



Sjabloon Eindrapportage Projecten programma Toetsing en Toetsgestuurd Leren Versie: Feb 2014

Nieuwe score regel voor digitale toetsen

Eindrapportage (4^e kwartaal 2013 + jan-feb 2014)

Gegevens penvoerende instelling
Drs. S. Klinkenberg
Vakgroep Psychologische Methodenleer, afdeling Psychologie,
Universiteit van Amsterdam
Weesperplein 4
1018 XA Amsterdam
S.Klinkenberg@uva.nl
020 525 6904
06 24 14 39 33

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

Het doel van dit onderzoeksproject was om de bruikbaarheid van de high speed high stakes score regel te onderzoeken. Deze regel voegt response tijd (RT) toe aan het bepalen van de behaalde score op een item door de het antwoord, de accuratesse (acc) op het item, gescoord als 0 of 1, te vermenigvuldigen met $(acc^2-1)*(d-RT)$. Hierbij is d de tijdlimiet voor het item. Met betrekking op de werking van de score regel hebben we ons de volgende vragen gesteld.

1. Leidt het gebruik van de HSHS scoreregel tot een verbetering van de betrouwbaarheid en validiteit van toetsen?
2. Zijn respondenten in staat om de optimale balans tussen accuratesse en snelheid te kiezen?
3. Zijn hierin belangrijke individuele verschillen te vinden?
4. Hangen die verschillen samen met achtergrondvariabelen als ervaring, sekse en vaardigheid?
5. Moeten respondenten leren om te gaan met deze scoreregel of kan deze zonder al te veel instructie direct worden toegepast?
6. Leidt het gebruik van de scoreregel tot een accuratere vaardigheidsschattingen in adaptieve testen met makkelijke items?

Aan de hand van een aantal onderzoeken is hier getracht antwoord op te geven.

Deze onderzoeken hebben geleid tot twee publicaties in internationale peer reviewed tijdschriften, een nog niet gepubliceerd artikel en twee onderzoeksrapporten.

- Klinkenberg, S., Straatemeier, M., Van der Maas, H. L. J. (2011). Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Comput. Educ.*, 57(2), 1813–1824.
- Jansen, B.R.J., Louwerse, J., Straatemeier, M., Van der Ven, S.H.G., Klinkenberg, S., Van der Maas, H.L.J. (2013). The influence of practicing maths with a computer-adaptive pro-gram on math anxiety, perceived math competence, and math performance. *Learning and Individual Differences*, 24, 190–197.
- Klinkenberg, S., Van der Maas, H.L.J. (2013). A dynamic paired comparison based computer adaptive testing method. Unpublished manuscript.
- Barkhof, J., Bekker, T., Bersma, M., Groenendijk, E., Maza, S. (2013). Oefening Baart Kunst (unpublished research report). University of Amsterdam, Netherlands.
- Özen, S., Pronk, A., Sanchez Maceiras, S., Stel, N., Van Wersch, T. (2012). De Invloed van de HSHS scoreregel op het Meten van Werkelijke Vaardigheid (Unpublished research report). University of Amsterdam, Netherlands.

Het gebruik van de HSHS scoreregel in zowel adaptieve als niet adaptieve omgevingen levert zowel valide als betrouwbare schattingen van vaardigheid op. Hoewel de betrouwbaarheid niet beter blijkt te zijn dan de som correct score regels is de validiteit wel degelijk hoger, met name bij adaptieve afname. Specifiek bij de afname van makkelijke items met een hoge kans op correct werkt de HSHS score regel beter. De dispositie op "risk taking" bleek, met betrekking tot het schatten van vaardigheid, de werking van de HSHS scoreregel niet te beïnvloeden. Respondenten beoordeelden de regel echter wel zeer gevarieerd, van zeer positief tot zeer negatief. Hoewel de frequentie verdeling van correcte en incorrecte antwoorden op responstijden aanwijzing geven dat respondenten antwoorden conform de scoreregel, geeft een aanzienlijk deel te kennen de regel niet goed te begrijpen. Of het gedrag komt niet overeen met de ervaring van de scoreregel of de Statistiekfabriek implementatie verschilt fundamenteel van de rekentuin implementatie. Dit laatste klopt deels aangezien in het Statistiekfabriek onderzoek er sprake was van een implementatie als oefententamen. Het wekenlang gebruik maken van de regel in de rekentuin kan mogelijk resulteren in een nauwgezetter begrip van de regel dan wanneer het slechts eenmalig in een oefententamen wordt toegepast. Dit zou er voor pleiten om studenten zich vertrouwd te laten raken met de regel voor hem toe te passen in summatieve toetsing.

Inleiding

Opzet en doel van het project

De doelstelling van dit project is om een theoretische oplossing, speciaal geschikt voor digitale toetsen, voor een tweetal klassieke toets problemen in de praktijk te onderzoeken. Daarmee geven we digitaal toetsen een kwalitatieve meerwaarde boven op de klassieke sequentiële papieren toetsen. Als de gekozen oplossing ook in de praktijk werkt zal dat leiden tot efficiëntere en eerlijkere digitale toetsen, met een grote meerwaarde voor computer adaptieve toetsen.

Belangrijkste te behalen resultaten

In het algemeen beogen we een praktische bevestiging te vinden voor de theoretische voordelen van het gebruik van scoreregels in toetsen en voor de High Speed High Stakes (HSHS) scoreregel in het bijzonder. Indien succesvol kan dit leiden tot een belangrijke algemeen toepasbare structurele verbetering van digitale toetsen. Met de oplossing van het snelheid-accuratesse probleem en het probleem van gokken worden toetsen kwalitatief beter en eerlijker. De winst van deze verbeteringen ten opzichte van andere maatregelen is naar onze verwachting relatief hoog.

Om dit te bereiken moeten we onze vragen eerst positief beantwoorden. De subprojecten moeten slagen. Per project is het beoogd resultaat als volgt.

- **Subproject 1** moet laten zien dat de HSHS regel leidt tot hogere validiteit en betrouwbaarheid. Als output moet Subproject 1 twee rapportages opleveren, bij positieve uitkomsten ook in de vorm van publicaties in internationale peer-reviewed tijdschriften. Het schrijven van deze publicaties valt echter buiten deze projectaanvraag. Voor beide implementaties A en B is al veel voorwerk gedaan.
- **Subproject 2** moet laten zien dat respondenten alleen voordelen en geen nadelen van de scoreregel ondervinden. We verwachten praktische richtlijnen voor het gebruik van de HSHS regel te kunnen opstellen. We vereisen tenminste 1 schriftelijk rapport, eventueel in de vorm van een artikel.
- **Subproject 3** moet bijdragen aan computer adaptieve testen met relatief makkelijke items. Ook dit moet leiden tot 1 rapportage. Een pilot simulatie studie is al met succes uitgevoerd.

Participerende instellingen

University of Amsterdam
Department of Psychology

Project leider
Drs. Sharon Klinkenberg
Weesperplein 4
1018 XA Amsterdam
Room DB 1.03
T +31 20 525 6904
E S.Klinkenberg@uva.nl

Stuurgroepleden
Prof. dr. Van der Maas
Weesperplein 4
1018 XA Amsterdam
Room DB 2.13
T +31 20 525 6678
E H.L.J.vanderMaas@uva.nl

De heer Klaas Visser
Weesperplein 4
1018 XA Amsterdam
Room DB 1.09
T +31 20 525 6772
E K.H.Visser@uva.nl

Projectperiode

1 maart 2011 - februari 2014
Kwartaal: okt – dec 2013, jan – feb 2014

Begroting en subsidie

Totale projectkosten	€ 77,766
Subsidie	€ 55,019

Beschrijving per werkpakket

Per werkpakket: een inhoudelijk evaluatief verslag van de activiteiten en resultaten/deliverables die zijn uitgevoerd respectievelijk bereikt, afgezet tegen de afspraken die zijn gemaakt in het Controlling Document.

Werkpakket 1: Management

In de oorspronkelijke aanvraag hebben we de omvang van het projectmanagement overschat. Gezien dit een onderzoeksproject is hebben we er voor gekozen om de kosten hiervoor terug te brengen naar 2,5% van de totale projectbegroting in plaats van de daarvoor geadviseerde 5%. De tussentijdse rapportages en een aanpassing van een fout in de begroting in het begin van het project hebben hierbij veel werk gekost en zijn ondergebracht bij overhead.

Werkpakket 2: Projectondersteuning

Voor de ondersteuning van het projectmanagement zullen we gebruik maken van het projectbureau van de UvA. Het projectbureau van de UvA zal de financiële administratie op zich nemen. Tot de taken behoren het indienen van declaraties bij Surf en de matching van 30% verwerken. Voor dit werkpakket worden geen kosten in rekening gebracht en dit zal daarom ook niet in de begroting worden opgenomen.

Dit onderdeel is in goede orde verlopen. Hoewel we bij aanvang van het project voor ogen hadden dat het projectbureau zou bijdragen aan het tot stand komen van de financiële kwartaalrapportages. Dit bleek niet het geval.

Werkpakket 3: Subproject 1

In subproject 1 komt de eerste vraagstelling aan de orde. Dit werkpakket bestaat uit het analyseren en rapporteren van de al verzamelde schaakdata en rekentuindata en het samenstellen van de rapporten. Het gaat hier om het uitvoeren van betrouwbaarheidsanalyses en convergente validiteitstoetsing.

	Uitgevoerd	Beraamd
Analyseren schaakdata	44	12 uur
Rapportage resultaten	44	12 uur
Analyseren Rekentuindata	0	12 uur
Rapportage resultaten	0	12 uur

Deliverables: Rapporten m.b.t. de betrouwbaarheid en validiteit analyses.
Kwaliteitscriteria: Afdeling Psychologische Methodenleer van de Universiteit van Amsterdam.

De publicatie van de rekentuindata bleek al voor de start van het project in 2011 gerealiseerd te zijn. In overleg met de programmamanager is besloten om de beschikbare uren over te hevelen naar het schaakonderzoek. De belangrijkste reden hiervan was dat we meer tijd hebben moeten besteden aan het opschonen en uitzoeken van de dataset. In kwartaal rapportage okt – nov 2012 is nogmaals een uitbreiding aangevraagd waar een akkoord op is gegeven door de programmamanager. In totaal ging het om een verhoging van 40 uur. Reden voor de uitbreiding was om een poging te doen om tot een publicabel artikel te komen. Het schaakartikel is tot op heden nog niet gepubliceerd.

De twee opgeleverde documenten zijn:

- Klinkenberg, S., Straatemeier, M., Van der Maas, H. L. J. (2011). Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Comput. Educ.*, 57(2), 1813–1824.
- Klinkenberg, S., Van der Maas, H.L.J. (2013). A dynamic paired comparison based computer adaptive testing method. Unpublished manuscript.

Werkpakket 4: Subproject 2

Dit subproject bestrijkt vraagstelling 2 t/m 4 waarin in eerste instantie gebruik gemaakt zal worden van de rekentuindata. Tevens zullen de gegevens uit de implementatie in het statistiekonderwijs gebruikt worden voor analyses. Het werkpakket bestaat uit het omzetten van een papieren versie naar een digitale statistiekttoets, analyseren van de data en het schrijven van het rapport.

	Uitgevoerd	Beraamd
Digitaliseren statistiekttoets	38	38 uur
Organiseren onderwijs	38	38 uur
Ondersteuning programmeurs	38	38 uur
Analyseren statistiekdata	39	38 uur
Rapportage statistiekresultaten	36	38 uur
Analyseren rekentuindata	36	38 uur
Rapportage rekentuindata	38	38 uur

Deliverables: Rapport over individuele verschillen in SAT strategie

Kwaliteitscriteria: De controle hiervan wordt gedaan door Prof. Dr. Han Van der Maas

De resultaten van dit werkpakket zijn allen succesvol opgeleverd. Het digitaliseren van de statistiekttoets en het implementeren in het onderwijs heeft geresulteerd in de Statistiekfabriek, opgeleverd in SURF TTL project "[Computer Adaptieve Monitoring in het statistiekonderwijs](#)". Daarnaast is het vraagstuk van individuele verschillen opgenomen in een publicatie over resultaten uit de Rekentuin. Daarnaast zijn er nog een paar aanvullende analyses gedaan die zijn opgenomen in de samenvatting van de onderzoeksresultaten.

- Jansen, B.R.J., Louwerse, J., Straatemeier, M., Van der Ven, S.H.G., Klinkenberg, S., Van der Maas, H.L.J. (2013). The influence of practicing maths with a computer-adaptive program on math anxiety, perceived math competence, and math performance. *Learning and Individual Differences*, 24, 190–197.
- Barkhof, J., Bekker, T., Bersma, M., Groenendijk, E., Maza, S. (2013). Oefening Baart Kunst (unpublished research report). University of Amsterdam, Netherlands.
- Özen, S., Pronk, A., Sanchez Maceiras, S., Stel, N., Van Wersch, T. (2012). De Invloed van de HSHS scoreregule op het Meten van Werkelijke Vaardigheid (Unpublished research report). University of Amsterdam, Netherlands.
- Klinkenberg, S. (2014). Samenvatting resultaten TTL tender: Nieuwe scoreregule digitaal toetsen. SURF, Netherlands.

Werkpakket 5: Subproject 3

Werkpakket 5 bestaat uit het uitvoeren van simulatie onderzoek, het toetsen van de effecten van verschillende scoreregels binnen de Rekenruimte en het schrijven van een rapport.

	Uitgevoerd	Beraamd
Programmeren simulaties	38	38 uur
Uitvoeren analyses	38	38 uur
Rapportage simulaties	38	38 uur
Opzetten experimentele condities in Rekenruimte	38	38 uur
Ondersteuning programmeurs	15	38 uur
Analyseren resultaten	15	38 uur
Rapportage experimentele manipulaties	8	38 uur

Deliverables: Rapport over effecten van scoreregels op de standard error of measurement.
Kwaliteitscriteria: Prof. Dr. Han Van der Maas.

De resultaten van dit werkpakket zijn opgenomen in de volgende documenten:

- Klinkenberg, S., Straatemeier, M., Van der Maas, H. L. J. (2011). Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Comput. Educ.*, 57(2), 1813–1824.
- Jansen, B.R.J., Louwerse, J., Straatemeier, M., Van der Ven, S.H.G., Klinkenberg, S., Van der Maas, H.L.J. (2013). The influence of practicing maths with a computer-adaptive program on math anxiety, perceived math competence, and math performance. *Learning and Individual Differences*, 24, 190–197.
- Klinkenberg, S., Van der Maas, H.L.J. (2013). A dynamic paired comparison based computer adaptive testing method. Unpublished manuscript.

Werkpakket 6: Evaluatie en afronding

De evaluatie en afronding zal uitgevoerd worden door de penvoerder en bestaat uit een bundeling van de deelrapporten en het samenstellen van een eindrapport voor SURF.

	Uitgevoerd	Beraamd
Schrijven inleiding en proces	38	38 uur
Samenvoegen resultaten	16	16 uur
Schrijven eindconclusie	16	16 uur

Deliverables: Eindrapport.

Kwaliteitscriteria: De eindrapportage zal ter controle worden voorgelegd aan SURF.

Voor de eindrapportage is een samenvattend document opgesteld waarin de belangrijkste resultaten van de verschillende onderzoeken in zijn samengevat.

- Klinkenberg, S. (2014). Samenvatting resultaten TTL tender: Nieuwe scoreregels digitaal toetsen. SURF, Netherlands.

Werkpakket 7: Disseminatie

Voor de disseminatie en kennisoverdracht is het voornamelijk de bedoeling de opgeleverde onderzoeken te publiceren. Verder zal de projectleider actief deelnemen aan congressen op het gebied van digitaal toetsen, computer adaptief toetsen en ICT en onderwijs om de opgedane kennis te presenteren.

De volgende activiteiten zullen ondernomen worden:

- Publicatie van onderzoeksgegevens in wetenschappelijke tijdschriften
 - Zie eerder genoemde publicaties
- Publicatie in EXAMENS (Tijdschrift voor de toetspraktijk)
 - [EXAMENS 2013-02](#)
- Presentaties op relevante wetenschappelijke en ICT congressen
 - Presentatie op Kennisnet Conferentie '[Maar wat werkt nu echt? Praktijkverhalen over ICT in het onderwijs](#)' door Prof. Han van der Maas.
 - TTL presentatie [Statistiekfabriek](#) door Caspar Groeneveld
 - Presentatie [EARLI SIG 2012](#)
- Presentaties georganiseerd door SURF
 - OWD 2012/2013
- Presentaties op (zelf-)uitnodiging bij instellingen
 - [ONBETWIST Conferentie: trends in toetsgestuurd leren](#)
 - Presentatie [PML](#) 2012
 - Presentatie SBO congres 2011
- Publicatie in SURF magazine
 - Niet uitgevoerd
- Publicatie op SURF space
 - <https://www.surfspace.nl/artikel/1226-computer-adaptive-practice-of-maths-ability/>
 - <https://www.surfspace.nl/artikel/635-workflow-bij-samenstellen-van-digitale-toetsen/>
- Deelname SURF programma-activiteiten
 - OWD, SIG DT dagen.

Aanvullend

- Niet wetenschappelijke publicaties
 - Publicatie in [COS](#)
 - Publicatie in [PAROOL](#)
 - Publicatie Lucas Raggars op [kennisnet](#)
 - Publicatie in [VIVES](#) 2012
- Voorzitter SIG Digitaal toetsen sinds 2012
- [SURF Masterclass digitaal toetsen](#) 2011 en 2012

Tevens zal er disseminatie plaatsvinden door de samenwerking met Prof. Dr. Maris van CITO en Prof. Dr. Glas van de universiteit Twente. Beiden zijn experts op het gebied van de psychometrie.

Prof. Dr. Maris is inmiddels met een team van AIO's, werkzaam bij de afdeling psychologische methodenleer aan het project: "Optimal learning benefit from educative games".

"The use of educational games during and outside school time is increasing. This research assumes that good educational games ought to optimise the learning benefit per unit time. But how do we do this and can that be measured? In this project modified versions of the popular TuxMath computer game will be used to do experimental research into this. Consortium partners: Cito, Oefenweb, University of Amsterdam."

Dit project is tot stand gekomen door het verwerven van een [Creative Industry grant](#). Er wordt binnen dit project voortgewerkt aan dit onderwerp en hoe educatieve games optimaal in het onderwijs zijn in te zetten.

In 2014 zal een bijdrage worden geleverd aan de Computer Assisted Assessment (CAA) conferentie te Zeist. Hiervoor is in het laatste kwartaal van dit project een samenvattend artikel geschreven. Dit is een iets beknoptere Engelse versie van de samenvatting voor het TTL project. Tevens zal er nog op deze conferentie gepresenteerd worden.

- Klinkenberg, S. (2014). High Speed High Stakes Scoring Rule - Assessing the performance of a new scoring rule for digital assessment. Presented at CAA2014: Research into E-Assessment, Zeist. Netherlands.

Overige zaken

Voor zover de volgende onderdelen niet in een apart werkpakket/fase aan de orde komen, dienen ook aan bod te komen:

Kennisdisseminatie

Besproken in WP 7

Effectmeting

Niet van toepassing

Exploitatieplan/implementatieplan

Het toepassen van de scoreregels in andere situaties vereist een computer afname waarin de responstijd kan worden geregistreerd en waarbij per item een tijdlimiet kan worden ingesteld. Het heeft daarbij de voorkeur om de afloop van tijd ook visueel weer te geven binnen het item gezien de feedbackloop. Bij een niet adaptieve toets kan men de som van HSHS scores $((acc*2-1) * (d-RT))$ per item hanteren. Hierbij moet de tijdlimiet d van tevoren worden ingeschat. Een langere tijdlimiet zou daarbij een item relatief eenvoudiger moeten maken. Voor een adaptieve afname zullen of de item moeilijkheden van tevoren bepaald moeten worden of zal een Elo update procedure zoals beschreven in Klinkenberg (2011) geïmplementeerd moeten worden. Beide oplossingen vereisen een aanzienlijke investering en beschikbaarheid van psychometrische kennis. Om hier een praktische oplossing voor te bieden wordt er binnen Oefenweb nagedacht over het beschikbaar maken van een [application programming interface](#) (API) waardoor een externe aanbieder gebruik zou kunnen maken van het algoritme. De ideeën hieromtrent zullen nog uitgewerkt moeten worden maar er kan hierbij gedacht worden aan een licentie model voor externe ontwikkelaars.

Evaluatie en lessons learned

Terugkijkend op het projectverloop is er een aantal punten te benoemen. Ten eerste kan ik stellen dat een onderzoeksproject zich niet zo goed laat vangen in het SURF project planningsformat. Vooraf valt slecht te voorzien wanneer je waar tijd aan kan besteden in een periode van drie jaar. Gaande weg besef je je vervolgens dat het geheel ook niet volgens het controlling document zal gaan verlopen terwijl het wel duidelijk is dat de uit te voeren onderdelen goed te doen zijn binnen het beschikbare budget en beschikbare tijd. Dit heeft tot gevolg gehad dat er veel tijd is gaan zitten in het aanpassen van de projectplanning en de daarbij behorende verantwoordingen. Ik ben ook nauw betrokken geweest bij het TTL project "[Computer Adaptieve Monitoring in het statistiekonderwijs](#)" en heb gezien hoe het projectmatige karakter daar goed werkt. Onderzoek in combinatie met onderwijs bleek slecht voorspelbaar. De voornaamste problemen binnen dit project waren dan ook het vrij maken van tijd op de momenten dat het volgens de planning stond aangegeven. Het TTL programma management van SURF is hier overigens altijd erg goed mee omgegaan. Het was goed mogelijk om beargumenteerd de standlijnen aan te passen zolang maar duidelijk gemaakt kon worden dat de project doelstellingen niet in gevaar zouden komen. Het verloop van het project zou je dan ook kunnen vergelijken met een cross loop. Op sommige momenten van het parcours ging het zeer traag terwijl op andere momenten het weer heel hard kon gaan. Al met al kunnen we stellen dat het project goed is afgerond. Alle doelstellingen zijn gehaald met een onder uitputting van iets meer dan €10.000,- en binnen de gestelde termijn. Naar mijn idee zijn er geen risico's geweest voor de afronding van dit project. Hoewel er veel deadlines zijn verschoven of verplaatst is het eind doel altijd helder in zicht geweest.

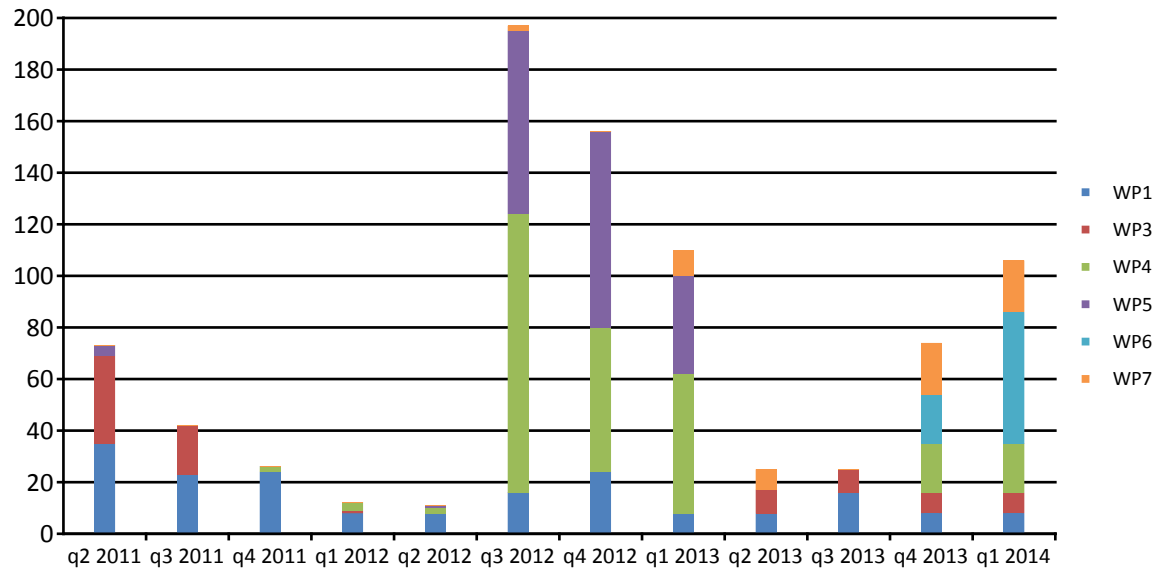
Conclusies

Het gebruik van de HSHS scoreregels in zowel adaptieve als niet adaptieve omgevingen levert zowel valide als betrouwbare schattingen van vaardigheid op. Hoewel de betrouwbaarheid niet beter blijkt te zijn dan de som correct score regels is de validiteit wel degelijk hoger, met name bij adaptieve afname. Specifiek bij de afname van makkelijke items met een hoge kans op correct werkt de HSHS score regel beter. De dispositie op "risk taking" bleek, met betrekking tot het schatten van vaardigheid, de werking van de HSHS scoreregels niet te beïnvloeden. Respondenten beoordeelden de regel echter wel zeer gevarieerd, van zeer positief tot zeer negatief. Hoewel de frequentie verdeling van correcte en incorrecte antwoorden op responstijden aanwijzing geven dat respondenten antwoorden conform de scoreregels, geeft een aanzienlijk deel te kennen de regel niet goed te begrijpen. Of het gedrag komt niet overeen met de ervaring van de scoreregels of de Statistiekfabriek implementatie verschilt fundamenteel van de Rekentuin implementatie. Dit laatste klopt deels aangezien in het Statistiekfabriek onderzoek er sprake was van een implementatie als oefentamen. Het wekenlang gebruik maken van de regel in de Rekentuin kan mogelijk resulteren in een nauwgezet begrip van de regel dan wanneer het slechts eenmalig in een oefentamen wordt toegepast. Dit zou er voor pleiten om studenten zich vertrouwd te laten raken met de regel voor hem toe te passen in summatieve toetsing.

We voelen ons gesterkt in het vertrouwen dat de HSHS scoreregels betrouwbare en valide resultaten oplevert. We blijven nog wel voorzichtig als het gaat om individuele verschillen en de attitude ten aanzien van de regel als het gaat om summatieve toetsing. Met betrekking tot formatieve toetsing denken we dat deze attitudes minder een rol spelen. De HSHS scoreregels belooft de brug te slaan tussen snelheid en accuratesse waardoor we denken op het goede spoor te zitten.

Afsluitende financiële verantwoording

Onderstaand is een globale analyse van de besteding in uren per werkpakket per kwartaal opgenomen. Opvallend hierbij is dat de minst productieve periodes het eerste en tweede kwartaal van 2012 zijn en het tweede en derde kwartaal van 2013, de implementatie van de onderwijsvernieuwingen voor onze afdeling. De totale eindrapportages van q4 2013 en q1 2014 zijn opgenomen in de bijlage.



Bijlage 2 Opgeleverde deliverables

Als gedurende de rapportageperiode een uitgebreid rapport is opgeleverd dient het volledige rapport in de bijlagen opgenomen te worden. In de hoofdtekst geef je een **korte** samenvatting van de belangrijkste conclusies.

De twee samenvattende artikelen zijn:

- Klinkenberg, S. (2014). [Samenvatting resultaten TTL tender: Nieuwe scoreregels digitaal toetsen](#). SURF, Netherlands.
- Klinkenberg, S. (2014). [High Speed High Stakes Scoring Rule - Assessing the performance of a new scoring rule for digital assessment](#). Presented at CAA2014: Research into E-Assessment, Zeist. Netherlands.

De project resultaten zijn:

- Klinkenberg, S., Straatemeier, M., Van der Maas, H. L. J. (2011). [Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation](#). Comput. Educ., 57(2), 1813–1824.
- Jansen, B.R.J., Louwerse, J., Straatemeier, M., Van der Ven, S.H.G., Klinkenberg, S., Van der Maas, H.L.J. (2013). [The influence of practicing maths with a computer-adaptive program on math anxiety, perceived math competence, and math performance](#). Learning and Individual Differences, 24, 190–197.
- Klinkenberg, S., Van der Maas, H.L.J. (2013). [A dynamic paired comparison based computer adaptive testing method](#). Unpublished manuscript.
- Barkhof, J., Bekker, T., Bersma, M., Groenendijk, E., Maza, S. (2013). [Oefening Baart Kunst](#) (unpublished research report). University of Amsterdam, Netherlands.
- Özen, S., Pronk, A., Sanchez Maceiras, S., Stel, N., Van Wersch, T. (2012). [De Invloed van de HSHS scoreregels op het Meten van Werkelijke Vaardigheid](#) (Unpublished research report). University of Amsterdam, Netherlands.