



**Universiteit
Leiden**
The Netherlands

**Design and evaluation of video portfolios : reliability,
generalizability, and validity of an authentic
performance assessment for teachers**

Bakker, M.E.J.

Citation

Bakker, M. E. J. (2008, December 2). *Design and evaluation of video portfolios : reliability, generalizability, and validity of an authentic performance assessment for teachers. ICLON PhD Dissertation Series.* Leiden University Graduate School of Teaching (ICLON). Retrieved from <https://hdl.handle.net/1887/13353>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/13353>

Note: To cite this publication please use the final published version (if applicable).

STELLINGEN

Behorende bij het proefschrift

‘Design and evaluation of video portfolios: Reliability, generalizability, and validity of an authentic performance assessment for teachers’

Mirjam Bakker, 2 december 2008

1. Een acceptabel niveau van betrouwbaarheid is gemakkelijker te bereiken wanneer het gaat om de beoordeling van meerdere performances in verschillende situaties dan wanneer het gaat om de beoordeling van performance in één specifieke situatie (dit proefschrift).
2. Om een acceptabel niveau van betrouwbaarheid te realiseren, kan volstaan worden met het inzetten van twee getrainde beoordelaars (dit proefschrift).
3. Beoordelaars vinden het moeilijk om een ‘voldoende’ of een ‘onvoldoende’ toe te kennen aan een performance, terwijl dit juist bij summatieve assessments een heel belangrijk onderscheid is (dit proefschrift).
4. Selectieve waarnemingen en persoonlijke voorkeuren van beoordelaars zijn niet volledig te vermijden in performance assessments (dit proefschrift).
5. Om tot een valide performancebeoordeling te komen, dient de beoordelaar ook inzicht te hebben in afwegingen en intenties die ten grondslag liggen aan het gedrag (dit proefschrift).
6. De overeenstemming tussen beoordelaars ten aanzien van toegekende scores is geen garantie voor overeenstemming in bewijzen en argumenten die beoordelaars gebruiken om toegekende scores te legitimeren (dit proefschrift).
7. Het feit dat competent handelen in sterke mate wordt gedefinieerd door de context waarin het handelen plaatsvindt, levert een probleem op ten aanzien van de selectie van taaksituaties waarin het handelen getoond moet worden.
8. In tegenstelling tot wat vaak wordt beweerd, is betrouwbaarheid niet alleen van belang bij summatieve assessment (assessment of learning), maar ook bij formatieve assessment (assessment for learning).
9. Het hanteren van heel specifieke scoringsregels bij performance assessments leidt ertoe dat het interpretatief vermogen van beoordelaars wordt ondermijnd.
10. In de (onderwijs)praktijk wordt te snel naar nieuwe vormen van assessment gegrepen zonder de methodologische kwaliteit van deze instrumenten voldoende te waarborgen en te onderzoeken.
11. In topsport zouden coaches meer gebruik moeten maken van informatie uit performance assessments om hun beslissingen te onderbouwen.