betrekking hadden op het betrekken van zowel de 'stillere' types als het afremmen van de 'haantjes de voorste' werden vervangen door één samenvattend item: *Mijn begeleider zorgt dat iedere deelnemer een actieve rol heeft tijdens een onderwijsmoment.*

Drie items in de lijst hadden betrekking op het vasthouden van de aandacht. Deze items werden dan ook vervangen door één samenvattend item: *Mijn begeleider houdt de aandacht tijdens een onderwijsmoment goed vast.*

Er waren twee items die betrekking hadden op of de begeleider controleert of de stof van het onderwijsmoment is begrepen. Deze twee items werden samengevat in één item: *Mijn begeleider controleert of de stof van een onderwijsmoment is begrepen.*

Er waren een aantal items, zeven in totaal, die taaltechnisch anders werden geformuleerd om te zorgen dat het item daadwerkelijk observeerbaar was voor de respondent en om te zorgen dat het item helder en duidelijk was omschreven zodat het niet voor onduidelijkheden zou zorgen bij het invullen van de vragenlijst.

Er werden zeven items verwijderd die of te specifiek waren geformuleerd op de cursusinhoud (gebruiken van de vier stappen methode, positie docent in de onderwijsruimte) of die bij nader inzien toch niet van toepassing waren op de cursusinhoud (directe toepasbaarheid van kennis in een onderwijsmoment).

Er waren twee items onder het kopje '*begeleiding*' die niet observeerbaar bleken, er is voor gekozen om deze items te verwijderen.

Twee items die betrekking hadden op het onderwerp 'beoordeling' werden verwijderd en vervangen door één samenvattend item: *Mijn begeleider zorgt voor summatieve en formatieve beoordelingen.*

Er werd één item toegevoegd aan de lijst: *Mijn begeleider geeft veel verschillende voorbeelden tijdens onderwijsmomenten.* Vanuit de gesprekken kwam dit als nieuwe input naar voren.

In de afrondende fase van het bepalen van de inhoudsvaliditeit met de docenten van het 'Teach the Teacher' programma vonden voornamelijk veranderingen plaats die betrekking hadden op de formulering van de items.

Vervolgens werd de lijst vanuit de Onderwijskunde besproken. De lijst werd verder onderverdeeld in onderwerpen. Doceren werd onderverdeeld naar opbouw van een onderwijsmoment: introductie, uitvoering en afronding. Drie items werden anders geformuleerd omdat de items mogelijk voor onduidelijkheid zouden zorgen bij het invullen van de vragenlijst.

Om ervoor te zorgen dat er bij de factoranalyse voldoende factoren werden onderscheiden werden zeven items verder uit elkaar gehaald.

De items moeten allemaal observeerbaar zijn voor de respondent. De vraag die bij het item over de 'arts als rolmodel' rijst is of een leerling kan beoordelen of een arts zich bewust is van zijn rolmodel. In de laatste fase vond discussie plaats over de items met betrekking tot de rolmodelfunctie van de docerend arts. Uiteindelijk werd besloten om de items als volgt te formuleren: *Mijn begeleider maakt bewust gebruik van zijn rolmodelfunctie als arts* en *Mijn begeleider maakt bewust gebruik van zijn rolmodelfunctie als opleider* en *Mijn begeleider maakt zijn rolmodelfunctie bespreekbaar*. Hier wordt nu onderscheid gemaakt tussen het rolmodel als arts en als opleider en het wordt observeerbaar gemaakt.

Na discussie over de laatste versie van de lijst werden de volgende gegevens toegevoegd. Er dienden persoonsgegevens van de arts en de respondent te worden verzameld en er werd gediscussieerd over de antwoordmogelijkheden. In bijlage (4) is de uiteindelijke digitale lijst te vinden.

Er zijn gesprekken gevoerd totdat er geen nieuwe input vanuit de betrokkenen kwam. Hiermee is de inhoudsvaliditeit vastgesteld. De lijst werd daarnaast beoordeeld door enkele medestudenten van het Ontwikkelplatform met het oog op de formulering van items en onduidelijkheid over woordkeuze. Het oordeel van de vier medestudenten was dat de items voldoende duidelijk en helder waren geformuleerd. Met bovenstaande stappen werd de inhoudsvaliditeit vastgesteld. Om gegevens te verzamelen met betrekking tot de constructvaliditeit werd de vragenlijst afgenomen onder aios. De leerlingen werd gevraagd om op een vijfpunts-Likertschaal aan te geven in hoeverre zij het eens waren met een stelling (volledig oneens- volledig eens). Met het uitvoeren van de dataverzameling kan worden gestart met het bepalen van de constructvaliditeit.

#### 4.3.2 Constructvaliditeit van het ontwikkelde instrumentarium

Om de constructvaliditeit van het instrument te bepalen werden een factoranalyse uitgevoerd om het aantal factoren te bepalen en daarmee het aantal schalen. Daarnaast werd een correlatieanalyse uitgevoerd om de samenhang tussen de items te bepalen. De resultaten van deze analyses worden in deze paragraaf besproken.

In hoofdstuk (3) werd al beschreven dat er een aantal analyses hebben plaatsgevonden voordat met de uiteindelijke factoranalyse werd aangevangen. De werkwijze is beschreven in hoofdstuk (3), de resultaten van deze vooranalyses zullen worden beschreven in deze paragraaf.

De eerste analyse controleerde of er sprake was van veel 'missing values'. De missende waarden bleken op een aantal items zo hoog te zijn dat werd besloten om deze items te verwijderen, op basis van dit criterium werden er zeven items uit de lijst verwijderd. De volgende item werden verwijderd uit de lijst:

*14. Mijn begeleider legt van te voren de materialen klaar die voor het onderwijsmoment nodig zijn*

*15. Mijn begeleider weet hoe de apparatuur werkt die hij tijdens het onderwijsmoment gebruikt*

*24. Mijn begeleider stelt vragen aan alle leerlingen tijdens het onderwijs*

*25. Mijn begeleider zorgt dat iedere leerling een actieve rol heeft tijdens een onderwijsmoment*

*26. Mijn begeleider geeft de ruimte om leerlingen op elkaar te laten reageren*

*49. Mijn begeleider zorgt dat hij mij in diverse situaties toetst*

*50. Mijn begeleider zorgt voor KPB's door verschillende beoordelaars*

52

Hetzelfde werd gedaan voor respondenten met veel missende waarden. Er werden vier respondenten verwijderd uit het databestand. Bij zeven items is er sprake van ambiguïteit. Deze items worden verwijderd uit de lijst. Daarnaast werd ook besloten om de items (27) en (28) uit de lijst te verwijderen. Item (28) laadt heel hoog maar blijft op zichzelf staan wanneer deze in de lijst betrokken wordt en laadt opvallend niet op factor (3) of (5). Mogelijk wordt dit item verkeerd geïnterpreteerd door de respondenten. In vervolgonderzoek kan dit item wel een plaats innemen. Het item moet dan mogelijk anders worden geformuleerd, of het item kan door het genereren van nieuwe items een schaal vormen met andere items en op die manier toch worden gemeten. Item (27) brengt twijfel met zich mee. De vraag is of dit item niet meer een aspect is dat bij een persoon thuishoort. Kun je enthousiasme aanleren of is dit een persoonlijkheidsaspect die je wel bezit of niet? De cursusleiders proberen het enthousiasme voor het onderwijs over te brengen maar of dit als werkelijk effect van de cursusinhoud moet worden bestempeld is de vraag.

In dit onderzoek wordt het bestempeld als een persoonlijkheidsaspect. In later onderzoek, wanneer het onderdeel toch als onderdeel van de cursusinhoud wordt bestempeld, kan ervoor worden gekozen om meerdere items te genereren en om het onderdeel enthousiasme te meten en dat er mogelijk een schaal 'enthousiasme' kan worden gevormd. In deze opzet zorgt het voor een complexe factorstructuur en wordt het item verwijderd. Tot slot werd een betrouwbaarheidsanalyse uitgevoerd om te kijken of er items waren die bij verwijdering zouden zorgen voor een aanzienlijk hogere betrouwbaarheid van de schaal. In bijlage (9) zijn de resultaten van deze analyse opgenomen. Op basis van deze analyse wordt besloten item (23) te verwijderen. Omdat het item ook inhoudelijk niet goed bij de schaal past wordt dit item verwijderd.

Naar aanleiding van voorgaande analyses wordt besloten om de vragenlijst aan te passen en bovenstaande items te verwijderen. De uiteindelijke vragenlijst bestaat uit 33 vragen en bevat acht schalen, een overzicht van deze lijst is te vinden in bijlage (10).

In tabel 4.1 worden de resultaten van de factoranalyse weergegeven. Bij elke analyse die werd uitgevoerd in de aanloop naar de verwerking van de resultaten was te zien dat items vaak op dezelfde factor laden. Wanneer we dit bij verschillende analyses constateren kunnen we spreken van een stabiele factorstructuur. Hieronder zal een beschrijving worden gegeven van de schalen.

De eerste schaal (6 items) heeft betrekking op het creëren van een veilig leerklimaat. Het omvat items die betrekking hebben op de mate waarin leerlingen zich tijdens een onderwijsmoment op hun gemak voelen, of leerlingen zenuwachtig zijn voorafgaand aan een onderwijsmoment en of er ruimte is om de eigen mening te ventileren. Deze schaal wordt bestempeld als: '*creëren van een veilig leerklimaat*'.

In de tweede schaal (6 items) draait het om aandacht en betrokkenheid bij het ontwikkelingsproces van de leerling. Er wordt aandacht besteed aan het stellen van doelen. In de

vragen van schaal (2) komen deze terug als het opstellen van persoonlijke leerdoelen, monitoren aan de hand van persoonlijke leerdoelen en weten wat de persoonlijke leerdoelen zijn (item 2,1 en 3). De vragen suggereren ook een vorm van controle bij de leerling, wat heeft de leerling geleerd tijdens het onderwijsmoment? Wat is de mening van de leerling over het onderwerp? De items hebben allemaal te maken met het begeleiden van het ontwikkelingsproces van een toekomstig arts. De tweede schaal wordt genaamd: *'begeleiding van het ontwikkelingsproces van de leerling'*.

De derde schaal (6 items) van deze lijst draait duidelijk om de organisatie van het onderwijsmoment. Het is belangrijk om van te voren goed over de organisatie van een onderwijsmoment na te denken om ervoor te zorgen dat de les vlekkeloos verloopt en voldoende kennis wordt opgedaan. Het is belangrijk om van te voren goed na te denken over het doel van het onderwijsmoment. Er moet een duidelijke opbouw in het onderwijsmoment zitten zodat de leerlingen weten waar ze aan toe zijn. Voorbeelden uit de praktijk kunnen gebruikt worden om situaties te verduidelijken en daarnaast om de relevantie van het onderwijsmoment aan te geven. In de organisatie dient een duidelijke casus te worden gekozen. Op welke manier bied ik het studiemateriaal aan zodat het geen saaie voordracht wordt? (voldoende afwisseling). Tot slot blijft item (30) over. Ook het stellen van vragen kan behoren tot de organisatie van het onderwijsmoment van te voren dient goed te worden nagedacht over de vragen die de docent stelt tijdens het onderwijsmoment, om de leerlingen verder te brengen bij het begrijpen van de stof of uit te dagen tot verdieping. De docent kan zich ook voorbereiden op de vragen die de leerlingen mogelijk gaan stellen. Vragen kunnen een andere draai aan het onderwijsmoment geven, vragen kunnen voor verdieping zorgen of voor de brug naar een volgend onderwijsmoment daarom past dit item in deze schaal: *'organisatie van het onderwijsmoment'*. De vierde schaal (3 items) benadrukt het belang en bewustzijn van het rolmodel. Deze items benadrukken het belang van het bewustzijn van de arts als rolmodel. Een arts

dient zich constant bewust te zijn van de rolmodelfunctie. Deze schaal wordt bestempeld als: *'bewustzijn rolmodel'*.

De vijfde schaal (4 items) is moeilijk te interpreteren. Item (40) en (41) hebben te maken met het op tijd aanwezig zijn voor een afspraak en het nakomen van afspraken die gemaakt zijn, het neigt hier naar het opbouwen van vertrouwen en het creëren van gelijkwaardigheid tussen de leerling en medisch specialist. Item (4) en (5) hebben te maken met vertrouwen in het feit dat de leerling feedback krijgt waar hij zich verder mee kan ontwikkelen, waar je je verder mee kunt ontwikkelen. Item (4) geeft aan dat er sprake moet zijn van constructieve feedback en item (5) geeft aan dat de feedback ook concreet gemaakt moet worden zodat de leerling handvatten krijgt om te handelen met de feedback die wordt gegeven. Impliciet heeft dit natuurlijk te maken met het scheppen van een veilige sfeer. Deze schaal wordt bestempeld als: *'integriteit en vertrouwen'*.

Alle items in de zesde schaal (4 items) hebben betrekking op het feedbackgesprek. Ook hier draait het opnieuw om het creëren van een veilig leerklimaat, of het creëren van een veilige sfeer tijdens een gesprek. Deze schaal krijgt de volgende naam: *'creëren van veiligheid tijdens een feedbacksessie'*.

De zevende schaal (2 items) is moeilijk te interpreteren. Item (39) controleert of de stof van het onderwijsmoment begrepen is. Item (38) vat de stof samen van het onderwijsmoment. Het is moeilijk om een eenduidige naam te vinden voor deze items samen. De schaal wordt dan ook bestempeld als: *'controle en herhaling'*.

De achtste en tevens laatste schaal (2 items) van deze lijst heeft betrekking op het geven van een beoordeling, het draait hier om het toetsen van de leerling. De laatste schaal wordt bestempeld als het onderdeel *'toetsing'*.

Item in de vragenlijst	Factoren							
	1	2	3	4	5	6	7	8
34 Ik durf vragen te stellen tijdens een onderwijsmoment met mijn begeleider	<b>0,88</b>	0,13	0,09	-0,03	-0,07	0	0,02	0,01
35 Ik heb het idee dat ik tijdens een onderwijsmoment met mijn begeleider de ruimte krijg om mijn eigen mening te geven	<b>0,88</b>	0,09	0,1	0,04	0,05	0,01	0,11	0,04
36 Ik voel mij op mijn gemak tijdens een onderwijsmoment met mijn begeleider	<b>0,85</b>	0,13	0,04	0,24	-0,03	0,23	0,04	0,04
33 Ik heb het gevoel dat mijn begeleider tijdens een onderwijsmoment ruimte laat om vragen te stellen	<b>0,8</b>	0,21	0,26	-0,09	0,19	0,08	-0,07	0,03
37 Ik ben niet zenuwachtig voorafgaand aan een onderwijsmoment met mijn begeleider	<b>0,77</b>	-0,01	-0,07	0,34	-0,05	0,25	-0,02	-0,08
11 Mijn begeleider geeft mij tijdens een feedbacksessie de ruimte om mijn eigen mening te geven	<b>0,68</b>	0,19	0,23	0,05	0,37	0,31	0	0
2 Mijn begeleider monitort mijn leerproces aan de hand van mijn persoonlijke leerdoelen	0,21	<b>0,82</b>	0,19	-0,03	0,07	0,07	0,22	0,2
6 Mijn begeleider vraagt na een patiëntcontact om mijn mening te geven over het verloop van de handeling en/of het gesprek	0,33	<b>0,75</b>	0,15	0,23	0,13	0	-0,04	0,12
1 Mijn begeleider geeft het belang aan van het opstellen van persoonlijke leerdoelen	0,35	<b>0,65</b>	0,11	-0,06	0,16	0,1	0,48	0,21
3 Mijn begeleider weet wat mijn persoonlijke leerdoelen zijn	0,11	<b>0,65</b>	0,11	0,00	0,15	0,22	0,48	0,24
8 Mijn begeleider vraagt na afloop van een patiëntcontact wat ik ervan heb geleerd	-0,14	<b>0,64</b>	0,33	0,34	0,21	0,25	0	0,17
7 Mijn begeleider vraagt, hoe ik de tijdens een onderwijsmoment geleerde stof, in praktijk ga brengen	0,18	<b>0,61</b>	0,21	0,47	0,23	0,05	0,13	-0,08
17 Mijn begeleider geeft de relevantie aan van het onderwijsmoment voor de praktijk	0,11	0,13	<b>0,85</b>	0,21	0,06	-0,01	0,14	0,1
16 Mijn begeleider geeft duidelijk het doel aan van een onderwijsmoment	0,08	0,2	<b>0,81</b>	0,05	0,02	-0,03	0,03	0,13
18 Mijn begeleider gebruikt een casus om een probleem te verduidelijken	-0,04	0,25	<b>0,67</b>	-0,13	0,27	0,16	-0,05	-0,15
19 Mijn begeleider hanteert een duidelijke opbouw in zijn onderwijsmomenten (introductie, uitvoering, afronding)	0,26	0,28	<b>0,64</b>	-0,05	0,19	-0,05	0,38	0,27
31 Mijn begeleider zorgt voor voldoende afwisseling tijdens een onderwijsmoment	0,41	-0,06	<b>0,64</b>	0,03	-0,22	0,09	0,28	0,06
30 Mijn begeleider stelt verschillende soorten vragen (kennisvragen, begripsvragen, toepassingsvragen)	0,37	-0,18	<b>0,5</b>	-0,18	-0,04	0,35	0,46	-0,07
42 Mijn begeleider maakt bewust gebruik van zijn rolmodelfunctie als arts	0,22	0,02	0,01	<b>0,84</b>	0,08	0,13	0,11	0,35
43 Mijn begeleider maakt bewust gebruik van zijn rolmodelfunctie als opleider	0,11	0,05	0,04	<b>0,79</b>	0,16	0,23	0,04	0,36
44 Mijn begeleider maakt zijn rolmodelfunctie bespreekbaar	0,03	0,34	0,03	<b>0,77</b>	0,06	0,19	0,24	-0,06
4 Mijn begeleider zorgt voor constructieve feedback	0,1	0,42	0,12	0,08	<b>0,79</b>	0,14	0,11	0,01
41 Mijn begeleider is op tijd aanwezig voor bijvoorbeeld een gesprek/onwijsmoment	0,02	0,06	0,07	0,09	<b>0,76</b>	-0,02	-0,00	0,44
40 Mijn begeleider komt zijn afspraken na	0,14	-0,04	-0,05	0,22	<b>0,74</b>	0,21	0,35	0,25
5 Mijn begeleider zorgt dat hij de gegeven feedback concreetiseert	-0,05	0,51	0,14	0,05	<b>0,64</b>	0,14	-0,21	-0,12
12 Ik voel mij op mijn gemak tijdens een feedbacksessie met mijn begeleider	0,38	0,24	0,09	0,29	0,28	<b>0,73</b>	-0,07	-0,05
10 Ik durf vragen te stellen tijdens een feedbackgesprek met mijn begeleider	0,25	0,36	-0,04	0,23	-0,06	<b>0,7</b>	0	0,25
9 Mijn begeleider geeft mij tijdens een feedbacksessie de ruimte om vragen te stellen	0,25	0,35	0,1	0,21	0,17	<b>0,6</b>	-0,09	0,41

	1	2	3	4	5	6	7	8
39 Mijn begeleider controleert of de stof van een onderwijsmoment is begrepen	-0,06	0,11	0,14	0,15	-0,06	0	<b>0,86</b>	0,02
38 Mijn begeleider vat de stof die besproken is tijdens het onderwijsmoment nog een keer kort samen aan het einde	0	0,19	0,15	0,17	0,17	-0,08	<b>0,8</b>	0,03
46 Mijn begeleider zorgt voor eindbeoordelingen	-0,03	0,33	0,05	0,15	0,07	0,14	0,13	<b>0,8</b>
47 Mijn begeleider geeft duidelijk aan wanneer er sprake is van een tussentijdse- of eindbeoordeling	-0,03	0,07	0,18	0,28	0,26	0,12	0	<b>0,76</b>

**Tabel 4.1:** Resultaten factoranalyse: Vragenlijst effectevaluatie 'Teach the Teacher'

Voor elke factor werd een betrouwbaarheidsanalyse uitgevoerd door het berekenen van Cronbach's alpha. De resultaten van deze analyse zijn te vinden in tabel 4.2. Uit de gegevens in tabel 4.1 kan worden geconstateerd dat de schalen over het algemeen een hoge betrouwbaarheid laten zien. De alpha scores liggen in de range van 0,83-0,92.

Schaal	Aantal items	Betrouwbaarheid (Cronbach's $\alpha$ )
Creëren van een veilig leerklimaat	6	,92
Begeleiding van het ontwikkelingsproces van de leerling	6	,92
Organisatie van het onderwijsmoment	6	,86
Bewustzijn rolmodel	3	,87
Vertrouwen en integriteit	4	,83
Creëren van veiligheid tijdens een feedback	4	,86
Controle en herhaling	2	,84
toetsing	2	,86

**Tabel 4.2:** Betrouwbaarheid van de schalen: Vragenlijst effectevaluatie 'Teach the Teacher'

Tot slot werd de constructvaliditeit van de vragenlijst vastgesteld door de samenhang tussen de schalen te bestuderen. De resultaten van de analyse zijn te vinden in tabel 4.3. Alle subschalen hangen positief met elkaar samen. Wanneer alle scores positief met elkaar correleren kan er worden gesproken van het meten van één construct. In dit stadium kan er nog geen krachtige uitspraak worden gedaan met betrekking tot het meten van één construct, bovendien zijn enkele correlatiescores zeer laag en vertonen dus een zwak lineair verband. Lage correlatiescores kunnen erop wijzen dat er sprake is van het meten van meerdere deelconstructen. Het 'Teach the Teacher' programma beoogt te doceren over 'de goede docent'. De vraag is of alle aspecten van een goede docent in één construct te meten zijn. Er is geen sprake van hele hoge correlaties, de hoogste score is 0,54. Als er sprake is van hoge correlatiescores is er geen sprake meer van onderscheid tussen factoren en meten de variabelen vaak hetzelfde. Er worden een aantal significante verbanden tussen schalen gevonden, zestien van de 28 verbanden zijn significant, dit houdt in dat de scores niet op toeval berusten. De hoogste correlaties worden bereikt tussen schaal (2) en de overige schalen

Schalen	Schaal 1	Schaal 2	Schaal 3	Schaal 4	Schaal 5	Schaal 6	Schaal 7	Schaal 8
Schaal 1	1,00	,40	,51**	,29	,21	,49	,07	,06
Schaal 2		1,00	,50**	,54**	,52**	,25**	,41*	,48**
Schaal 3			1,00	,21	,26	,26	,38*	,30
Schaal 4				1,00	,37*	,52**	,28	,50**
Schaal 5					1,00	,40*	,18	,45**
Schaal 6						1,00	,06	,37*
Schaal 7							1,00	,25
Schaal 8								1,00

\* Correlatie is significant bij een  $\alpha$  van 0,05 (2-zijdig)  
\*\* Correlatie is significant bij een  $\alpha$  van 0,01 (2-zijdig)

**Tabel 4,3;** Correlaties tussen de schalen

Uit bovenstaande bevindingen kan worden opgemaakt dat het gelukt is om een basisinstrument te ontwikkelen voor het meten van het effect van het 'Teach the Teacher' programma module (1) dat voldoet aan de eisen van inhoudsvaliditeit, constructvaliditeit en betrouwbaarheid. Er wordt gesproken van een basisinstrument omdat nog niet aan alle eisen van validiteit is voldaan, de predictieve validiteit van het instrument is bijvoorbeeld nog niet berekend. Het is van groot belang dat het instrument verder wordt getest met een nieuwe steekproef om te kijken of dezelfde resultaten worden behaald. In het laatste hoofdstuk zullen de conclusies nader geanalyseerd worden en worden bediscussieerd.

richtinggevend konden dienen bij de ontwikkeling van het nieuwe instrument. Vervolgens werd bij vraagstelling (2) gestart met de ontwikkeling van het nieuwe instrument. Het is gelukt om een instrument te ontwikkelen wat aan de eisen van inhoudsvaliditeit, constructvaliditeit en betrouwbaarheid voldoet. In hoofdstuk 5 zullen de conclusie en de discussie van dit onderzoek worden beschreven.

#### 4.4 Samenvatting hoofdstuk 4

In dit hoofdstuk werden de resultaten van het onderzoek besproken. Vraagstelling (1) werd beantwoord door de uitvoering van een literatuurstudie.

Uiteindelijk bleken er geen instrumenten aanwezig die volledig geschikt waren om het effect van het 'Teach the Teacher' programma module (1) te meten. Wel heeft een selectie plaatsgevonden, op basis van het conceptueel kader, van items uit bestaande instrumenten die als



## 5 Conclusie en Discussie

### 5.1 Inleiding

Dit onderzoek geeft antwoord op de centrale onderzoeksvraag:

*Is het mogelijk om een valide en betrouwbaar instrument te ontwikkelen, gericht op het meten van de door de medisch specialist geleerde vaardigheden in module (1) van het 'Teach the Teacher' programma?*

Deze vraag werd aan het begin van het onderzoek opgesplitst in twee deelvragen. Paragraaf 5.2 en 5.3 bespreken de conclusies van de twee vraagstellingen. Paragraaf 5.4 gaf de discussie over de resultaten en methoden en de aanbevelingen voor verder onderzoek.

### 5.2 Conclusies vraagstelling 1

Deze paragraaf bespreekt de conclusies van de eerste vraagstelling. Deze vraagstelling luidt:

*In hoeverre zijn er instrumenten aanwezig om het effect van didactische basisvaardigheidstrainingen in het medisch onderwijs te bepalen en welke schalen uit deze instrumenten kunnen worden gebruikt voor het meten van het 'Teach the Teacher' programma module (1)?*

Uit het literatuuronderzoek kwam naar voren dat er geen geschikte instrumenten bestaan die volledig zijn te benutten voor het meten van het effect van het 'Teach the Teacher' programma module (1). Uiteindelijk werden er elf instrumenten geselecteerd die aan de criteria voldeden voor opname in de literatuurstudie. De overgebleven elf instrumenten werden uitgebreid beschreven en

bediscussieerd. Vanuit deze analyse werd ook bepaald welke items eventueel in aangepaste vorm bruikbaar zouden zijn voor de lijst. Op basis van het conceptueel kader dat aan het einde van hoofdstuk 2 werd geschetst werden er items uit de instrumenten geselecteerd.

Opvallend was dat vier van de elf geselecteerde instrumenten zijn gebaseerd op het SFDP (Stanford Faculty Development Program) raamwerk. Deze instrumenten vertonen dan ook veel overlap, vaak werden dezelfde items geselecteerd zodat de selectie niet veel nieuwe informatie opleverde. Eén van de instrumenten was afkomstig uit een andere sector, vanuit de Tandheelkunde en bleek uiteindelijk ook gebaseerd te zijn op het SFDP-raamwerk. Uit het onderzoek in de onderwijskundige databases kwamen geen nieuwe instrumenten voort die bruikbaar waren voor deze specifieke medische context. Ook juist vanwege deze specifieke context, en de manier van onderwijs geven (tijdens de dagelijkse patiëntenzorg), zijn de instrumenten uit het algemene onderwijs vaak niet geschikt voor een meting in het medisch onderwijs.

Concluderend kan men stellen dat er geen instrumenten geschikt zijn voor het meten van het effect van het 'Teach the Teacher' programma module (1). De instrumenten die voortkwamen uit de literatuurstudie bevatten een aantal bruikbare items die richtinggevend waren bij de ontwikkeling van het nieuwe instrument.

### 5.3 Conclusies vraagstelling 2

Deze paragraaf bespreekt de conclusies van de tweede deelvraag. Deze deelvraag luidt:

*Wat is de betrouwbaarheid en validiteit van het ontwikkelde instrument voor het bepalen van het effect van een training in didactische basisprincipes aan medisch specialisten?*

Vanuit de resultaten van vraagstelling 2 kunnen een aantal conclusies worden getrokken.

Met behulp van observaties en de bespreking van de items uit de vragenlijst met twee docenten van het 'Teach the Teacher' programma werd de inhoudsvaliditeit van het instrument vastgesteld.

De constructvaliditeit werd bevestigd door een factor- en correlatieanalyse uit te voeren.

Met het uitvoeren van een factoranalyse met Varimax rotatie met een Kaiser's criterium van 1,00 werd de constructvaliditeit bepaald. In totaal werden er acht factoren onderscheiden die 80,05% van de variantie verklaarden.

Er werden in totaal acht factoren onderscheiden die als volgt werden omschreven:

Schaal 1: *'creëren van een veilig leerklimaat'*

Schaal 2: *'begeleiding van het ontwikkelingsproces van de leerling'*

Schaal 3: *'organisatie van het onderwijsmoment'*

Schaal 4: *'bewustzijn rolmodel.'*

Schaal 5: *'integriteit en vertrouwen'*

Schaal 6: *'creëren van veiligheid tijdens een feedbacksessie'*

Schaal 7: *'controle en herhaling'*

Schaal 8: *'toetsing'*

Tot slot werd voor het bepalen van de constructvaliditeit een correlatieanalyse uitgevoerd tussen de schalen van de vragenlijst. Alle subschalen hangen positief met elkaar samen. De correlaties liggen in de range van 0,06-0,54. Wanneer alle scores positief met elkaar correleren kan er worden gesproken van het meten van één construct. In dit stadium kan er nog geen krachtige uitspraak worden gedaan met betrekking tot het meten van één construct, bovendien zijn enkele correlatiescores zeer laag en vertonen dus een zwak lineair verband. Lage correlaties kunnen erop wijzen dat er sprake is van het meten van meerdere deelconstructen. Er is geen sprake van hele hoge correlaties, de hoogste score is 0,54. Als er sprake is van hoge correlatiescores is er geen sprake meer van onderscheid tussen factoren en meten de variabelen vaak

hetzelfde. Het 'Teach the Teacher' programma module (1) beoogt te doceren over 'de goede docent'. De vraag is of alle aspecten van een goede docent in één construct te meten zijn. Er worden een aantal significante verbanden tussen schalen gevonden, zestien van de 28 verbanden zijn significant. De hoogste correlaties worden bereikt tussen schaal (2) en de overige schalen.

Naast het analyseren van de validiteit werd ook de betrouwbaarheid bestudeerd. Er werd een hoge betrouwbaarheid bereikt per schaal en voor het gehele instrument. Alle schaal scores liggen boven de 0,83. Concluderend kan worden gesteld dat het onderzoek er in is geslaagd een basisinstrument voort te brengen dat voldoet aan de eisen van inhoudsvaliditeit, constructvaliditeit en betrouwbaarheid. In de toekomst is het mogelijk om een instrument te ontwikkelen voor module (1) van het 'Teach the Teacher' programma dat voldoet aan alle eisen van validiteit en betrouwbaarheid. Het is daarnaast gelukt om een instrument te ontwikkelen dat op niveau (3): *'Evaluatie van gedrag'* van het evaluatiemodel van Kirkpatrick kan meten. Er zijn echter nog wel een aantal aandachtspunten waarbij rekening moet worden gehouden in de verdere ontwikkeling van het instrument en bij het lezen van de conclusies. In de volgende paragraaf zullen deze discussiepunten worden besproken.

## 5.4 Discussie

Zoals in de conclusies van dit onderzoek al werd weergegeven is het onderzoek erin geslaagd om een antwoord te geven op de centrale onderzoeksvraag. Echter zijn er nog een aantal discussiepunten waarmee rekening dient te worden gehouden wanneer de conclusies worden bestudeerd. In deze paragraaf zullen de methoden van onderzoek en de resultaten die zijn voortgekomen uit het onderzoek worden bediscussieerd.

#### 5.4.1 Discussie deelvraag 1 en aanbevelingen voor verder onderzoek

Uit het literatuuronderzoek kwamen uiteindelijk elf bruikbare instrumenten voort die geschikt bleken voor opname. In de vooranalyse bleken zeven andere instrumenten in eerste instantie ook geschikt voor opname maar bij nader onderzoek bleek er geen onderzoek beschikbaar over de psychometrische kwaliteiten van deze instrumenten. Dit is een opvallende bevinding omdat de medische wereld predikt te willen werken met 'evidence based methoden' (zie hoofdstuk 1). De verwachting is dat er in de toekomst mogelijk verder onderzoek zal worden verricht naar deze instrumenten daarom worden de instrumenten kort genoemd.

Het is opvallend dat er in de literatuur uiteindelijk weinig instrumenten worden gevonden voor het meten van de 'opleiderskwaliteiten van docenten', en bijzonder opvallend dat er geen instrumenten worden gevonden in de algemene onderwijssector. De verwachting was dat er in het regulier onderwijs meer instrumenten zouden worden gevonden die aan de psychometrische eisen voldeden. Ook in het onderwijs is steeds meer aandacht voor het werken met 'evidence based' methoden. Ook in het algemene onderwijsveld dienen dus instrumenten te worden ontwikkeld die voldoen aan de eisen van betrouwbaarheid en validiteit.

Bij het uitvoeren van de literatuurstudie is gebruik gemaakt van verschillende bronnen en methoden: verschillende databases zijn doorzocht, de 'sneeuwbalmethode' is gebruikt en experts in het medisch onderwijs zijn geraadpleegd. Mogelijk had de literatuurstudie nog verder kunnen worden uitgebreid met het gebruik van zogenaamde Medical Subject Heading Terms (MeSH). Dit systeem is overeenkomstig met de Thesaurus bij Gedrags- en Maatschappijwetenschappen. Omdat het medisch onderwijs, en specifiek het evalueren van artsen als docent, nog volop in ontwikkeling is was de verwachting dat er met dit systeem waarschijnlijk niet meer instrumenten zouden worden gevonden dan nu het geval is. Waarschijnlijk zijn de MeSH termen op dit gebied nog niet volledig uitgekristalliseerd en zal de studie mogelijk zelfs minder

resultaten opleveren. Daarentegen is er binnen het onderzoek gebruik gemaakt van verschillende informatiebronnen, dit zorgt al voor een hogere betrouwbaarheid van het onderzoek en het uitsluiten van missende data.

Op basis van het conceptueel kader vond een selectie van items uit de elf instrumenten plaats. Opvallend is dat bepaalde kernpunten uit het conceptueel kader een grote selectie van items oplevert en andere kernpunten niet. Criterium 1: '*Er moet sprake zijn van een duidelijke opbouw in het onderwijsmoment*', levert geen items uit de geselecteerde instrumenten op. Het kernpunt werd belangrijk geacht op basis van de onderwijskundige modellen van Gagné en Van Gelder en dit aandachtspunt komt ook terug in de cursusinhoud van het 'Teach the Teacher' programma en de theorie rond het leren van volwassenen, '*herkenbaar begrippenkader*'. Mogelijk is er bij de ontwikkeling van de instrumenten voornamelijk gebruik gemaakt van kennis en observaties vanuit de medisch onderwijskundige praktijk en minder op basis van modellen vanuit de onderwijskunde of gebruik gemaakt van andere onderwijskundige- en psychologische theorieën. Dolmans et al. (2004) meldden al eerder dat de SFDP-26 niet sterk theoretisch is gefundeerd. Aangezien vier van de elf instrumenten gebaseerd zijn op de SFDP-26 zijn deze in principe ook niet sterk theoretisch gefundeerd. Ook is er sprake van veel redundantie in de items die geselecteerd werden omdat deze instrumenten gebaseerd zijn op het SFDP-raamwerk, dit zien we terug bij criterium (2), (4), (5), (6), (8) van het conceptueel kader. Dat er verschillen zijn tussen de geselecteerde instrumenten en het ontwikkelde instrument kan mogelijk ook te wijden zijn aan verschillen in het medisch onderwijs tussen Nederland en de landen waar de instrumenten uit afkomstig zijn. Verder treedt er in ieder geval een verschil op met de andere instrumenten in inhoud van de items omdat specifiek items werden opgesteld die betrekking hadden op de cursusinhoud. Vaak treedt er een verschil op tussen hoe vaardigheden worden aangeleerd in de cursus en hoe het uiteindelijk wordt uitgevoerd op de werkvloer, bepaalde

vaardigheden worden geabsorbeerd en toegepast andere vaardigheden worden niet belangrijk genoeg geacht en niet toegepast op de werkvloer.

#### 5.4.2 Discussie deelvraag 2 en aanbevelingen voor verder onderzoek

Met betrekking tot de data die werd verzameld voor dit onderzoek onder aios, co-assistenten en anios, kan het volgende worden opgemerkt. In de analyse werd voor de missende waarden gebruik gemaakt van een schatting. Bij de missende waarden werd het gemiddelde op het item ingevuld, er van uitgaande dat dit ongeveer de trend is, en dat de respondent met de missende waarde dit antwoord mogelijk zou hebben gegeven. Bij een factoranalyse mag eigenlijk geen gebruik worden gemaakt van geïmputeerde data daarom is het van belang om hier bij het lezen van de conclusies rekening te houden. In een volgend onderzoek kan ervoor gekozen worden om de optie: 'kan ik niet beoordelen' weg te laten. Dan zal er wel uitsluitend gewerkt moeten worden met leerlingen (aios) die al voldoende tijd met de medisch specialist hebben doorgebracht om een goede beoordeling te kunnen geven van de items.

In de vooranalyse zijn een aantal items verwijderd die veel missende waarden vertoonden. De 'optie kan ik niet beoordelen' is aangewezen als missende waarde bij de codering van het bestand. In vervolgonderzoek moet worden geanalyseerd wat de oorzaak is dat de respondenten deze items vaak als 'kan ik niet beoordelen' hebben ingevuld. Het kan twee redenen hebben dat de optie 'kan ik niet beoordelen' is ingevuld. De oorzaak kan bij de respondent liggen die mogelijk te kort met de medisch specialist heeft gewerkt om een objectief oordeel te geven. De tweede verklaring kan zijn dat het item niet duidelijk is geformuleerd en dat er mogelijk onduidelijkheid is over wat er beoogt wordt met het item. Vanuit de schriftelijke opmerkingen die bij de vragenlijst werden geplaatst bleek dat een enkele leerling een te korte periode met een medisch specialist had doorgebracht om

alle vragen goed te kunnen beoordelen, de leerling had dan geen idee of een medisch specialist de vaardigheid toepaste op de werkvloer of niet. In vervolgonderzoek dient rekening te worden gehouden met de keuze voor de beoordelaar. Mogelijk is deze lijst niet geschikt voor beoordeling door een co-assistent. Omdat de co-assistent soms een zeer korte periode met een medisch specialist meeloopt kan dit een aanwijzing zijn om er in een volgend onderzoek voor te kiezen om alleen van beoordelingen gebruik te maken door aios.

Het probleem van de missende waarden kan ook liggen op het niveau van het item, de vraag is niet duidelijk genoeg geformuleerd en de respondent weet niet wat de onderzoeker bedoeld. Mogelijk moet volgende keer wanneer de lijst verder is uitgekristalliseerd en wanneer gebruik wordt gemaakt van één type respondent: de aios, de optie 'kan ik niet beoordelen' uit de lijst worden gehaald. Dan moet wel rekening gehouden worden dat wanneer de leerling een item eigenlijk niet kan beoordelen mogelijk invult met zo maar een score, op deze manier kan er nog een vertekend beeld ontstaan en kunnen er uiteindelijk verkeerde conclusies worden getrokken. Mogelijk worden de problemen deels al opgelost wanneer alleen aios worden gevraagd van wie de opleider het 'Teach the Teacher' programma heeft gevolgd. Dan wordt de inhoud van de items ook duidelijker voor de aios. De medisch specialist zal dan meer gedrag laten zien dat in de cursus werd geleerd, dan nu het geval was. De items die nu werden verwijderd kunnen in vervolgonderzoek mogelijk nog worden toegevoegd aan de lijst om te kijken of dezelfde resultaten worden bereikt. Vanuit de schriftelijke opmerkingen bij de vragenlijst komen mogelijke verklaringen voort waarom de verwijderde items niet werden ingevuld. Een aantal respondenten geeft aan dat het onderwijs vaak niet gepland is, dat er sprake is van informele onderwijsmomenten en dat onderwijsmomenten vaak op een onverwacht moment plaatsvinden. Uit de opmerkingen blijkt dat de onderwijspraktijk nog vaak anders wordt ingevuld dan in het 'Teach the Teacher' programma wordt aangeleerd. In deze onderzoeksgroep is ook gebruik gemaakt van respondenten waarvan niet duidelijk is of de medisch

specialist de cursus wel of niet heeft gevolgd. In een volgend onderzoek dient hier duidelijk wel rekening mee te worden gehouden. Bij de ingevulde vragenlijsten werden ook een aantal opmerkingen over de KPB's geplaatst. Mogelijk worden KPB's wel meer gecontroleerd bij leerlingen waarvan de specialist wel de cursus heeft gevolgd. In het 'Teach the Teacher' programma wordt het belang van de KPB uitgelegd en ook verteld hoe deze gebruikt dient te worden. Uit de opmerkingen blijkt dat het belang van KPB's duidelijk nog niet is doorgedrongen op elke afdeling en dat ook de medisch specialist deels verantwoordelijk is voor het feit dat de KPB's voldoende worden ingevuld. KPB's kunnen van groot belang zijn voor de groei en ontwikkeling van een leerling. De leerling hoort op tijd waar de komende weken aan moet worden gewerkt. Een KPB kan gebruikt worden voor tussentijdse beoordeling. Vanuit de opmerkingen wordt duidelijk dat de gedachte heerst dat de leerling zelf verantwoordelijk is voor de beoordeling.

Bij de vooranalyse werden een aantal items verwijderd omdat deze ambiguïteit vertoonden. Vervolgens werd een betrouwbaarheidsanalyse uitgevoerd en geanalyseerd of er items waren die verwijderd konden worden zodat de schaal voor een hogere alpha zou zorgen. Uiteindelijk werd er één item, item 23, verwijderd. Ook item 44 zou zorgen voor een hogere betrouwbaarheid wanneer dit item werd verwijderd. De alpha zou dan van 0,87 naar 0,96 worden verhoogd. Omdat het item goed in de schaal paste is het uiteindelijk niet verwijderd maar in een volgende analyse kan mogelijk gekozen worden om het item toch nog te verwijderen.

Om de inhoudsvaliditeit te verhogen had in het eerste stadium een analyse plaats kunnen vinden voor de mate van relevantie van de opgestelde items. Docenten binnen het 'Teach the Teacher' programma hadden benaderd kunnen worden met de vraag of zij op een driepunts-schaal konden aangeven wat de relevantie van het item was ten opzichte van de cursusinhoud. Daarentegen heeft er nu overleg plaatsgevonden met twee verschillende

cursusleiders en is uiteindelijk ook een goede inhoudsvaliditeit bereikt met betrekking tot de items in de vragenlijst.

De steekproef was zeer klein. Uiteindelijk zijn er 44 enquêtes beantwoord. Van deze 44 respondenten vielen uiteindelijk nog vier respondenten af die veel missende waarden vertoonden. Voor een factoranalyse dienen er in ieder geval meer respondenten te zijn dan items in de vragenlijst. Field (2005) geeft aan dat een steekproef van 300 meestal afdoende is voor een factoranalyse. In het onderzoek had ook de keuze gemaakt kunnen worden om de eerste dataverzameling gelijk uit te voeren onder aios in het UMCG en daarna nog een aantal 'reminders' te sturen. Deze opzet had mogelijk meer data opgeleverd in het eerste stadium. In eerste instantie is uitgegaan van de ideale situatie om in het beginstadium data te verzamelen onder de aios die bij een medisch specialist werden opgeleid die het 'Teach the Teacher' programma module (1) hebben gevolgd. Helaas leverde dit heel weinig data op. In vervolgonderzoek dient hier rekening mee te worden gehouden. De aios zijn een lastige doelgroep om te bereiken. De aios worden met veel verzoeken benaderd voor het invullen van enquêtes. Mogelijk dient het beleid binnen het UMCG hier ook te worden aangepast zodat de respons voor onderzoek weer zal verhogen. Ondanks dat er van te voren rekening werd gehouden met de 'validiteitsbedreiger': non-respons was het responspercentage toch teleurstellend. Bij vrijwel alle methoden van dataverzameling krijgt de onderzoeker te maken met non-respons (Baarda en De Goede, 2006). "Van non-respons is sprake als niet van alle eenheden die tot de samengestelde steekproef behoren de benodigde antwoorden kunnen worden verkregen. We kunnen een steekproef nog zo zorgvuldig proberen samen te stellen, in de sociale wetenschappen stuiten we bijna altijd op het probleem van de non-respons: sommigen, soms velen, doen aan de zo mooi getrokken steekproef niet mee" (Swanborn, 2002, p.155). Eerst moeten de, in dit geval aios of co-assistenten, gevraagd worden om mee te doen aan het onderzoek en daarna moeten de respondenten ook

werkelijk de vragenlijst gaan invullen. Veel mensen hebben geen behoefte om mee te doen aan een onderzoek. Het gevolg kan zijn dat de mensen die wel meedoen tot een selecte groep behoren en daardoor kan uiteindelijk een vertekend beeld ontstaan in de uitkomsten van het onderzoek. Wanneer je ergens een uitgesproken mening over hebt, in positieve of negatieve zin, is een persoon vaker geneigd om zijn mening te geven dan wanneer hij geen uitgesproken mening heeft over een bepaald onderwerp (Baarda en De Goede, 2006). Bij de eerste ronde van dataverzameling werd geprobeerd de non-respons te beperken door een presentatie over het onderzoek te geven in de cursussen. De medisch specialist heeft de onderzoeker gezien en de ruimte gekregen om persoonlijk vragen te stellen aan de onderzoeker. Op deze manier werd de medisch specialist geënthousiasmeerd voor het onderzoek en daarmee ook zijn aios of co-assistenten. Daarnaast wordt in de cursus officieel toestemming gevraagd aan de medisch specialist om zichzelf te laten beoordelen door een aios of co-assistent, op deze manier ontstaan achteraf geen misverstanden wanneer een aios of co-assistent de medisch specialist beoordeelde. Daarnaast wordt er toestemming gevraagd om de secretariaten van de afdeling waar de medisch specialist werkzaam is te benaderen zodat de medisch specialist zo weinig mogelijk zelf moest regelen voor het onderzoek. Dit proces werd ondersteund door het opstellen van brieven voor het secretariaat en voor de aios of co-assistent, op deze manier werden de doelgroepen van het onderzoek op de hoogte gesteld van het onderzoek en de onderzoeksprocedure.

Een aantal dagen voor de sluitingsdatum van de dataverzameling hebben de aios, co-assistenten en medisch specialisten een herinnering per e-mail ontvangen met de oproep om de lijst alsnog in te vullen. Om het invullen van de vragenlijst te vergemakkelijken is de lijst digitaal gezet en via een online link in de e-mail te bereiken. Bij de tweede ronde van dataverzameling is een begeleidende brief opgesteld door een hoogleraar Medisch Onderwijs, omdat de naam van de student Onderwijskunde niet bekend was in het UMCG zou dit mogelijk tot minder respons kunnen leiden.

Daarnaast werd bij de vragenlijsten opgemerkt dat enkele respondenten moeite hadden met het invullen van de naam van de medisch specialisten die werd geëvalueerd. Mogelijk kunnen deze opmerkingen deels in verband worden gebracht met de hoge non-respons. In vervolgonderzoek dient dit punt mee te worden genomen. In eerste instantie zou ook de betrouwbaarheid worden berekend per individuele opleider, hoeveel vragenlijsten zijn er nodig per opleider om een betrouwbare score te bereiken? Daarom werd ook de naam van de medisch specialist gevraagd. Om te zorgen dat er sprake is van anonimiteit binnen dit onderzoek is ervoor gekozen om geen scores per specialisme te berekenen. In volgend onderzoek met een grotere steekproef kan dit mogelijk wel worden bereikt.

Wanneer er gebruik wordt gemaakt van een grotere steekproef zal het ook mogelijk zijn om meerdere metingen (meer aios per opleider) te bereiken en de lijst te laten invullen voor één specifieke opleider. Op deze manier kan in volgend onderzoek ook een uitspraak worden gedaan over hoeveel beoordelaars er nodig zijn per opleider om een betrouwbare meting te verrichten. Helaas werden er in dit onderzoek maar bij twee medisch specialisten meerdere formulieren, in totaal twee formulieren bij twee medisch specialisten, ingevuld waardoor er geen score kon worden berekend.

In eerste instantie werd gewerkt met een onderzoeksgroep waarvan de onderzoeker zeker wist dat de medisch specialist het 'Teach the Teacher' programma had gevolgd. In de tweede vorm van dataverzameling werden aios binnen het UMCG benaderd om de medisch specialist te beoordelen maar waarvan niet duidelijk was of de specialist het 'Teach the Teacher' programma had gevolgd of niet. Mogelijk treedt hierbij ook een vertekend beeld op in de data. Daarentegen wordt ervan uitgegaan dat iedere goede docent in principe over de vaardigheden die in het 'Teach the Teacher' programma worden aangeleerd moet beschikken. Wanneer deze analyse wordt herhaald moet gebruik worden gemaakt van een grotere steekproef en

gebruik worden gemaakt van specialisten die het 'Teach the Teacher' programma hebben gevolgd.

In de eerste vorm van dataverzameling werd ook de medisch specialist gevraagd om een zelfevaluatie uit te voeren. De medisch specialist diende zichzelf dan te beoordelen op het geleerde in de cursus. De respons viel ook binnen deze vorm van dataverzameling tegen. Vaak was er wel een vragenlijst ingevuld door de medisch specialist zelf en niet door de leerling of andersom. Zodoende konden er geen uitspraken worden gedaan over de vergelijking tussen de scores van de medisch specialist zelf en de leerling. In vervolgonderzoek kan dit mogelijk wel worden meegenomen en worden geanalyseerd.

Uiteindelijk werd de factoranalyse uitgevoerd en daaruit bleek dat sommige kernpunten uit het conceptueel kader wel terug kwamen in de factoren en andere kernpunten niet of deze waren terug te vinden onder andere factoren. De kernpunten die duidelijk terugkomen in de lijst zijn: '*het creëren van veiligheid*', in de lijst valt het zelfs uiteen in veiligheid tijdens een feedbacksessie en veiligheid tijdens het onderwijs. '*bewustzijn van het rolmodel*' krijgt een plaats in het instrument. De '*organisatie van het onderwijsmoment*' komt terug als factor. Onder deze factor wordt in de items aandacht besteed aan 'het stellen van doelen' en een 'duidelijke opbouw in het onderwijsmoment'. Deze twee kernpunten worden hier als onderdeel gezien van de organisatie van een onderwijsmoment. 'Evaluatie van het onderwijsmoment' komt terug in de factor '*toetsing*'. Het geven van systematische feedback komt terug onder de factor '*vertrouwen en integriteit*'. Allen criterium 3: '*aansluiten bij de voorkennis van de leerling*' krijgt geen plaats in het instrument. Mogelijk komt hier de theorie vanuit de onderwijskundige modellen niet overeen met de werkelijke medische onderwijssituatie.

In dit onderzoek werden alleen de inhouds- en constructvaliditeit bepaald. In het eerste stadium zou ook de criteriumvaliditeit van het instrument worden bepaald. Helaas doorkruisten praktische punten en het tijdsbestek

voor dit onderzoek met betrekking tot de analyses, al snel het bepalen van de criteriumvaliditeit. In volgend onderzoek zal gelijk in het beginstadium een nulmeting moeten worden verricht zodat op een later tijdstip de criteriumvaliditeit kan worden bepaald. De optimale situatie is om meerdere metingen te verrichten bijvoorbeeld voor de cursus, kort na de cursus en een langere periode na de cursus. Op deze manier kan worden geanalyseerd wat de effecten van de cursus op korte- en lange termijn zijn. In zo'n onderzoeksopzet dient wel te worden nagedacht over het verbinden van medisch specialisten en de bijbehorende leerlingen aan het onderzoek. Hoe kan ervoor worden gezorgd dat leerlingen en medisch specialisten zich voor een langere tijd aan het onderzoek committeren?

Uit de onderzoeksresultaten kan geconcludeerd worden dat ondanks het feit dat er betrouwbare- en valide metingen zijn bereikt voorzichtig met de resultaten moet worden omgesprongen. Met het oog op de missende waarden en de geïmputeerde data dient de lezer zich bewust te zijn van het gebruik van schattingen. In vervolgonderzoek dienen missende waarden te worden voorkomen. De items die werden verwijderd kunnen in vervolgonderzoek eventueel worden toegevoegd aan de bestaande vragenlijst om te kijken of dezelfde resultaten worden bereikt. Er wordt aangeraden om de geconstrueerde vragenlijst nogmaals af te nemen in de praktijk met een grotere steekproef. Wanneer opnieuw betrouwbare- en valide scores worden bereikt kan de effectmeting met betrekking tot de cursus worden uitgevoerd.

## 5.4 Samenvatting hoofdstuk 5

In dit laatste hoofdstuk werden de conclusies en discussie van deze masterthesis beschreven. Het onderzoek is erin geslaagd om een betrouwbaar- en valide basisinstrument voort te brengen voor het meten van het effect van het 'Teach the Teacher' programma module (1). Vanuit de

discussie blijkt echter wel dat de onderzoeksresultaten in het licht van de discussiepunten moeten worden bestudeerd en dat er verder onderzoek plaats zal moeten vinden voordat het instrument geschikt is om werkelijk in de praktijk te gaan gebruiken. De hoofdpunten die uit de discussie naar voren kwamen was dat er in vervolgonderzoek voorkomen moet worden dat er sprake is van veel missende waarden zodat er geen gebruik gemaakt hoeft te worden van geïmputeerde data. De items die werden verwijderd kunnen in vervolgonderzoek eventueel worden toegevoegd aan de bestaande vragenlijst om te kijken of dezelfde resultaten worden bereikt. Daarnaast dient het instrument meerdere malen te worden toegepast in het praktijk om te kijken of dezelfde valide- en betrouwbare resultaten worden bereikt.



## Literatuurlijst

Abu-Zidan, F.M., & Premadasa, I.G. (2002). Instructional skills of surgical tutors. *Singapore Medical Journal*, 43(12), 610-613.

Althouse, L.A., Stritter, F.T., & Steiner, B.D. (1999). Attitudes and approaches of influential role models in clinical education. *Advances in health sciences education*, 4, 111-122.

Aronson, D.T., & Briggs, L.J. (1983). *Contributions of Gagné and Briggs to a prescriptive model of instruction*. In Reigeluth, C.M. (Ed.). *Instructional design theories and models: An overview of their current status*, (p. 75-100). Hillsdale: Lawrence Erlbaum Associates publishers.

Baarda, B.D. & Goede, M.P.M. de (2001). *Basisboek methoden en technieken. Handleiding voor het opzetten van onderzoek*. Groningen: Stenfert Kroese.

Baarda, B.D. & Goede, M.P.M. de. (2006). *Basisboek methoden en technieken. Handleiding voor het opzetten en uitvoeren van kwantitatief onderzoek*. Groningen: Wolters-Noordhoff.

Bandura, A. (1977). *Social learning theory*. New Jersey: Prentice-Hall.

Beckman, T.J., Lee, M.C., Rohren, C.H., & Pankratz, V.S. (2003). Evaluating an instrument for the peer review of inpatient teaching. *Medical teacher*, 25(2), 131-135.

Bernaards, C.A. (2000). *Proefschrift: Nonresponse and factor analysis, with applications to rating scale data*. Utrecht: Universiteit Utrecht, faculteit Sociale Wetenschappen.

Boendermaker, P.M. (2003). *Meesterschap, van verkenning naar herkenning van de goede huisartsopleider*. Maarsse: Elsevier gezondheidszorg.

Boendermaker, P.M., & Rimmelts, P. (2008). De CTEI: kwaliteit of kwantiteit of van allebei een beetje? Zorgen over een veel gebruikt meetinstrument voor de docentkwaliteit van de klinische docenten. *Tijdschrift voor medisch onderwijs*, 27(3), 147-149.

Boeree, G. (z.j.) Albert Bandura 1925-present. Geraadpleegd op zondag 22 juni 2008 via URL: <http://webpace.ship.edu/cgboer/bandura.html>

Bolhuis, S.M. (1988). *Leren*. In: Gent, B, van. & Notten, A.L.T. (Red.), *Inleiding tot de volwasseneducatie*. (p. 69-85). Meppel/Amsterdam: uitgeverij Boom.

Boon, T. van den & Geerts, G. (1999). *Van Dale, groot woordenboek der Nederlandse taal*. Utrecht/Antwerpen: Van Dale lexicografie.

Website bouwstenenonline, geraadpleegd op donderdag 21 augustus 2008, 15.29 uur via URL: <http://studion.fss.uu.nl/Bouwstenenonline/3a7factoranalyse.doc>

Buchel, T.L., & Edwards, F.D. (2005). Characteristics of effective clinical teachers. *Family Medicine*. 37(1), 30-35.

Busari, J.O., Scherpbier, A.J.J.A., Vleuten, C.P.M, van der., & Essed, G.G.M. (2006). A two-day Teacher-Training programme for medical residents: investigating the impact on teaching ability. *Advances in health sciences education*. 11(2),133-144.

Businessballs, (z.j.). Kirkpatrick's learning and training evaluation theory. Geraadpleegd op: maandag 23 juni, 2008 om 10:40 uur via URL: <http://www.businessballs.com/kirkpatricklearningevaluatiomodel.htm>

Kaderbesluit CCMS 2004, Website KNMG, geraadpleegd op: zaterdag 10 mei 2008, 11.38 uur via URL: [http://knmg.artsennet.nl/content/resources/AMGATE\\_6059\\_100\\_TICH\\_L698161609/AMGATE\\_6059\\_100\\_TICH\\_R1560811149673000//](http://knmg.artsennet.nl/content/resources/AMGATE_6059_100_TICH_L698161609/AMGATE_6059_100_TICH_R1560811149673000//)

Cohen, R., MacRae, H., en Jamieson, C. (1996). Teaching effectiveness of surgeons. *The American Journal of Surgery*, 171, 612-614.

Copeland, H.L., & Hewson, M. (1999). *Clinical teaching effectiveness instrument: Development and Psychometric testing*, Paper presented at the annual meeting of the American educational research Association (Montreal, Quebec, April, 1999).

Copeland, H.L., & Hewson, M.G. (2000). Developing and testing an instrument to measure the effectiveness of clinical teaching in an academic medical center, *Academic medicine*. 75,(2), 161-166.

Cramer, G. & Van der Kamp, M. (1990). *Forum 5. Leren op latere leeftijd*. Amsterdam/Lisse: Swets en Zeitlinger B.V.

DeVellis, R.F. (2003). *Scale development, theory and applications* (2<sup>e</sup> editie). Thousand Oaks/ London/New Delhi: Sage publications.

Dolmans, D.H.J.M., Wolfhagen, H.A.P., Gerver, W.J., Grave, W, de., & Scherpbier, A.J.J.A. (2004). Providing physicians with feedback on how they supervise students during patient contacts. *Medical teacher*, 26(5), 409-414.

Dolmans, D.H.J.M., Stalmeijer, R.E., Wolfhagen, H.A.P. (2006). Feedback ter bevordering van de professionele ontwikkeling van docenten. *Tijdschrift voor medisch onderwijs*, 25(6), 267-272.

Donnelly, M.B., Woolliscroft, J.O., (1989). Evaluation of clinical instructors by third-year medical students. *Academic Medicine*. 64, 159-164.

Field, A.P. (2005). *Discovering statistics using SPSS* (2e editie). London: Sage publishers.

Flikweert, E.R., Miedema, M.T., & Briët, J.W. (2007). Beoordelen van opleiders door arts-Assistenten. *Tijdschrift voor medisch onderwijs*, 26,( 1) , p. 3-10

Frank, J.R. (Ed.) (2005). *The CanMeds 2005 physician competency framework, better standards, better physicians, better care*. Ottawa: The Royal College of Physicians and Surgeons of Canada.

Gelder, L, van., Oudkerk Pool, T., Peters, J., & Sixma, J. (red.) (1979). *Didactische analyse, studieboek*. Groningen: Wolters-Noordhoff.

Gibson, S.K. (2004). Social learning (cognitive) theory and implications for human resource development. *Advances in developing human resources*, 6 (2), 193-210.

Glaser, R. (1962). Psychology and institutional technology. In R. Glaser (red.), *Training research and education* (pp. 1-30). Pittsburgh: University of Pittsburgh Press.

Godfrey, J., Dennick, R., & Welsh, C. (2004). Training the trainers: do teaching courses develop teaching skills? *Medical education*, 38, 844-847.

Haan, M. de., Boendermaker, P.M., & Heij, I. (red.) (2002). *Het medisch ambacht, opleiden en leren in de praktijk van de (verpleeg)huisarts*. Maarssen: Elsevier gezondheidszorg.

Hamdy, H., Williams, R., Tekian, A., Benjamin, S., El Shazali, H., Bandaranayake, R. (2001). Application of "VITALS": Visual indicators of teaching and learning success in reporting student evaluations of clinical teachers. *Education for health, 14*(2),267-276.

Harden, R.M., & Crosby, J. (2000). AMEE Guide no 20: The good teacher is more than a lecturer- the twelve roles of the teacher. *Medical teacher, 22*(4),334-347.

Harden, R.M., Grant, J., Buckley, G., & Hart, I.R. (2000). Best evidence medical education. *Advances in Health Sciences Education, 5* (1), p. 71-90.

Hem-Stokroos, H.H, van der., Vleuten, C.P.M, van der., Daelmans, H.E.M., Haarman, H.J.T.M., & Scherpbier, A.J.J.A. (2005). Reliability of the clinical teaching effectiveness instrument. *Medical education, 39*, 904-910.

Hewson, M.G.A.B., & Jensen, N.M. (1990). An inventory to improve clinical teaching in the general internal medicine clinic. *Medical education, 24*, 518-527.

Irby, D., & Rakestraw, P.(1981). Evaluating clinical teaching in medicine. *Journal of medical education, 56*(3),181-186.

James, P.A., & Osborne, J.W. (1999). A measure of medical instructional quality in ambulatory settings: The MedIQ. *Family Medicine, 31*(4), 263-9.

James,P.A., Kreiter,C.D., Shipengrover, J., & Crosson, J. (2002). Identifying the attributes of instructional quality in ambulatory teaching sites: A validation study of the MedEd IQ. *Family medicine, 34* (4),268-73.

Jellema, F. (2003). Measuring training effects: the potential of 360-degree feedback. Enschede: Twente University press.

Johnson, N.R. Chen , J., (2006). Medical students evaluation of teaching quality between obstetrics and gynecology residents and faculty as clinical preceptors in ambulatory

gynecology, *American journal of obstetrics and Gynecology, 195*, 1479-83.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kallenberg, A.J., Grijspaarde, L. van der., Braak, A. ter., Horzen, C.J. van. (2003). *Leren (en) doceren in het hoger onderwijs*. Utrecht: Uitgeverij EMMA BV.

Kirkpatrick, D.L. (1996). Techniques for evaluating training programs. *Training and development, 50*(1), 54-59.

Website KNMG, herziening Nederlandse Artseneed, samenwerking VSNU & KNMG, augustus 2003, geraadpleegd op zaterdag 14 juni 2008, 12.00 uur via URL: [http://knmg.artsennet.nl/uri/?uri=AMGATE\\_6059\\_100\\_TI\\_CH\\_R1290401169290340](http://knmg.artsennet.nl/uri/?uri=AMGATE_6059_100_TI_CH_R1290401169290340)

Website KNMG, II.04 Nederlandse Artseneed,richtlijn, 2004, geraadpleegd op maandag 16 juni 2008, 20.41 uur via URL: [http://knmg.lad.artsennet.nl/uri/?uri=AMGATE\\_6059\\_100\\_TICH\\_R155017769634208&xsl=AMGATE\\_6059\\_100\\_TI\\_CH\\_R15495243970383&PHPSESSID=f5781e5e48d417e236020880c06ffb6f](http://knmg.lad.artsennet.nl/uri/?uri=AMGATE_6059_100_TICH_R155017769634208&xsl=AMGATE_6059_100_TI_CH_R15495243970383&PHPSESSID=f5781e5e48d417e236020880c06ffb6f)

Knowles, M.S., Holton III, E.F., & Swanson, R.A. (2005). *The adult learner. The definitive classic in adult education and human resource development*. San Diego: Elsevier.

Kyriacou, C. (1991). *Essential teaching skills*. Oxford: Blackwell Education.

Litzelman, D.K., & Shea, J.A. (2005). Chapter 7: Evaluation of the clerkship: Clinical teachers and program. In: Guidebook for Clerkship Directors, 3rd Edition, Fincher,R.M.E. (Eds.) geraadpleegd op: Zaterdag 19 juli 2008, 15:07 uur via URL: <http://familymed.uthscsa.edu/ACE/guidebook.htm>

Litzelman, D.K., Stratos, G.A., Marriott, D.J., Skeff, K.M. (1998). Factorial validation of a widely disseminated educational framework for evaluating clinical teachers. *Academic medicine*, 73(6), 688-695.

Lombarts, M.J.M.H., Bucx, M.J.L., Rupp, I., Keijzers, P.J., Kokke, S.I.M., & Schlack, W. (2007). Een instrument voor de evaluatie van opleiderskwaliteiten van stafleden, *Nederlands Tijdschrift voor de Geneeskunde*, 151(36), 2004-8.

Malling, B., Bested, K.M., Skjelsager, K., Ostergaard, H.T., & Ringsted, C. (2007). Long-term effect of a course on in-training assessment in postgraduate specialist education. *Medical teacher*, 29, 966-971.

Maslowski, R. (2001). *School culture and School performance*. Enschede: Twente University Press.

McGrath, C., Wai Kit Yeung, R., Comfort, M.B., McMillan, A.S. (2005). Development and evaluation of a questionnaire to evaluate clinical dental teachers (ECDT). *British Dental Journal*, 198(1), 45-48.

Website Medische vervolgopleidingen, geraadpleegd op zaterdag 10 mei 2008, 11.40 uur via URL:  
<http://www.medischevervolgopleidingen.nl/pages/homepage/S2/homepage.aspx>

Metz, J.C.M., Scherpbier, A.J.J.A., & Vleuten, C.P.M., van der. (1995). *Medisch onderwijs in beweging*. In: Metz, J.C.M., Scherpbier, A.J.J.A., & Van der Vleuten, C.P.M. (red.), *Medisch onderwijs in de praktijk* (p. 5-15). Assen: Van Gorcum.

Nijman, D.J.M. (2004). *Supporting transfer of training: Effects of the supervisor*. Enschede: Print Partners Ipskamp.

Notzer, N., & Abramovitz, R. (2008). Can brief workshops improve clinical instruction? *Medical education*, 42, 152-156.

Website Open Universiteit Nederland, basis kwalificatie onderwijs, geraadpleegd op: Maandag 16 juni 2008, 11:47 uur via URL:  
<http://www.ou.nl/eCache/EMO/1/93/228.html>

Oppenheimer, L., Keely, E., & Marks, M. (2006). An encounter card to evaluate teachers in clerkship. *Medical education*, 40, 459-489.

Website Orde van medisch specialisten, geraadpleegd op zaterdag 10 mei 2008, 12.00 uur via URL:  
[http://orde.artsennet.nl/content/resources/AMGATE\\_605\\_9\\_397\\_TICH\\_L39759101/AMGATE\\_6059\\_397\\_TICH\\_R14\\_31721314139520//](http://orde.artsennet.nl/content/resources/AMGATE_605_9_397_TICH_L39759101/AMGATE_6059_397_TICH_R14_31721314139520//)

Website QANU, onderwijsvisite Geneeskunde, 2004, geraadpleegd op: zaterdag 14 juni 2008, 12.21 uur via URL:  
[http://www.qanu.nl/comasy/uploadedfiles/GeneeskundeR\\_apportwebsite.pdf](http://www.qanu.nl/comasy/uploadedfiles/GeneeskundeR_apportwebsite.pdf)

Ramsbottom-Lucier, M.T., Gillmore, G.M., Irby, D.M., Ramsey, P.G. (1994). Evaluation of clinical teaching by general internal medicine faculty in outpatient and inpatient settings. *Academic medicine*, 69(2), 152-154.

Robson, C. (2007). *How to do a research project, a guide for undergraduate students*. Oxford: Blackwell Publishing.

Rubak, S., Mortensen, L., Ringsted, C., en Malling, B.M. (2008). A controlled study of the short- and long term effects of a Train the Trainers course. *Medical education*, 42, 693-702.

Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't: It's about integrating individual clinical expertise and the best external evidence, *BMJ*, 312(7023), 71-72.

Sanders, K., & Ards, J. (2008). Waarom trainingen (niet) effectief zijn. *Leren in organisaties*, 8 (4), 14-18.

Segers, J. (2002). *Methoden voor de maatschappijwetenschappen*. Assen: Koninklijke van Gorcum BV.

Snell, L., Tallett, S., Haist, S., Hays, R., Norcini, J., Prince, K., Rothman, A., Rowe, R. (2000). A review of the evaluation of clinical teaching: new perspectives and challenges. *Medical Education*. 34, 862-870.

Swanborn, P.G., (2002). *Basisboek sociaal onderzoek*. Amsterdam: Uitgeverij Boom.

Sutkin, G., Wagner, E., Harris, I., & Schiffer, R. (2008). What makes a good clinical teacher in medicine? A review of the literature. *Academic medicine*. 83 (5), 452-466.

Website UMCG, cursushoud voor medici, geraadpleegd op zondag 27 juli 2008, 18.06 uur via URL: <http://wencke4.housing.rug.nl/winkel/winkelframemedici.htm>

UMCG (2006). nota: Het UMCG en het medisch opleidingscontinuüm. Groningen: UMCG.

Website UMCG, Wenckebach Instituut Ontwikkelplatform, geraadpleegd op zondag 25 juli 2008, 20.06 uur via URL: <http://www.wenckebachinstituut.nl/documenten/medici/Ontwikkelplatform/Opdrachten/Bij-ennascholing.htm>

Website UMCG, geraadpleegd op dinsdag 27 mei 2008, 12.01 uur via URL: <http://www.umcg.nl/azg/nl/azg/>

Warries, E. & Pieters, J.M. (1992). *Inleiding Instructietheorie*. Amsterdam/Lisse: Swets & Zeitlinger B.V.

Website Wenckebachinstituut, geraadpleegd 11 april 2008, 21.09 uur via URL: <http://wencke4.housing.rug.nl/documenten/medici/VervolgDocentprofessionalisering.htm>

Wenckebach instituut. (2008). *Cursusboek 2008 Teach the teacher, cursus 1: didactische basisprincipes*. Groningen: Wenckebachinstituut UMCG.

Wenckebachinstituut, folder 'Het Wenckebach Ontwikkelplatform', Match tussen zorg en wetenschap en onderwijs, UMCG.

Werkgroep docentprofessionalisering Medisch onderwijs OCG-DMW en NVMO, juli 2007, *Competenties Docenten in het medisch onderwijs*, gedownload van website NVMO, Geraadpleegd op zaterdag 10 mei 2008, 16.00 uur via URL: <http://www.nvmo.nl/Werkgroepen/Docentprofessionalisering/tabid/80/Default.aspx>

Werkgroep docentprofessionalisering Medisch onderwijs OCG-DMW en NVMO, mei 2008, *Competenties van Docenten en opleiders in het medisch opleidingscontinuüm*, gedownload van website NVMO, Geraadpleegd op maandag 16 juni 2008, 20.59 uur via URL: [http://www.nvmo.nl/portals/0/werkgroepen/Professionalisering/competententiebeshrijvingen%20DEFINITIEF%20\(mei%202008\).pdf](http://www.nvmo.nl/portals/0/werkgroepen/Professionalisering/competententiebeshrijvingen%20DEFINITIEF%20(mei%202008).pdf)

Williams, B.C., Litzelman, D.K., Babbott, S.F., Lubitz, R.M., Hofer, T.P. (2002). Validation of a global measure of faculty's clinical teaching performance. *Academic medicine*, 77(2), 177-180.

Zuberi, R.W., Bordage, G., & Norman, G.R. (2007), Validation of the SETOC Instrument- Student evaluation of teaching in outpatient clinics. *Advances in health sciences education*, 12, 55-69.

Zwierstra, R.P., Venekamp, R., & Koppelman, G.H. (2006). Docentprofessionalisering: de stand van zaken anno 2006. *Tijdschrift voor medisch onderwijs*, 25(6), 279-284.

## **Bijlagen**

De bijlagen kunt u opvragen via

[ontwikkelplatform@wenckebachinstituut.nl](mailto:ontwikkelplatform@wenckebachinstituut.nl)

o.v.v. bijlagen scriptie Vanessa Werners