



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 619349



www.project-watchme.eu



@Project_WatchMe

The logo for the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), featuring the letters "NWO" in a stylized font with a red swoosh above the "O".

NWO

Nederlandse Organisatie voor Wetenschappelijk Onderzoek



Universiteit Utrecht

Gevalideerd beoordelen van een gepersonaliseerd portfolio

Marieke van der Schaaf


Department of Education, Utrecht University

+ Electronisch portfolio


<http://www.leraar24.nl/video/1367> (and /1368)

MATEUM


MATHEMATISCHE TECHNOLOGIE




DASHBOARD




INSERT



PROGRESS



PROFILE



MY EPASS

EPASS

Home / Dashboard

Personal data

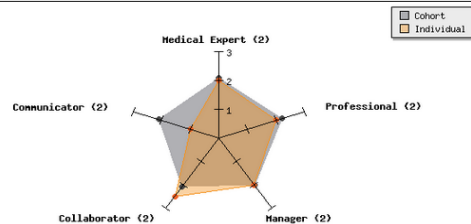


Name: Suzanne Schut demo student
 Department: Diakonessenhuis Undergraduate
 Medical Education Utrecht
 E-mail address: s.schut@maastrichtuniversity.nl
 Phone number: -
 Total attachments: 0MB / 50MB

[Edit information](#)

EPASS

My competencies



JIT Feedback

EPA 1: History taking, general impressions and general examination

There is room for improvement on this EPA.

[Click to see Improvement Feedback](#)

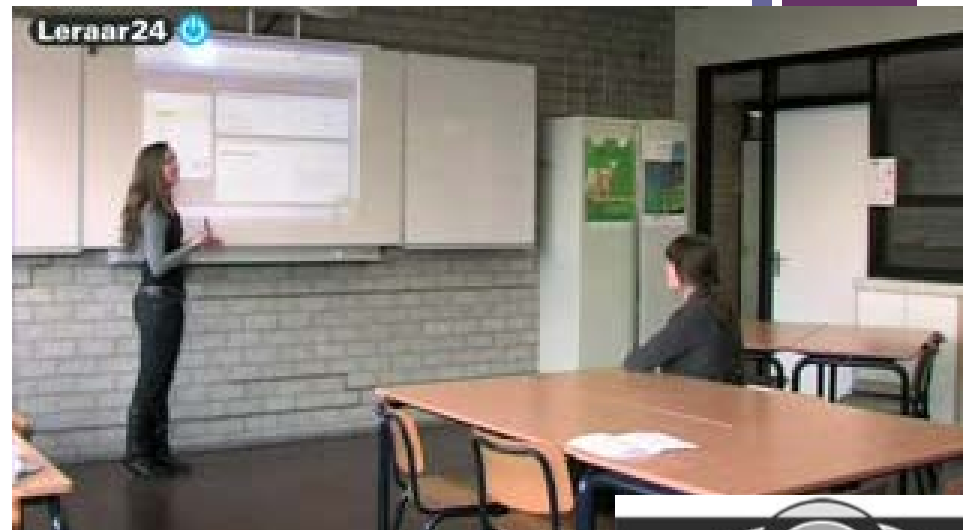
Your Supervisor added a few comments on this EPA.

[Click to see Supervisor Feedback](#)

You currently have a trend of decreasing scores for this EPA

[Click to see Trend Feedback](#)

EPA 2: Respiration problem



Universiteit Utrecht



Gepersonaliseerde assessment en feedback geeft (Sadler, 1989; 2010):

- Inzicht in performance
- Monitoring process
- Suggesties voor verbetering

Feeds into learners' major feedback questions (Hattie & Timperley, 2007):

- Where am I going? (goals, feedup)
- How am I going? (feedback)
- Where to next? (feedforward)



Electronic portfolios: mixed success

- Onderbenutting van potentiële data
- Authentiek, persoonlijk uniek karakter, bemoeilijkt scoring
- Standardisatie is onmogelijk (no golden standard or external criterion).
- **Negatieve impact op betrouwbaarheid en validiteit van de beoordeling**





KWALITEITSEISEN PERFORMANCE ASSESSMENT

(Stokking, Van der Schaaf, Erkens & Jaspers, 2004)



Validiteit	<p>Meten wat men beoogt te meten</p> <ul style="list-style-type: none">•Ontwikkeling zichtbaar maken•Onderscheiden tussen meer en minder bekwamen•Duiden waar lerenden toe in staat zijn <p>Voorkom construct underrepresentation en construct irrelevant variance</p>
Betrouwbaarheid	Herhaalbaar, reproduceerbaar
Praktische bruikbaarheid	Uitvoerbaar Gewenste functies vervullen
Aanvaardbaarheid	Beoordelingsprocedure en criteria helder Procedure en criteria objectief





HOE KWALITEIT VERBETEREN

Twee studies

- 1) Inzicht krijgen in beoordelingsprocessen
- 2) Hoe kunnen we de kwaliteit van beoordelingsprocessen verbeteren

- 3) Van der Schaaf, M.F., Stokking, K.M., & Verloop, N. (2005). Cognitive representations in raters' assessment of teacher portfolios. *Studies in Educational Evaluation*, 31, 27-55.

Van der Schaaf, M.F., Stokking, K.M., & Verloop, N. (2005). *De invloed van cognitieve representaties van beoordelaars op hun beoordeling van docentportfolio's*. *Pedagogische Studiën*, 82 (1), 3-26.

Van der Schaaf, M.F., Baartman, L.K.J., & Prins, F.J. (2011). Exploring the role of assessment criteria during teachers' collaborative judgment processes of students' portfolios. *Assessment and Evaluation in Higher Education*.





Hoe denken beoordelaars?



STUDIE 1

6 experts

18 docent portfolios





Resultaten

- Beoordelaars bouwen een holistisch consistent beeld op
- Het beeld is vaak beïnvloed door eigen ervaringen van de beoordelaars
- Beoordelaars zochten naar concreet bewijs (bv. Videomateriaal) om hun opgebouwde beeld te bevestigen
- Hoe meer beoordelaarsparen overeenstemden in uitspraken over de inhoud van de beoordeling en de mate van negativiteit/positiviteit van die uitspraken, hoe hoger hun interrater betrouwbaarheid ($r = .44; .53$).





Wat denken beoordelaars?

Portfoliopresentaties van 28 Vwo-leerlingen

6 Docenten die in moderation beoordeelden

Procedure

- ZONDER BEOORDELINGSFORMULIER (n=10).
- MET BEOORDELINGSFORMULIER (n=18).



+ Beoordelingsformulier



1. Geeft een beeld van sterke en minder sterke kanten.
2. Geeft aan waardoor sommige zaken meer/minder goed gaan.
3. Dit zijn mijns inziens reële sterke en minder sterke argumenten gezien de vaardigheden van de leerling.
4. Onderbouwt sterke kanten.
5. Dit zijn mijns inziens geschikte bewijzen
6. Geeft aan hoe er gewerkt is aan mindere punten.
7. Geeft aan wat dit heeft opgeleverd en waardoor het dat heeft opgeleverd.
8. Geeft aan hoe hij/zij vaardigheden en/of attitude wil verbeteren (doelen)





Kwaliteit van argumentatie in beoordelings processen

(Clark & Sampson, 2005; based on Toulmin, 1958)



Taped conversations and judgment forms, analysed on Content and quality:

- **Geen of irrelevante evidentie**
- **Persoonlijke ervaringen als bewijs**
- **Een bewijs**
- **Afweging gebaseerd op meerdere bewijzen**
- **Check op het beoordelingsproces**



Inhoud van processen



Beoordelingsformulier	Zonder	Met
Bewijs in het portfolio en de presentatie ¹	40, 4%	51, 7%
Competenties en ontwikkeling van studenten	24, 3%	19, 7%
Over student in het algemeen ²	17, 7%	6, 6%
Irrelevant / niet over student	16, 5%	17, 2%

1) $U = 53$, $z = -2.25$, $p < .05$, $r = -.41$ [medium effect]

2) $U = 56$, $z = -2.11$, $p < .05$, $r = -.39$ [medium effect]



Kwaliteit van argumentatie in beoordelingsprocessen

Beoordelingsformulier	Zonder	Met
Geen beoordeling	43,9 %	50,5%
Geen / irrelevante evidentie	12,8 %	8,2%
Persoonlijke ervaring¹	15,7 %	3,9%
Een bewijs	22,0 %	24,0%
Meerdere bewijzen	0,3%	0,7%
Check beoordelingproces²	5,2 %	12,3%

1) $U = 24$, $z = -3.58$, $p < .00$, $r = -.66$ [large effect]

2) $U = 52$, $z = -2.26$, $p < .05$, $r = .42$ [medium effect]





Discussie

- Beoordelingsprocessen van beoordelaars bevatten bedreigingen voor de validiteit van de beoordeling
- Construct-irrelevante variantie en construct onderrepresentatie vaak gebaseerd op geen of irrelevant bewijs en persoonlijke ervaringen van de beoordelaars
- Het gebruik van een beoordelingsformulier (gestructureerde benadering) komt tegemoet aan deze problematiek.
- RUBRICS! – hanteer ze holistisch





Hartelijk dank voor uw
aandacht!

Contact

m.f.vanderschaaf@uu.nl

Marieke van der Schaaf
Department of education, Utrecht University



Universiteit Utrecht