

Een jaar na het winnen van de Prijs voor Examens

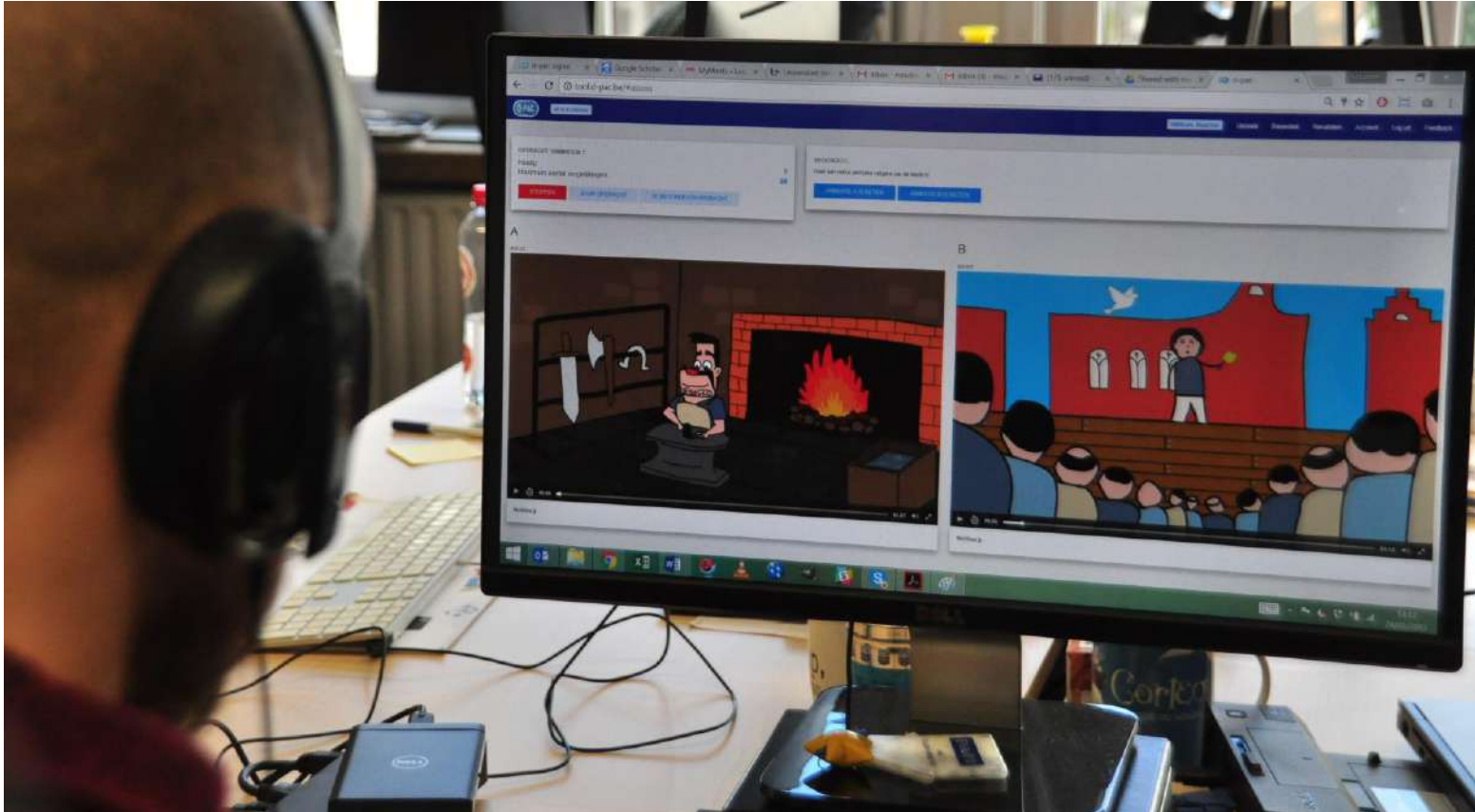
Marije Lesterhuis & Roos Van Gasse

NVE congres 2019



D-PAC

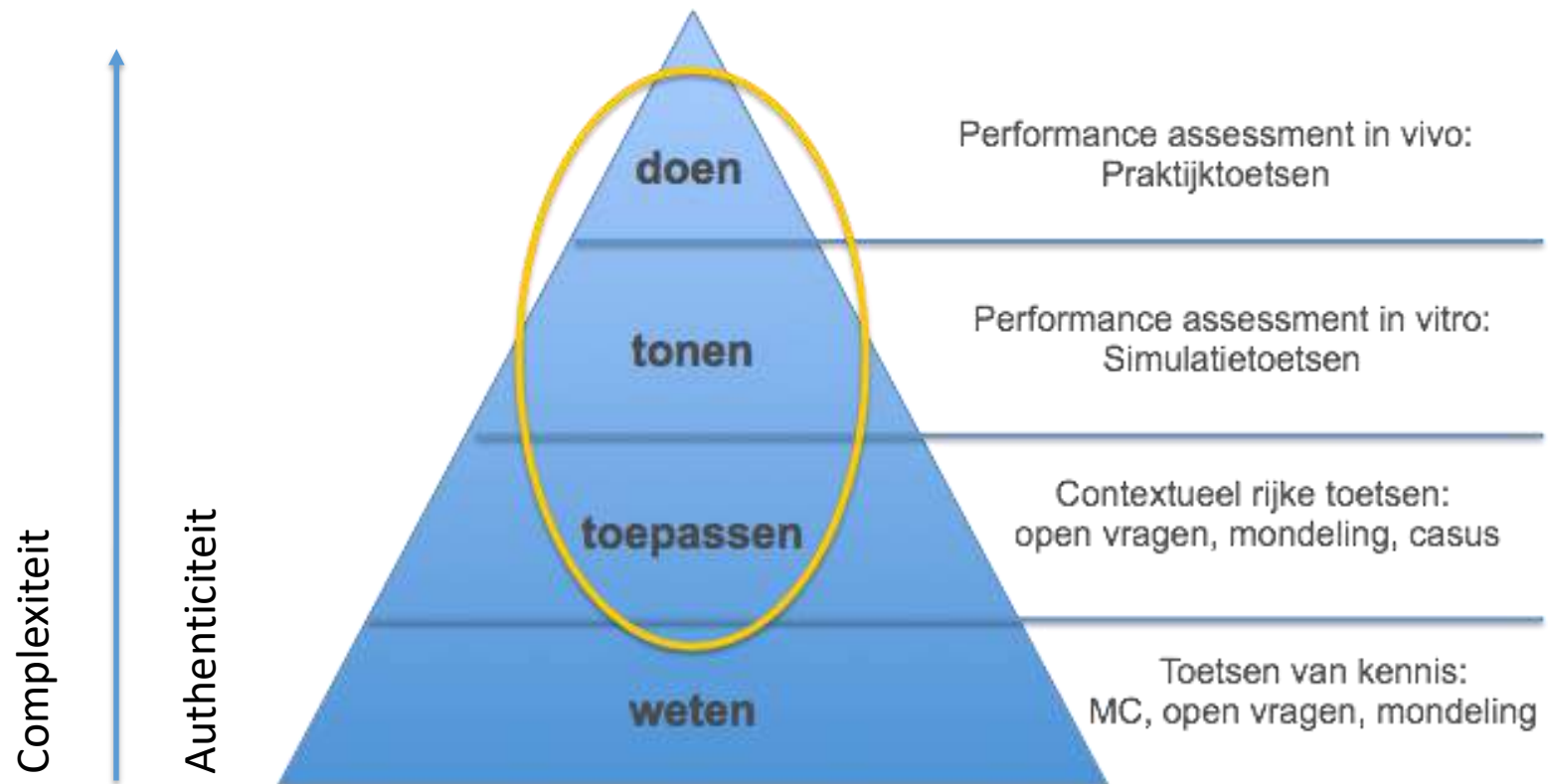
Digital Platform for the Assessment of Competences



NVE congres 2019

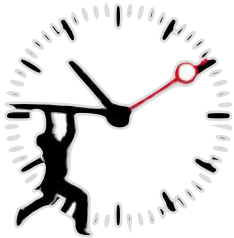


Aanleiding



Aanleiding

Problematiek competenties beoordelen



Beoordelen van
competenties kost
(te) veel tijd



Beoordelaars
meten niet wat ze
willen meten



Beoordelaars
komen niet
overeen

Aanleiding

Wat maakt het beoordelen zo lastig

- op welke aspecten let je wel en welke niet?
- welk oordeel geef je, wat is de range?
- is deze beter dan de vorige?
- geef ik nog wel dezelfde scores als in het begin?
- geven mijn collega's dezelfde oordelen?
- ...

Aanleiding

Hoe donker is dit vlak op een schaal van 1 tot 10?



Aanleiding

Ondersteuning aan beoordelaars

Rubric voor het schrijven van een paper
Geesteswetenschappen, Bachelor niveau 2

Criteria niveau 2	Onvoldoende (score 35)	Matig (score 50)	Voldoende (score 65)	Goed (score 80)	Excellent, er valt niets te verbeteren (score 100)
Formuleren van een probleem-, vraagstelling binnen de aangereikte context Weging 10%	De probleem-, vraagstelling ontbreekt, is niet helder, of niet afgebakend, of niet haalbaar, of staat niet in relatie tot de aangereikte context.		De probleem-, de vraagstelling is concreet en staat in relatie tot de aangereikte context, maar kan beter afgebakend en/of scherper geformuleerd worden.	De probleem-, de vraagstelling is concreet, goed afgebakend, scherp geformuleerd en staat in relatie tot de aangereikte context.	
Inbedden van de vraag in de aangereikte literatuur Weging 15%	Er is geen relevante selectie gemaakt uit de aangereikte literatuur met het oog op de probleem-/vraagstelling. Er wordt nauwelijks of geen relatie gelegd tussen de vraag en de literatuur.		Uit de aangereikte literatuur is een relevante selectie gemaakt met het oog op de probleem-/vraagstelling. Er wordt een relatie gelegd tussen de vraag en de literatuur, maar deze kan explicieter worden gemaakt en/of beter worden verantwoord.	Uit de aangereikte literatuur is een relevante en weloverwogen selectie gemaakt met het oog op de probleem-/vraagstelling. De vraag is overtuigend en nauwkeurig ingebed in de literatuur.	
Toelichten van de aangereikte methode Weging 10%	De methode is niet, onvoldoende of onjuist toegelicht.		De methode is expliciet, bondig en correct toegelicht.	De methode is expliciet, bondig, overtuigend en nauwkeurig toegelicht.	
Presentatie en begrip van de gebruikte bronnen Weging 15%	De beschrijving van de gebruikte bronnen is onnauwkeurig, onvolledig en/of inhoudelijk onjuist. Uit de beschrijving blijkt dat de student geen of weinig begrip heeft van de gebruikte bronnen.		De beschrijving van de bronnen is over het algemeen nauwkeurig en inhoudelijk juist. Uit de beschrijving blijkt dat de student in grote lijnen begrip heeft van de gebruikte bronnen.	De beschrijving van de bronnen is nauwkeurig en inhoudelijk juist. Uit de beschrijving blijkt dat de student een gedetailleerd begrip heeft van de gebruikte bronnen.	
Analyseren van de bronnen op basis van de aangereikte methode Weging 10%	De aangereikte methode wordt niet of onjuist toegepast. De analyse getuigt van onnauwkeurige of onjuiste observaties en er is geen aanzet tot interpretatie.		De aangereikte methode wordt correct toegepast. De analyse getuigt van adequate observaties en er is een aanzet tot interpretatie.	De aangereikte methode wordt overtuigend toegepast. De analyse getuigt van nauwkeurige en genuanceerde observaties. De interpretatie is overtuigend.	
Structureren en argumenteren op inhoudelijk niveau Weging 10%	De tekst is geen consistent verhaal dat toewerkt naar beantwoording van de hoofdvraag. Een logische opbouw van denkstappen in afgebakende paragrafen en alinea's ontbreekt. Uitspraken worden niet of nauwelijks ondersteund met argumenten.		De tekst is een voornamelijk consistent verhaal dat toewerkt naar beantwoording van de hoofdvraag. Er is een logische opbouw van denkstappen in afgebakende paragrafen en alinea's. Uitspraken worden over het algemeen ondersteund met adequate argumenten.	De tekst is een consistent verhaal dat op systematische en overtuigende wijze toewerkt naar beantwoording van de hoofdvraag. Er is een logische opbouw van denkstappen in afgebakende paragrafen en alinea's. Uitspraken worden ondersteund met adequate argumenten.	
Conclusie: Beantwoorden van de hoofdvraag en evalueren van de methoden	De hoofdvraag wordt niet of onvoldoende beantwoord.		De hoofdvraag wordt beantwoord maar de uitwerking van het antwoord blijft oppervlakkig.	De hoofdvraag wordt zorgvuldig beantwoord en de werkwijze en resultaten worden geëvalueerd.	

Aanleiding

Ondersteuning beoordelaars?

Rubric voor het schrijven van een paper
Geesteswetenschappen, Bachelor niveau 2

Criteria niveau 2	Onvoldoende (score 35)	Matig (score 50)	Voldoende (score 65)	Goed (score 80)	Excellent, er valt niets te verbeteren (score 100)
Formuleren van een probleem-, vraagstelling binnen de aangereikte context Weging 10%	De probleem-, vraagstelling ontbreekt, is niet helder, of niet afgebakend, of niet haalbaar, of staat niet in relatie tot de aangereikte context		De probleem-, de vraagstelling is concreet en staat in relatie tot de aangereikte context, maar kan beter afgebakend en/of scherper geformuleerd worden.	De probleem-, de vraagstelling is concreet, goed afgebakend, scherp geformuleerd en staat in relatie tot de aangereikte context.	
Inbedden van de vraag in de aangereikte literatuur Weging 15%	Er is geen relevante selectie gemaakt uit de aangereikte literatuur met het oog op de probleem-/vraagstelling. Er wordt nauwelijks of geen relatie gelegd tussen de vraag en de literatuur.		Uit de aangereikte literatuur is een relevante selectie gemaakt met het oog op de probleem-/vraagstelling. Er wordt een relatie gelegd tussen de vraag en de literatuur, maar deze kan explicieter worden gemaakt en/of beter worden verantwoord.	Uit de aangereikte literatuur is een relevante en weloverwogen selectie gemaakt met het oog op de probleem-/vraagstelling. De vraag is overtuigend en nauwkeurig ingebed in de literatuur.	
Toelichten van de aangereikte methode Weging 10%	De methode is niet, onvoldoende of onjuist toegelicht.		De methode is expliciet, bondig en correct toegelicht.	De methode is expliciet, bondig, overtuigend en nauwkeurig toegelicht.	
Presentatie en begrip van de	De beschrijving van de gebruikte		De beschrijving van de bronnen is	De beschrijving van de bronnen is	

Aanleiding

Problemen met rubrics en criterialijsten

Lastig om te construeren

- Definiëring van het construct
- Specifieke versus globale criteria

Geen garantie voor betrouwbare en valide oordelen

- Nog steeds ruimte voor interpretatie van criteria
- > lage interbeoordelaarsbetrouwbaarheid
- Verschil in strengheid/mildheid van de beoordelaar
 - Halo-effecten (vooral bij expertbeoordelaars)
 - Volgorde-effecten

Aanleiding

Comparatief beoordelen

Thurstone (1927): Law of Comparative Judgement

“Men is beter en betrouwbaarder in het vergelijken van twee objecten dan in het toekennen van absolute scores aan een enkel object”

Aanleiding

Comparatief beoordelen

Welk van de twee vlakken is het donkerst?



Aanleiding

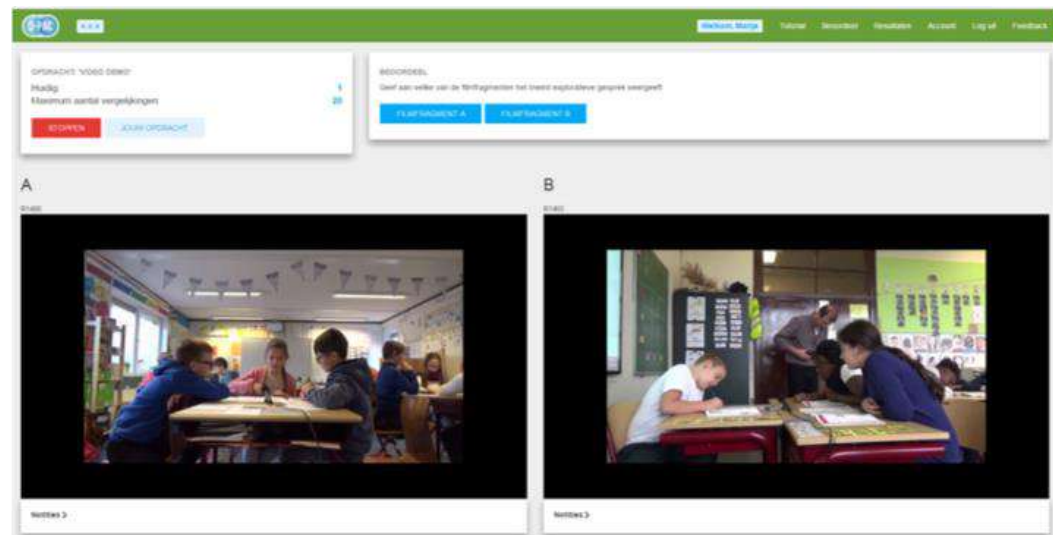
Comparatief beoordelen



Aanleiding

Comparatief beoordelen

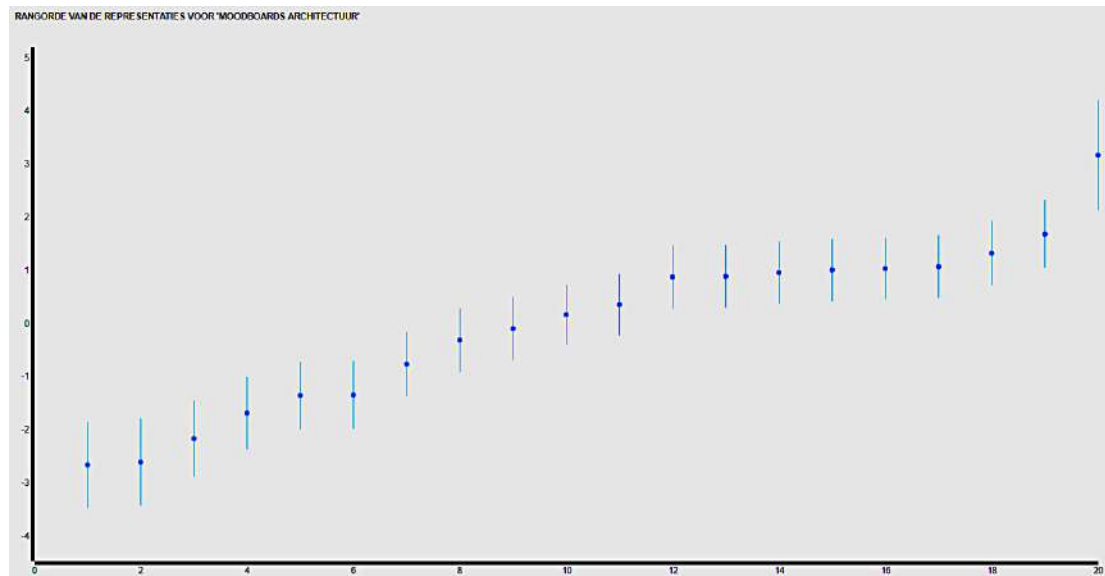
- Elk product komt meerdere keren terug in een random samengesteld paar
- Meerdere beoordelaars vellen holistisch oordeel (evt. ook met feedback)



Aanleiding

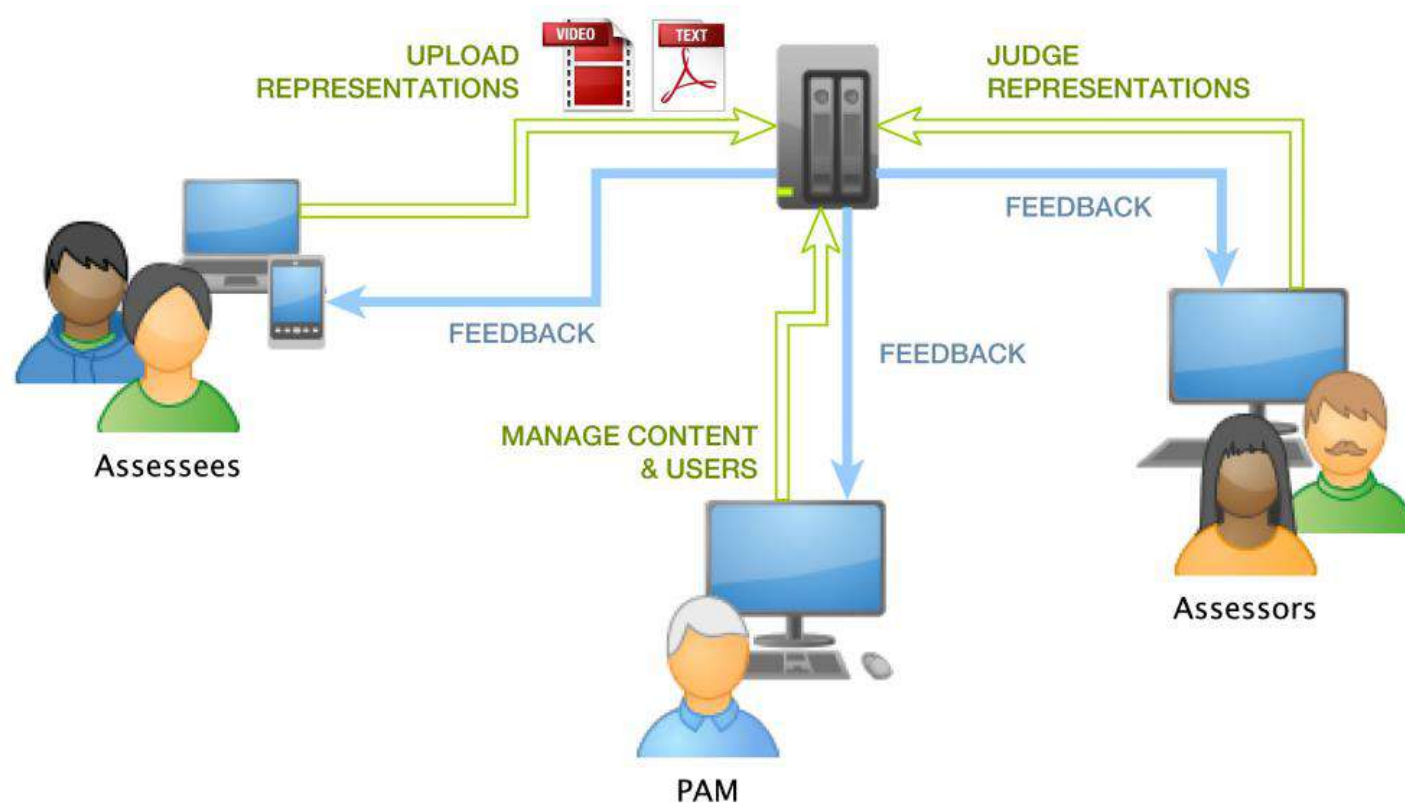
Comparatief beoordelen

- Rangorde van representaties (scores in logits)
- Beoordelaars die afwijken van het model (misfits)
- Representaties die afwijken van het model (misfits)



D-PAC

De tool



D-PAC

De tool

- Surf naar <http://sandbox.d-pac.be>
- Ga naar 'beoordeel'

Vier jaar onderzoek

- wie gebruikt het en waarvoor?
- betrouwbaarheid & efficiëntie
- validiteit
- peer assessment
- voor de punten

Wie gebruikt het?



Waarvoor wordt het gebruikt?

Basisonderwijs

- beschrijvende teksten
- gespreksvaardigheden (video)
- inclusieve werkvormen (video)

Voortgezet onderwijs:

- argumentatieve teksten
- creatieve teksten
- formele brieven
- synthese teksten
- probleemoplossend vermogen wiskunde
- schetsen voor beeldende vaardigheden

Hoger onderwijs

- wetenschappelijk rapporteren
- essays
- onderzoeksverslagen
- labverslagen
- behandelingsplannen
- zelfreflectie-opdrachten
- portfolio
- interactieve installaties (live)
- statistische output (afbeelding)
- personal branding (afbeelding)
- moodboards (afbeelding)
- ER diagrammen (afbeelding)
- interviewvaardigheden (video)
- Presentatievaardigheden (video)
- animatiefilmpjes (video)
- Stages (video)

Overig

- projectvoorstellen
- te investeren technologie
- te ontwerpen verkeerslicht
- Werkbelasting inschatting
- prioritering features
- statements t.a.v. school
- metaforen

Betrouwbaarheid en efficiëntie

Betrouwbaarheid van minimaal .70:

- gemiddeld 12 vergelijkingen per representatie
minimaal 9, maximaal 20

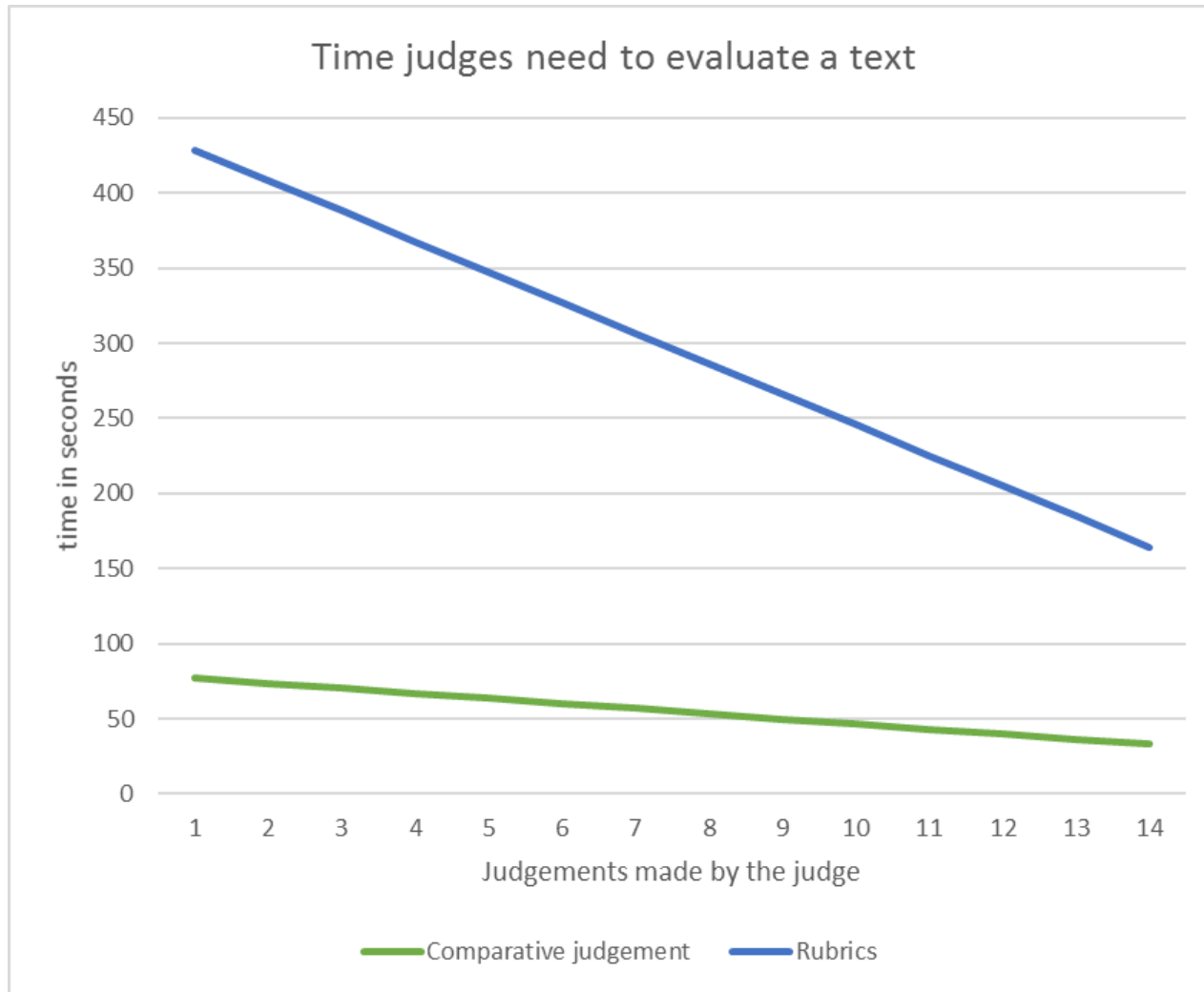
Betrouwbaarheid van minimaal .80:

- gemiddeld 17 vergelijkingen per representatie
minimaal 13, maximaal 25

Voorbeeld met 20 studenten:

*$20 * 12 \text{ vergelijkingen} / 2 = 120 \text{ vergelijkingen}$*

Betrouwbaarheid en efficiëntie

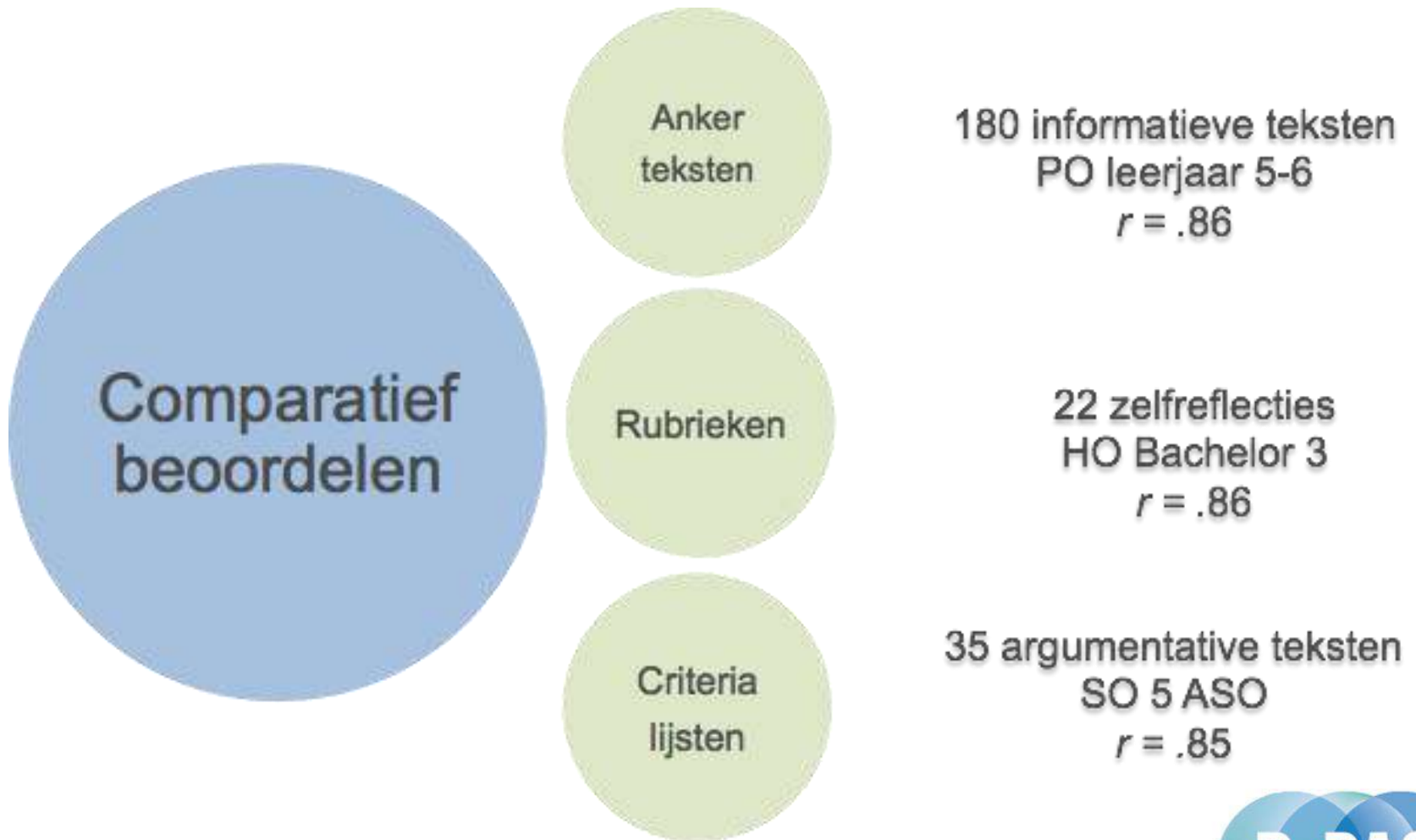


NVE congres 2019



Validiteit

Beoordelen we hetzelfde als met andere methoden?

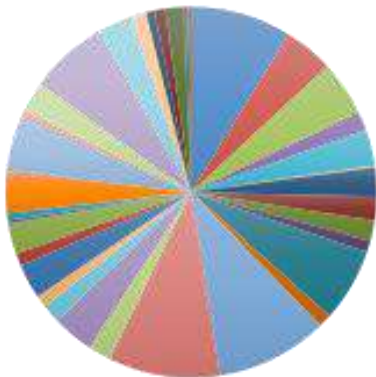


Validiteit

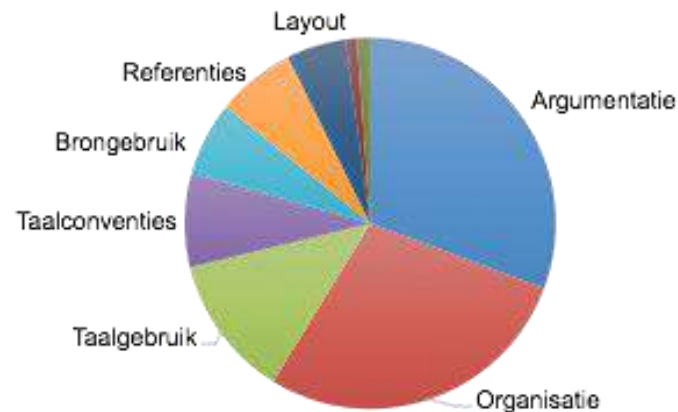
Welke aspecten beïnvloeden de comparatieve beoordelingen?

Argumentatief schrijven in het vo

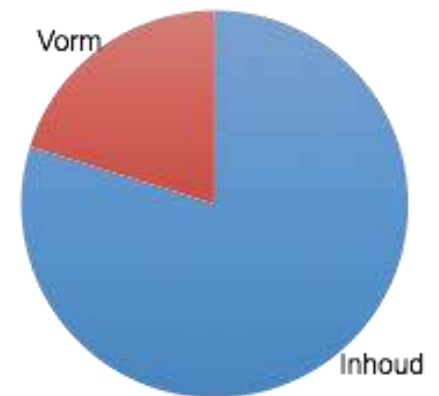
Genoemde aspecten



Argumentatief schrijven

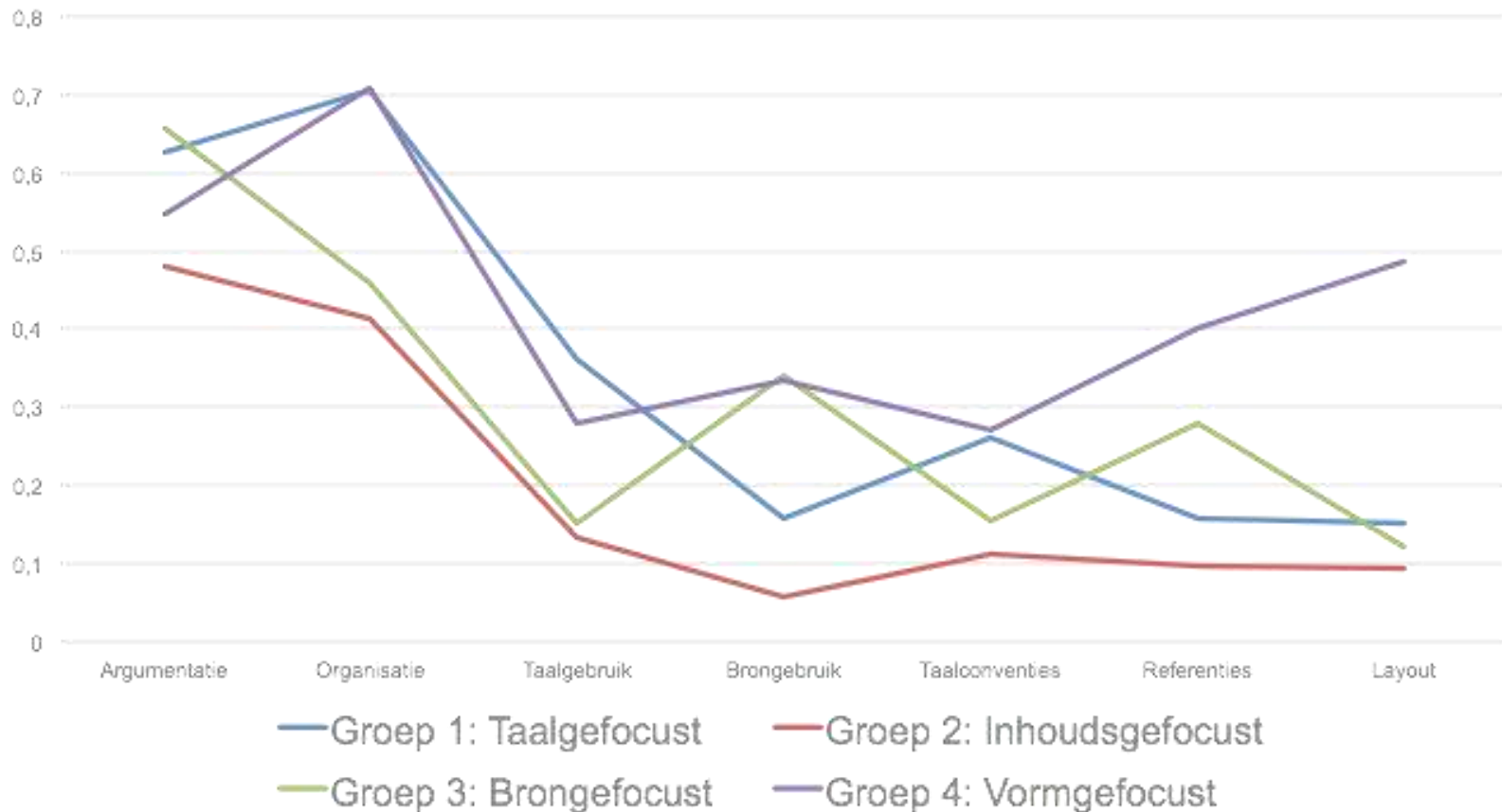


Inhoud versus vorm



Validiteit

De meerwaarde van meerdere beoordelaars



Peer assessment

Voordelen

Eenvoudig studenten beoordelen even betrouwbaar als docenten (SSR beide groepen = .77; correlatie = .65)

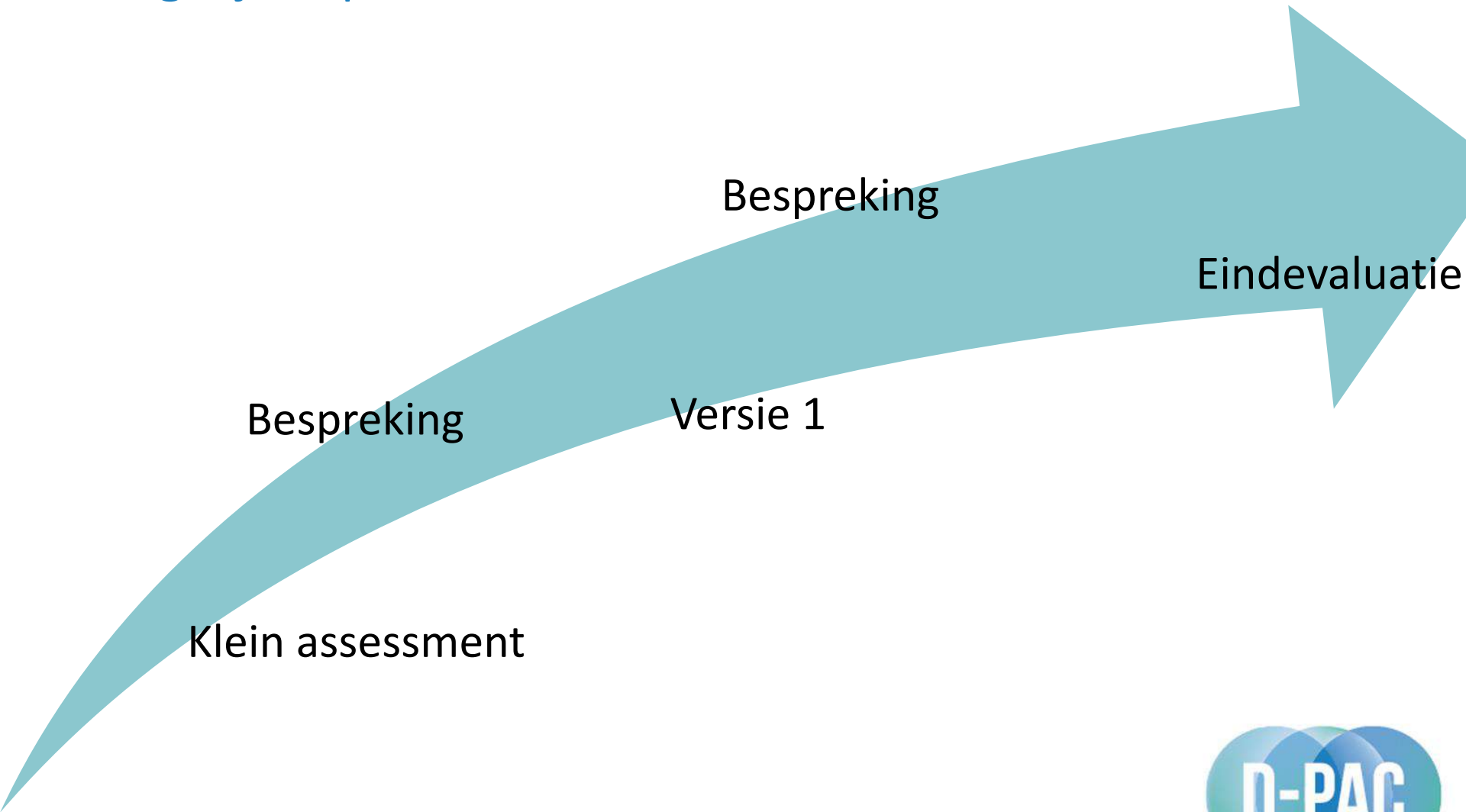
Veilig want alles gaat anoniem

Leerkansen:

- Vergelijken
- Feedback formuleren
- Feedback krijgen
 - Sterke en verbeterpunten
 - Positie in rangorde
- Voorbeelden
- Mediërende rol docent

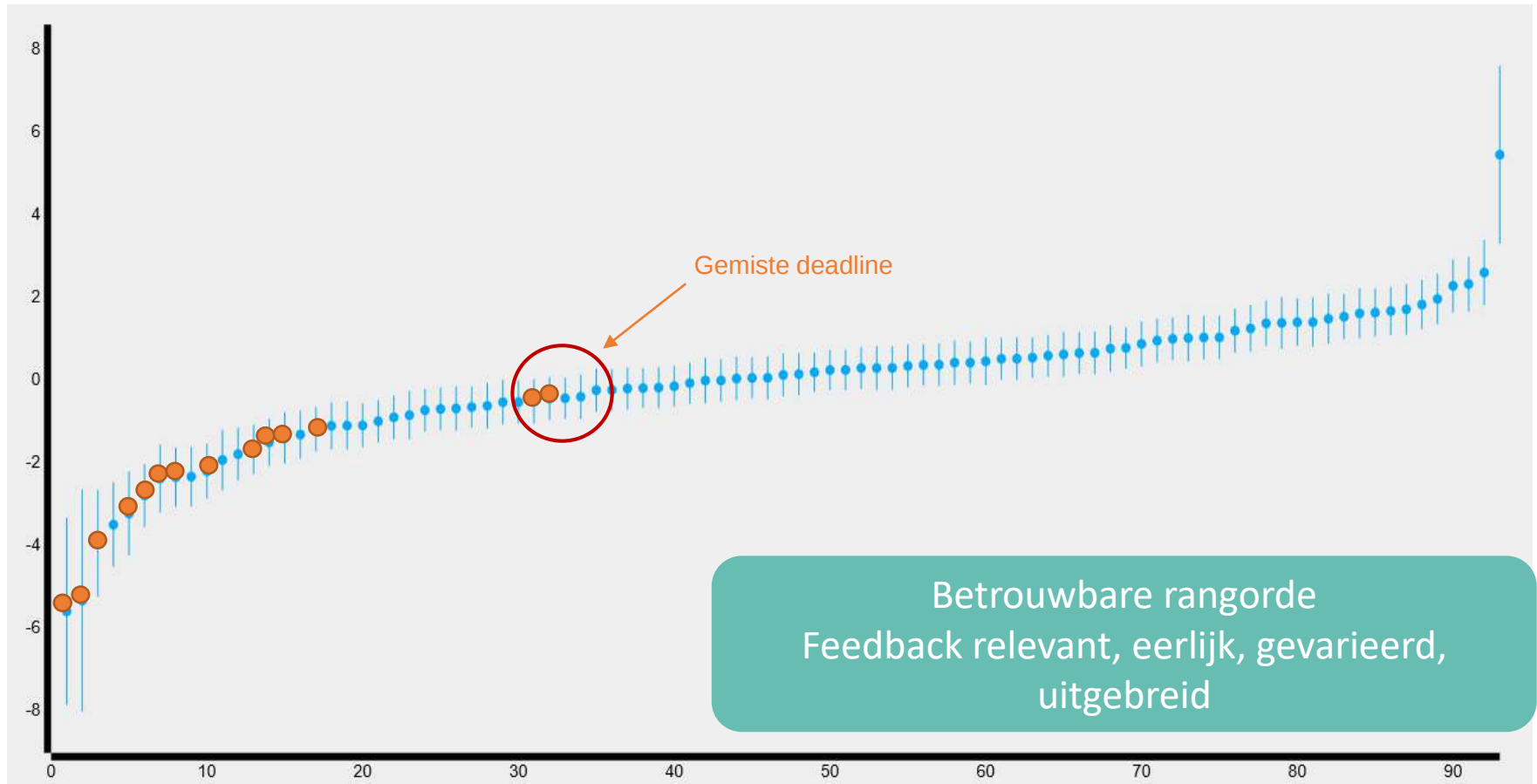
Peer assessment

Mogelijke opzet



Peer assessment

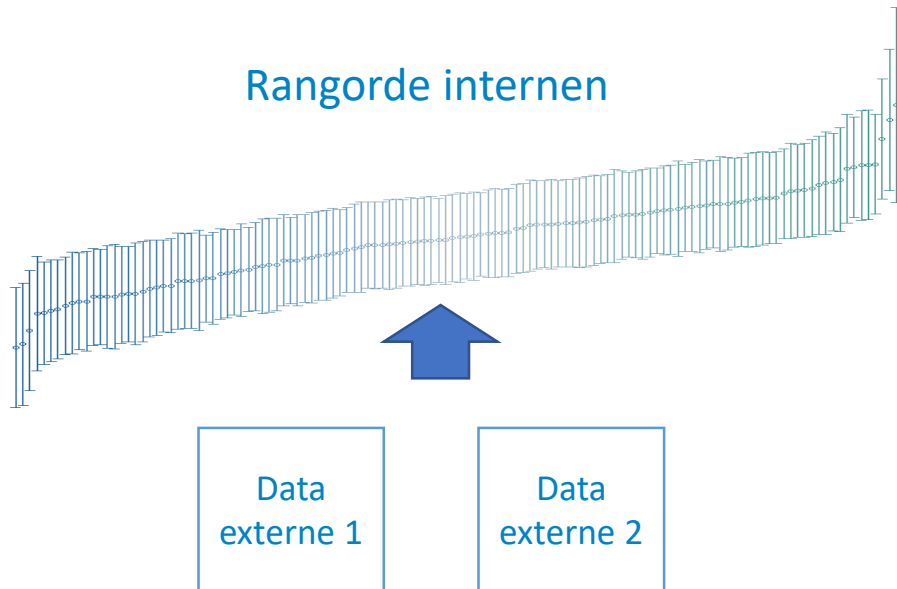
Een tijdswinst voor de docent



Voor de punten

Een focus op professionalisering

Rangorde internen

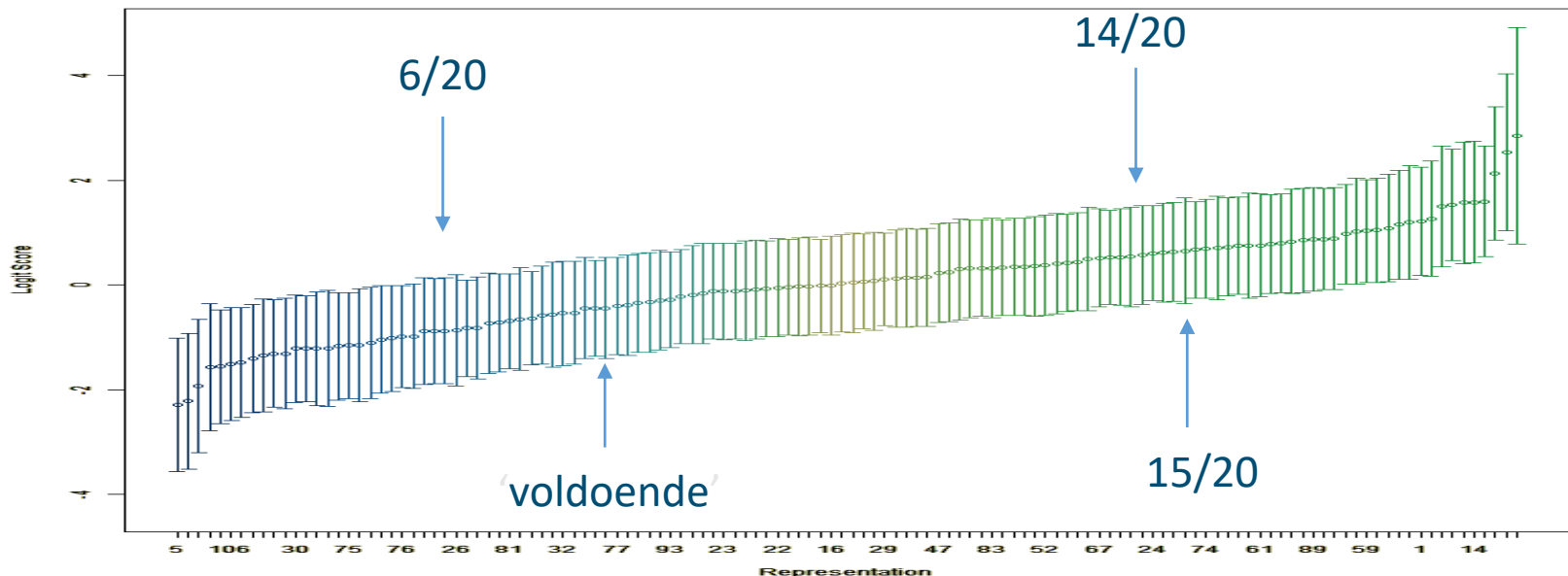


Voorbeeld Ahovoks

- 9 interne beoordelaars
- 16 externe beoordelaars
- 37 teksten
- Pre-test: 5 misfits
- Professionaliseringsdag
- Post-test: 1 misfit

Voor de punten

Naar een cijfer

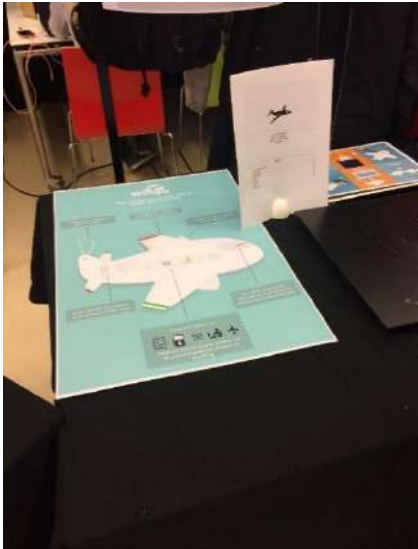


- Op basis van twee gekende scores de rest berekenen
- Vooraf of achteraf

Voor de punten

Jury's

- Beoordelen interactieve installaties
- Beoordelen afstudeerprojecten architectuur



Afronding

Waarom comparatief beoordelen?

☒ Vooraf

- Geen rubric ontwikkeling
- Geen training van beoordelaars nodig

☒ Gedurende het assessment

- Je maakt gebruik van de expertise van de beoordelaars
- Holistische beoordelingen
- Geen orde-effecten
- Gemakkelijk, snel en leerzaam voor peer assessment

☒ Achteraf

- Een valide en betrouwbare rangorde/scores
- Mogelijkheid studenten op te volgen voor peer assessment
- Informatie of afwijkende beoordelaars en beoordelingen

Vragen?



marije@d-pac.be

roos@d-pac.be

www.D-PAC.be