



# Eindrapportage Projecten programma Toetsing en Toetsgestuurd Leren

## **Toetsing in Beeld**

Eindrapportage (1 maart 2014)

UMC Utrecht

Cécile Ravesloot & Anouk van der Gijp namens het projectteam

De eindrapportage is een zelfstandig leesbaar document, met een evaluatief karakter. Dit betekent dat een lezer, die geen van de andere stukken over het Project gelezen of gezien heeft, zich een goed beeld moet kunnen vormen van het gehele project. Uit de rapportage dient eenvoudig af te leiden zijn hoe de bereikte eindresultaten zich verhouden tot het Controlling Document.

Bij elke publicatie van de (eind)resultaten van het project dient de Projectuitvoerder melding te maken van de medefinanciering van Stichting SURF. Daartoe dient hij de tekst zoals opgenomen in Bijlage 6 van de Subsidievoorwaarden in zijn publicatie op te nemen.

## Samenvatting

In het onderzoeksproject "Toetsing in Beeld" (2011-2014) is onderzocht of en op welke wijze de kwaliteit van de toetsing van radiologische beeldinterpretatievaardigheden met behulp van 3D beelden verhoogd zou kunnen worden.

De *radiologie* is het vakgebied dat zich bezig houdt met het vaststellen van ziektes (*diagnosticeren*) aan de hand van beelden van de inwendige mens, *radiologische beelden*. Dankzij digitale en technische ontwikkelingen kunnen tegenwoordig in enkele seconden honderden aaneensluitende dwarsdoorsnedenbeelden van het menselijk lichaam worden gemaakt. Dit maakt het mogelijk om deze volumebeelden in stack mode te bekijken, waarbij de arts als het ware door de patiënt heen scrollt. Ter vergelijking, tot zo'n 10 jaar geleden werden radiologische dwarsdoorsnede beelden bekeken als naast elkaar afgedrukte plaatjes van het menselijk lichaam (2D). Hoewel de kliniek sterk veranderd is, vond de toetsing van radiologische beeldinterpretatie tot op heden nog steeds plaats aan de hand van deze zogenaamde 2D beelden. Men kan zich afvragen of deze vorm van toetsing nog wel aansluit bij de klinische praktijk.

In dit project wordt de veronderstelling dat toetsing van radiologische beeldinterpretatie kwalitatief verbeterd zou kunnen worden door implementatie van 3D beelden in de toetsing onderzocht. Verschillende aspecten van 3D toetskwaliteit worden bij twee expertiseniveau's onderzocht, tweedejaars geneeskundestudenten en co-assistenten (werkpakket 1). Deze worden vergeleken met de kwaliteit van traditionele 2D toetsen. Verder wordt het construct radiologische beeldinterpretatie kwalitatief onderzocht (werkpakket 2). Met name wordt aandacht besteed aan de verschillen in benodigde vaardigheden van 2D versus 3D beeldinterpretatie.

In werkpakket 1 zijn in twee studies de kwaliteit van 2D en 3D vragen in radiologietoetsen met elkaar vergeleken. De eerste studie vond plaats tijdens de radiologische anatomietoets voor tweedejaars geneeskundestudenten en de tweede studie bij de radiologietoets voor co-assistenten radiologie. In het laatste geval betrof het vragen aan de hand van beelden over veelvoorkomende (sub)acute ziektebeelden. Beide toetsen bevatten parallelle 2D en 3D beeldvragen, welke qua niveau en inhoud aan elkaar gelijk waren. Na afloop van de toetsen is aan de studenten een enquêteformulier voorgelegd. De volgende kwaliteitsaspecten van 2D en 3D beeldvragen worden met elkaar vergeleken: *betrouwbaarheid, gepercipieerde authenticiteit, moeilijkheid en praktische bruikbaarheid*. In een aanvullende studie werd bij een subgroep van de deelnemers aan studie 1 (tweedejaars radiologietoets) ook een anatomietoets afgenomen. Deelnemers aan deze anatomietoets moesten in een menselijk lichaam op de snijzaal dezelfde structuren als uit de radiologietoets aanwijzen dan wel benoemen. Omdat er twee versies waren van de radiologietoets konden de vragen gekruist worden voorgelegd aan de deelnemers. De scores op de anatomietoets werden gecorreleerd aan de score op 2D en 3D radiologietoets, als maat voor de *externe validiteit*.

In werkpakket 2 hebben naast een literatuurstudie naar het construct radiologische beeldinterpretatie twee experimenten plaatsgevonden. Eén bij de tweedejaarsstudenten en één bij de coassistenten uit werkpakket 1. Aan de tweedejaars geneeskundestudenten werd een radiologische anatomietoets voorgelegd met 20 2D en 20 3D beeldvragen. Tijdens de toets werden hun digitale handelingen en hun oogbewegingen vastgelegd. Aan de coassistenten werden 2D en 3D beelden van veel voorkomende acute ziektebeelden voorgelegd, met de vraag of ze deze beelden wilden beoordelen terwijl ze hun gedachten hardop uitspraken. Op deze manier werd inzicht verkregen in het *construct beeldinterpretatie* en in het bijzonder de verschillen tussen *2D en 3D beeldinterpretatie*.

Alle hoofdactiviteiten, de meeste analyses en de rapportage ervan zijn afgerond. Per werkpakket zullen de belangrijkste conclusies ervan hier worden besproken.

### **Conclusies en opbrengsten werkpakket 1:**

De resultaten van de onderzoek in werkpakket 1 hebben informatie gegeven over de verschillen in diverse kwaliteitsaspecten van 2D en 3D beeldvragen op twee verschillende expertiseniveaus. Overall blijken 3D beeldvragen het op alle gemeten kwaliteitsaspecten beter te doen dan 2D vragen: hogere betrouwbaarheid, hogere gepercipieerde authenticiteit en betere correlatie met een externe validatiemaat (snijzaaltoetsen). De resultaten zijn beschreven in het handboek "digitale toetsing van radiologische beeldinterpretatie". Daarnaast is er veel ervaring opgedaan met het maken van 2D en 3D beeldvragen. Deze tips en tricks zijn gebundeld en eveneens ontsloten in het handboek. Een wetenschappelijk artikel over de resultaten is in concept klaar en zal binnenkort worden ingediend bij een medisch onderwijskundig wetenschappelijk tijdschrift. Een ander artikel over de resultaten van de pilottoetsen (welke voorafgaand aan het beschreven hoofdonderzoek hebben plaatsgevonden) is in augustus ingediend en is na revisie opnieuw ingediend bij Academic Radiology. Een derde artikel over de kwaliteitsaspecten bij 2D en 3D beeldvragen bij coassistenten is in draft versie klaar en zal nadat de analyses van de aanvullende inclusie heeft plaatsgevonden worden voltooid en ingediend bij een internationaal medisch onderwijskundig tijdschrift.

### **Conclusies en opbrengsten werkpakket 2:**

In werkpakket 2 is meer inzicht verkregen in het construct beeldinterpretatie in de radiologie. Het literatuuronderzoek, aangevuld met expertinterviews en resultaten van een hardopdenkstudie bij de coassistenten, heeft geresulteerd in een conceptueel raamwerk over de kennis en vaardigheden die belangrijk zijn voor beeldinterpretatie. Dit resultaat is gepubliceerd in een internationaal medisch onderwijskundig tijdschrift: *Advances in Health Sciences Education* (in press, online available). Uit de analyses van werkpakket 2 blijkt dat perceptie een grotere rol heeft in 3D beeldinterpretatie, terwijl synthese (het stellen van een diagnose met eventueel advies en actie) juist een groter aandeel heeft in 2D beeldinterpretatie. De eerste analyses van werkpakket 1.1 ondersteunen dit verschil in beeldinterpretatie tussen 2D en 3D beelden: studenten uit de studie hebben significant meer tijd nodig om een relevante structuur te fixeren in een 3D beeld ten opzichte van een 2D beeld. Ook besteden zij significant meer tijd en meer fixaties op de irrelevante structuren in 3D beelden ten opzichte van 2D beelden. Hieruit kan geconcludeerd worden dat het vinden van de relevante structuur (dit kan worden geschaard onder perceptie) meer tijd in beslag neemt. Er lijkt vooralsnog geen correlatie te zijn tussen het ruimtelijk inzicht van de deelnemers en beeldmanipulatievaardigheden. Hier zullen nog aanvullende analyses naar worden gedaan. De resultaten van werkpakket 2.1 en 2.2 zijn gerapporteerd in het handboek en zullen in de vorm van een wetenschappelijk artikel ingediend worden bij een internationaal medisch onderwijskundig tijdschrift.

## **Inleiding**

### ***Opzet en doel van het project***

Het interpreteren van beelden en daarmee *beeldinterpretatievaardigheid* is essentieel in menig opleiding en vakgebied. Door de digitale ontwikkeling kunnen beelden naast in 2D nu ook multi-dimensionaal worden gerepresenteerd. Traditioneel worden beeldvaardigheden in 2D getoetst. Er is weinig bekend over toetsing met multi-dimensionale beelden, maar de eerste resultaten wijzen uit dat het innovatief en authentiek is, met een positieve bijdrage aan de psychometrische kwaliteit van beeldvaardigheidstoetsing. Het doel van dit project is praktijkgerichte evidentie te zoeken voor deze veronderstelling. De experimenten zullen worden uitgevoerd binnen de radiologie, het medisch specialisme dat zich bezig houdt met de interpretatie van medische beelden, en daarmee exemplarisch is voor andere beeldgeoriënteerde vakgebieden.

Projectdoelen: (1) verhoging van de psychometrische kwaliteit en de validiteit van toetsing van beeldvaardigheden; (2) inzicht in de benodigde cognitieve processen bij het interpreteren van multi-dimensionale beelden.

### ***Belangrijkste te behalen resultaten***

Dit project beoogt inzicht te geven in hoe de mogelijkheden van digitaal toetsen kunnen ingezet worden om de kwaliteit van het toetsen van beeldinterpretatievaardigheden te verbeteren. Het onderzoek zal worden uitgevoerd op het gebied van de radiologie, waarbij specifiek wordt onderzocht wat het verschil in kwaliteit is tussen traditionele en innovatieve digitale toetsvormen. Verder wordt inzicht verkregen in de cognitieve processen die bij beeldinterpretatievaardigheden een rol spelen. De resultaten en ervaringen van dit project zullen d.m.v. presentaties op diverse platforms en d.m.v. handboeken worden aangeboden aan een breed publiek, maar zal primair zijn gericht op medisch onderwijs en aanverwante vakken.

Het toetsprogramma VQuest zal voor derden via een licentie beschikbaar komen.

### ***Participerende instellingen***

UMC Utrecht

- Prof. dr. J.P.J. van Schaik (projectleider)
- drs. C.J. Ravesloot (werkpakketleider 1, correspondentiepersoon)
- dr. ir. K.L. Vincken (projectlid)
- dr. I.C. van der Schaaf (stuurgroeplid)
- dr. R.L.A.W. Bleys (projectlid)
- drs. A. van der Gijp (projectlid)
- Prof dr. ThJ ten Cate (stuurgroeplid)

Universiteit Utrecht, faculteit sociale wetenschappen

- dr. M.F. van der Schaaf (co-projectleider en werkpakketleider 2)

### ***Projectperiode***

1 maart 2011 tot en met 1 februari 2014.

### ***Begroting en subsidie***

Totale begroting: € 166.090

Toegekende subsidie: € 120.000

## **Beschrijving per werkpakket**

### *Werkpakket 1:*

Het onderzoek naar kwaliteitsverschillen tussen 2D en 3D beeldvragen bij tweedejaarsgeneeskundestudenten is afgerond.

### **Radiologietoets bij tweedejaars geneeskundestudenten**

Op 10 en 12 april 2012 heeft de radiologietoets voor tweedejaars geneeskundestudenten met 3D beelden in VQuest plaatsgevonden. Er werden 20 vragen met 2D beelden en 20 vragen met 3D beelden over radiologische anatomie gesteld. Op 10 april 2012 werd de toets afgenomen bij 129 studenten en op 12 april bij 149 studenten. Tweehonderdvierenveertig studenten hebben het informed consent getekend. Na afloop van de toets werden digitale vragenlijsten verstuurd over o.a. praktische bruikbaarheid en authenticiteit. Deze vragenlijst is ingevuld door 138 studenten. De analyses van de toetsresultaten van de toets van 10 en 12 april zijn afgerond. De antwoord- en scoringsmodellen zijn geoptimaliseerd. En op basis hiervan zijn resultaten met betrekking tot de betrouwbaarheid en de moeilijkheid van 2D versus 3D vragen beschikbaar. De gevonden resultaten van beide toetsen (10 en 12 april) zijn vergelijkbaar en dat duidt erop dat deze reproduceerbaar zijn. De ingevulde vragenlijsten van de deelnemers van beide toetsen zijn geanalyseerd. En resultaten van over de gepercipieerde praktische bruikbaarheid, tijdsbesteding, moeilijkheid en authenticiteit van 2D en 3D vragen zijn bekend.

### **Anatomietoets bij selectie deelnemers radiologietoets**

Van de 244 tweedejaars geneeskundestudenten die aan de radiologietoets hadden meegedaan waren 33 studenten bereid om mee te doen aan het vervolgonderzoek: anatomietoets op de snijzaal. Op 12 en 13 juni 2012 kregen de studenten de toetsitems van de versie van de radiologietoets, die zij nog niet gemaakt hadden (studenten van 12 april kregen de toetsitems van 10 april en vice versa). In totaal bestond de toets uit 40 vragen. Na afloop van de toets kregen de studenten een digitale vragenlijst die door alle deelnemers werd ingevuld. De definitieve scoreberekeningen en betrouwbaarheidsanalyses zijn verricht. Ook zijn er correlatieberekeningen gedaan tussen de toetsresultaten van de deelnemers aan de snijzaaltoetsen met hun 2D en 3D beeldvraagscore op de radiologietoetsen van 10 en 12 april 2012. Dit geeft inzicht in de externe validiteit van de 3D beeldvragen. De vragenlijsten zijn geanalyseerd en dit leverde aanvullende resultaten op over de gepercipieerde praktische bruikbaarheid, tijdsbesteding, moeilijkheid en authenticiteit van 2D en 3D vragen t.o.v. de snijzaaltoetsvragen (externe validatie). De resultaten zijn samen met de resultaten van onderzoek 1.1 in een concept artikel beschreven, welke zal worden aangeboden aan een radiologisch tijdschrift, nadat het eerste artikel over de pilotresultaten opnieuw is ingediend.

### **Radiologietoets bij coassistenten**

Voor het onderzoek naar kwaliteitsverschillen tussen 2D en 3D beeldvragen bij de coassistententoetsen is een database met 60 parallelvragen gemaakt (30 unieke casusvragen, waarvan er 'n 2D en 'n 3D variant bestaat). Er zijn hiermee zes gepaarde toetsen met elk 10 casusvragen samengesteld (5 2D en 5 3D). De toetsparen zijn wat betreft de 10 casusvragen identiek, behalve dat als de ene toets van het tweetal de 2D variant van een bepaalde casusvraag bevat, dan bevat de paralleltoets de 3D variant van deze casus en vice versa. Op 1 september 2013 waren met succes 33 toetsen afgenomen bij in totaal 93 coassistenten (bijna elke 2 weken vanaf begin maart 2012) in kleine groepjes coassistenten (1 tot 7 studenten). Van de 93 coassistenten hebben er 89 toestemming gegeven voor gebruik van hun gegevens voor het onderzoek. Aanvankelijk was beoogd om 100 coassistenten te includeren, maar bij de tussentijdse analyses bleek dat voor met name de scoreberekeningen en betrouwbaarheidsanalyses een hogere inclusie toch wenselijk was. De inclusie van coassistenten is dus na 1 september doorgezet. Op dit moment (21 februari 2014) zijn er bij 123 coassistenten toetsen afgenomen, waarvan er 115 informed consent hebben getekend. Er zijn dus voorlopige conclusies verkregen over de betrouwbaarheid, moeilijkheid en de gepercipieerde authenticiteit en moeilijkheid

van de 2D en 3D beeldvragen. Daarnaast is er veel ervaring op gedaan met het maken van 2D en 3D beeldvragen.

### *Evaluatie werkpakket 1*

Alle beschreven in het Controlling Document beschreven activiteiten zijn uitgevoerd en voltooid. De resultaten van de onderzoeken zijn op diverse podia (congressen en symposia) gepresenteerd en gepubliceerd. Het schrijven en aanbieden van de artikelen aan wetenschappelijke tijdschriften zal nog doorgaan na beëindiging van het project. Ook zal op 18 maart 2014 een symposium plaatsvinden waarin samen met het project "digitaal toetsen met beelden, een extra dimensie" de resultaten aan een breed publiek zullen worden gepresenteerd. Er zal dan ook een workshop over het maken en afnemen van een digitale beeldinterpretatietoets worden gehouden. Naast de onderzoeksresultaten is er ook veel geleerd over de logistieke, praktische en technische aspecten van een beeldinterpretatietoets. Alle resultaten van het project zijn verwerkt in het handboek "Digitale toetsing van radiologische beeldinterpretatie".

### *Werkpakket 2*

In dit werkpakket werd onderzocht wat het construct beeldinterpretatie in de radiologie behelst en wat de verschillen zijn tussen 2D en 3D beeldinterpretatie. Alle onderzoeken zijn afgerond.

Ter voorbereiding op de experimenten is een literatuuronderzoek gedaan naar het construct "beeldinterpretatievaardigheden binnen de radiologie". Dit construct is gevalideerd en geoptimaliseerd in een interviewronde met experts (5 radiologen met onderwijskundige expertise, 2 onderwijskundigen met onderzoekservaring in het vakgebied radiologie en 2 psychologen met respectievelijk cognitieve en perceptuele expertise). Vervolgens is het raamwerk getoetst aan de uitingen van coassistenten tijdens het beoordelen van radiologische 2D en 3D beelden en op basis hiervan verfijnd. Het eindproduct is een raamwerk waarin kennis en vaardigheden staan beschreven die radiologen moeten hebben respectievelijk beheersen om beelden te kunnen interpreteren. De resultaten van deze studie zijn gepubliceerd in een internationaal medisch onderwijskundig tijdschrift: *Advances in Health Sciences Education*.

#### Onderzoek 2.1:

In deze studie werden tien studenten geïncludeerd. Zij kregen allen een anatomietoets met 20 2D beelden en 20 3D beelden. In beide categorieën bestond de helft uit aanwijsvragen en de helft uit benoemvragen. Tijdens het maken van de toets werden de oogbewegingen en de digitale handelingen van de studenten geregistreerd. Voorafgaand aan de toets legden de studenten twee ruimtelijk inzicht testen af. De oogbewegingen werden gekoppeld aan de digitale handelingen en aan de overlays (een 3D segmentatie van de relevante anatomische structuur). Op deze manier kon worden berekend wanneer en hoe lang de studenten hadden gekeken naar de relevante structuur (de gevraagde anatomische structuur) en wanneer en hoe lang naar de irrelevante structuur (de overige structuren in het 3D beeld). Helaas konden de gegevens van twee deelnemers niet meegenomen worden in de analyses, in één geval als gevolg van ontoereikende registratie van de digitale handelingen en in één geval als gevolg van ontoereikende registratie van de oogbewegingen. Van de overige acht deelnemers werd het kijkgedrag bij de 3D beelden vergeleken met de 2D beelden. De eerste resultaten zijn gerapporteerd in het handboek.

#### Onderzoek 2.2:

Aan deze mixed-methods studie hebben 20 coassistenten deelgenomen. De deelnemers maakten eerst twee ruimtelijk inzicht testen en vervolgens maakten ze een radiologietoets met 2D en 3D beeldvragen over acute ziektebeelden (de casus kwamen uit onderzoek 1.3, maar de deelnemers kregen altijd andere vragen dan ze al hadden gemaakt). Tijdens de toets uitten zij hun gedachten hardop en hier

werd een geluidsopname van gemaakt. Alle bewoordingen van de 20 deelnemers werden uitgetypt. Vervolgens werden de bewoordingen gecodeerd met behulp van een codeboek, dat gebaseerd was op het raamwerk uit het voorbereidende onderzoek. De bewoordingen van vijf deelnemers werden door twee onderzoekers onafhankelijk gecodeerd. De overeenstemming tussen de beoordelaars was redelijk tot goed ( $\kappa = 0.73$  tot  $0.83$ ). Vervolgens werden de coderingen van de 3D beeldvragen vergeleken met de 2D beeldvragen in een within-subject analyse, zodat in beeld werd gebracht welke kennis en vaardigheden vooral aangesproken worden bij 2D beeldinterpretatie en welke juist bij 2D beeldinterpretatie. Verder werd de correlatie tussen de score op de ruimtelijk inzicht testen en het aantal uitingen over beeldmanipulatie berekend. De resultaten zijn gerapporteerd in het handboek.

### *Evaluatie werkpakket 2*

Alle in het Controlling Document beschreven activiteiten zijn uitgevoerd en voltooid. De resultaten van het voorbereidende onderzoek is op diverse congressen gepresenteerd en zijn gepubliceerd in *Advances in Health Sciences Education*. De resultaten van de twee experimenten (onderzoek 2.1 en 2.2) zullen worden gepresenteerd op het afsluitende symposium. De resultaten van deze twee studies zullen na het project gepresenteerd worden op congressen en de wetenschappelijke artikelen zullen ingediend worden bij wetenschappelijke tijdschriften. Dit werkpakket heeft nieuwe inzichten gegeven over het construct beeldinterpretatie en de verschillen tussen 2D en 3D beeldinterpretatie. Deze inzichten zullen in onze instellingen en hopelijk ook in andere instellingen gebruikt worden voor het ontwikkelen van beeldinterpretatie-onderwijs en toetsing.

## **Overige zaken**

### ***Kennisdisseminatie***

De kennis en ervaring die is opgedaan is op de volgende manieren naar buiten gebracht:

#### *Handboek*

"Digitale toetsing van radiologische beeldinterpretatie"

#### *Wetenschappelijke artikelen:*

Gepubliceerd:

- 1) Van der Gijp et al, Radiological Image Interpretation, Advances in health sciences, 2013

Ingediend:

- 1) Raveslout et al, Improved Testing of Radiological Image Interpretation Skills using Volumetric Images, major revision;
- 2) Glenn Stuijzand et al, (submitted) Medical Students' Working Memory Load in Multiplanar Image Interpretation: Insights from Human-Computer Interactions and Eye Movements
- 3) Raveslout et al., Innovatieve toetsing in beeld", Medisch Contact.

Concept:

- 1) Raveslout et al. External validation of radiological assessment using 3D volumetric images,
- 2) Raveslout et al. Using 3D images in testing radiology clerks
- 3) Van der Gijp et al. Radiologists' cognitive processes during interpretation of multiplanar CT images versus two-dimensional CT images
- 4) Van der Gijp et al. Search strategy in multiplanar versus two-dimensional radiologic image interpretation related to performance of medical students: an eye tracking study

#### *Workshops*

Surf academy Summatieve toetsing 2011

Workshop op symposium "digitale toetsing van radiologische beeldinterpretatie"

#### *Congresbijdragen*

Radiologendagen (2012&2013), EARLI (2012), NVMO (2012), AMEE (2011), SURFonderwijsdagen (2013), MIPS (2013), AUR (2014)

#### *Website*

www.vquest.nl

#### *Symposium*

"Digitale toetsing van radiologische beeldinterpretatie" op 18 maart 2014.

Zie voor referenties van de opgeleverde deliverables (presentaties, posters of papers) in bijlage 2.

### ***Effectmeting***

De effecten zijn behaald op het onderzoeksterrein en slechts kwalitatief te meten (zie beschrijvingen hierboven).

### ***Exploitatieplan/implementatieplan***

De nieuwe digitale toetsvorm is blijvend geïmplementeerd in de Utrechtse geneeskundeopleiding voor alle radiologietoetsen. Daarnaast is in het vervolgproject "digitaal toetsen met beelden, een extra dimensie" de nieuwe manier van radiologische toetsing op landelijke schaal bij de toetsing van aios radiologie geïmplementeerd. Verder wordt in samenwerking met de UMCU Holdings gewerkt aan een



businessmodel voor VQuest, het digitale toetsprogramma waarmee de beeldinterpretatietoetsen worden afgenomen.

### **Evaluatie en lessons learned**

Overall is het project goed verlopen. In de looptijd van drie jaar zijn alle onderzoeken uitgevoerd en is er veel bereikt op het gebied van verspreiding van de resultaten en toepassing van de resultaten in de bestaande organisatie. Ook is er vanuit andere partijen een vraag naar de nieuwe toetstool ontstaan. Op dit moment worden de nodige stappen gezet om VQuest in de markt te zetten. Zowel in werkpakket 1 als in werkpakket 2 zijn een aantal activiteiten later afgerond dan aanvankelijk beoogd. Er is daarom 2 keer verlenging van het project aangevraagd en gekregen. Dit is de kwaliteit van de resultaten zeker ten goede gekomen. De problemen waar we tegenaan liepen en die tot vertraging hebben geleid zijn: 1) een update van het programma VQuest om aan de nieuwe dataload vraag te kunnen voldoen 2) de omvangrijkheid van de data die gegenereerd zijn met de hardopdenkonderzoeken. 3) de technische uitdaging van het koppelen van de eyetracking data aan de geregistreerde handelingen.

Aanvankelijk zijn we met een zeer kleine projectgroep van 5 personen gestart en lopende het project zijn daar vier personen gedurende kortere of langere tijd aan toegevoegd. Dit maakte dat de lijntjes onderling erg kort waren. Er kon vlug worden geschakeld en beslissingen worden genomen. Bovendien waren de projectleden allemaal erg gemotiveerd om tot een goed resultaat te komen. Soms werden er dan ook vele extra uren gewerkt om een deadline te halen.

## **Conclusies**

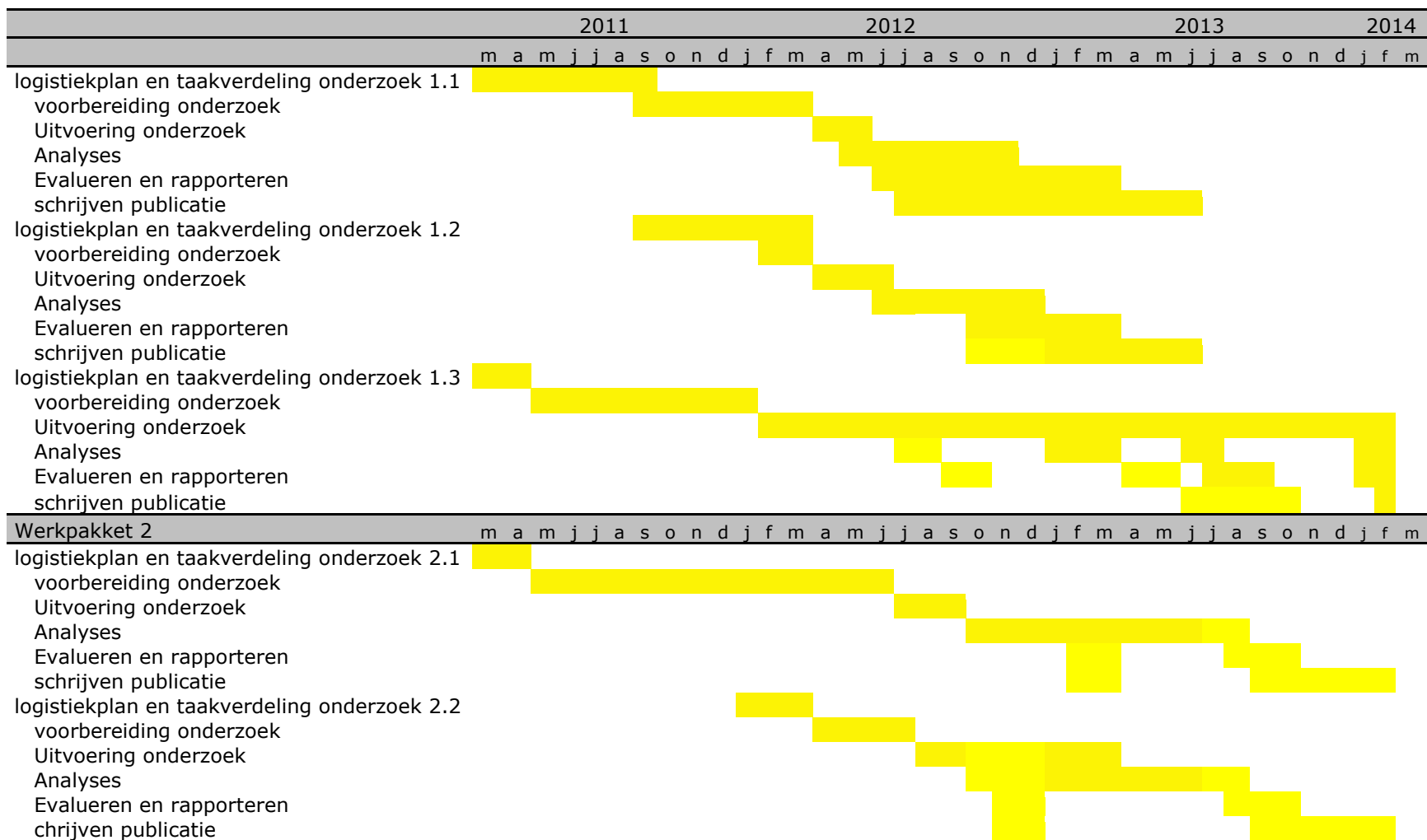
De verrichtte onderzoeken wijzen uit dat 3D beeldvragen op diverse aspecten van kwaliteit beter scoren. Zo is de betrouwbaarheid van deze vragen herhaaldelijk hoger dan voor 2D beeldvragen. Daarnaast correleerden ze beter met een externe validatiemaat. Daarnaast blijkt de praktische bruikbaarheid van 3D beeldtoetsing goed te zijn en zijn studenten ook tevreden over de nieuwe manier van toetsing. Verder is er inzicht verkregen in het construct radiologische beeldinterpretatie, waarbij de diverse kennis en vaardigheden zijn geïdentificeerd die een rol spelen. Het verschil tussen 2D en 3D beeldinterpretatie is dat bij 3D beelden meer denkprocessen worden gebruikt en de nadruk meer op perceptie ligt, terwijl de nadruk bij 2D beelden ligt op het stellen van de diagnose. Over het algemeen zijn studenten langer bezig met het maken van een 3D beeldvraag dan met een 2D beeldvraag (ongeveer twee keer zo lang bij tweedejaars geneeskundestudenten). Het is dus belangrijk om steeds af te wegen of het toetsdoel deze tijdsinvestering verantwoordt.

### **Afsluitende financiële verantwoording**

Geef een toelichting op en analyse van het financiële verloop van het project. Gebruik voor het financiële overzicht het aangeleverde SURFformat. Mag ook als losse Excel bijlage aangeleverd worden. Vergeet niet zaken als rekeningnummer, kenmerk waaronder SURF die geld moet overmaken etc.

## Bijlage 1 Standlijnenoverzicht

De geelgearceerde activiteiten zijn in de afgelopen periode verricht.



## **Bijlage 2 Opgeleverde deliverables**

### 2014

Raveslout, C., Van der Gijp, A., Van der Schaaf, M., Bleys, R.L.A.W., Mol, C.P., Huige, J.C.B.M., Ten Cate, O., Van Schaik, J.P.J., & Vincken, K.L. Handboek "Digitale toetsing van radiologische beeldinterpretatie" (maart 2014)

Symposium "Digitale Toetsing van Radiologische Beeldinterpretatie" georganiseerd door SURF-projecten "Toetsing in Beeld" en "Digitaal toetsen met beelden, een extra dimensie" 18 maart 2014, UMC Utrecht

Van der Gijp, A, Van der Schaaf, MF., Van der Schaaf, I.C., Raveslout, C.R., Huige, J.C.B.M, Van Schaik, J.P.J. & Ten Cate, Th. J. (2014). *Image interpretation in radiology: a conceptual framework of knowledge and skills. AUR Baltimore USA*

### 2013

Raveslout, C., Van der Gijp, A., Van der Schaaf, M., Ten Cate, O., Van Schaik, J.P.J., Mol, C., Tipker, C., Maas, M., & Vincken, K.L. (2013). Quality improvement of radiological image interpretation skills assessment through digital MPR images in Medical Education. Radiologendagen 19&20 sept 2013

Van der Gijp, A, Van der Schaaf, MF., Van der Schaaf, I.C., Raveslout, C.R., Huige, J.C.B.M, Van Schaik, J.P.J. & Ten Cate, Th. J. (2013). *Image interpretation in radiology: a conceptual framework of knowledge and skills. MIPS 2013, August 14-16, Washington United States.*

Van der Gijp, A. & Raveslout, C.J., (SURFonderwijsdagen 2013) Pucha Kucha, Toetsing in Beeld & Toetsen met beelden, een extra dimensie.

### 2012

Raveslout, C.J., Van der Schaaf, M.F., Van der Gijp, A., Mol, C.P., Ten Cate, Th.J, Van Schaik, J.P.J., & Vincken, K.L. (2012). *Authenticiteit en betrouwbaarheid van toetsing van radiologische beeldinterpretatievaardigheden met multi-slice beelden.* Poster presentation. Radiologendagen Sept 26 and 27, Den Bosch, The Netherlands.

Raveslout C.J., Van der Schaaf, M.F., Haaring, C., Kruitwagen, C.L.J.J., Beek, F.J.A., Ten Cate Th.J, Van Schaik, J.P.J. (poster). Development of radiological expertise measured with progress testing in radiology. Symposium "teacher expertise", October 11, 2012, Utrecht University.

Raveslout, C., Van der Schaaf, M.F., Vincken, K. ,& Van Schaik, J. (2012). Quality improvement of assessing radiological skills with digital multi-slice images in medical education. Earli sig 1 meeting, Brussels. First author: C. Raveslout, other co-authors: K. Vincken and J. van Schaik. Brussels on 28-31st August, 2012.

Van der Gijp, A., Van der Schaaf, M.F., Van der Schaaf, I.C., Raveslout C.R., Van Schaik, J.P.J., & Ten Cate, Th.J., (wetenschappelijk paper). NVMO proceedings, 15 en 16 november 2012, Maastricht.

Raveslout, C.J., Van der Schaaf, M.F., Van der Gijp, A., Huige, J., Ten Cate, Th.J., Van Schaik, J.P.J., Vincken, K.L., Ervaringen toetsing radiologische beeldinterpretatie met digitale beeldvragen. NVMO, 15 en 16 november 2012 Maastricht; bijdrage aan: Kok E et al. Symposium Visueel diagnostische vaardigheden in de medische praktijk: van onderzoek naar onderwijs.

## 2011

Raveslout C.J., Van der Schaaf M.F., Vincken K.L., Ten Cate Th.J., Van Schaik J.P.J. (2011) Digital assessment of radiological skills with digital multi-dimensional images in medical education. AMEE sept 2011 Wenen

Van der Schaaf, M.F., Raveslout, C.J., & Vincken, K.L. (2011). Toetsing in Beeld. Posterprestatie . Abstractboek Onderwijs Research Dagen, Maastricht 8-10 juni 2011, p. 244-245.

Raveslout, C.J., Van der Schaaf, M.F., & Vincken, K.L. (2011). Toetsing in Beeld. Workshop "Good practices van de Universiteit Utrecht" Onderwijsparade 10 maart 2011 "Toetsen en leren, de kracht van beoordelen"

Van der Schaaf, M.F., Raveslout, C.J., & Vincken, K.L. (2011). Toetsing in Beeld. Conferentie: Onderwijscentraal, Utrecht

Vincken, K.L., Raveslout, C.J., Van der Schaaf, M.F., workshop "toetsen met VQuest" tijdens SURF academy Workshop "Aan de slag met summatieve toetsen." Penners, M. feb 2011

## 2010

Raveslout, C., Schaaf, M.F. van der, Haaring, C., Beek, F.J.A., Cate, Th.J. ten & Schaik, J.P.J. van (2010). De Voortgangstoets Radiologie (VGT), een betrouwbare en constructvalide kennistoets? In Proceedings 15e Radiologendagen (pp. 13). Veldhoven.

Raveslout, C., Schaaf, M.F. van der, Cate, Th.J. ten, Haaring, C., Beek, F.J.A. & Schaik, J.P.J. van (2010). De Voortgangstoets Radiologie (VGTR), een betrouwbare en constructvalide kennistoets? In Proceedings NVMO Congres (pp. 146-147). Egmond aan Zee.

Raveslout, C., Schaaf, M.F. van der, Haaring, C., Beek, F.J.A., Cate, Th.J. ten & Schaik, J.P.J. van (2010). Dutch radiology progress test: knowledge learning curve in radiology residents. In Assessment of Competence in Medicine and the Healthcare Professions (pp. 42). Miami, Florida.