Portfolios, the Pied Piper of Teacher Certification Assessments: Legal and Psychometric Issues


Education Policy Analysis Archives
http://epaa.asu.edu/epaa/v11n45/

90% of SCDE’s use portfolios to make decisions about candidates
40% do so as a certification or licensure requirement (Salzman et al., 2002)

“Portfolios as Tests”

- Portfolios used in a high-stakes context are technically testing devices
-need to meet psychometric standards:
  - Validity
  - Reliability
  - Fairness
  - Absence of bias


High Stakes Tests

- Tests used for decisions, such as for employment, licensure, or a high school graduation
- When tests are used for high-stakes decisions, they will be subject to legal scrutiny


Portfolios for Learning

- Excellent tools
  - Reinforcing learning
  - Making formative decisions about teacher candidate
    - Knowledge
    - Skills
    - Dispositions
    - Growth


Requirements and Caveats

- Graduation and certification decisions in a standards-based environment
- Requirements for tests
- Caveats for portfolios

1. Job-Related and Authentic

- Stand the test that the portfolio is job-related and authentic.
- Equivalent to what teachers do with technology in the classroom.

2. Representative, Relevant, Proportional

- “All criteria used to evaluate the portfolio must be relevant to the job [of teaching].”

3. Documentation and Appeals Process

- Adequate documentation in place
- Tells candidates
  - how and when to prepare the portfolio
  - how it will be reviewed
  - who is allowed to help them and how much help they can receive
  - the consequences of failure and the opportunities for remediation
  - what their due process rights and procedures are if they wish to challenge the review results.”

4. Candidate Support System

- Adequate instructional opportunities provided to candidates
  - to succeed in meeting the requirements of the portfolio/test
  - to remediate when performance is inadequate.
- Embed portfolio preparation into instructional program (i.e., coursework).
- “Any requirements outside of the instructional program could be subjected to a claim based on instructional/curricular validity.”

4. Faculty Support System

- “The entire faculty need to buy into, and support, portfolio preparation activities of the students and provide remedial opportunities for components that are found lacking.”

5. Realistic Cut-Score

- The most difficult aspect of portfolio design
- Identify the specific score or characteristics
- Sort teachers into the dichotomous categories of competent and not competent based on their portfolios
6. Alternatives

“Alternatives must be provided to candidates who cannot successfully complete requirements, or the SCDE must be able to demonstrate why no alternatives exist”


7. Protected Populations
8. Reliable Scoring

7. Are a disproportionate number of protected populations (minorities, handicapped, women) unsuccessful? SCDE must show why no alternatives exist.


More limited and focused use of portfolios

- Suggest a more limited and focused use of portfolios
- To measure specific, job-related skills
- Realize long known values of portfolio assessment without burdening faculty and students with excessive requirements that have limited use and without taking serious psychometric and legal risks


Showcase portfolios vulnerable to legal challenge

When used as a measure of job performance themselves, or when evaluated using criteria that are related in only tangential ways to authentic job tasks, or when they are not substantially related to standards required for state program approval, or when they are prepared as an extra-curricular activity, or when they contain student-selected evidence, or when they are not adequately monitored for reliability or bias


Portfolios by Level of Inference

<table>
<thead>
<tr>
<th>Inference Level</th>
<th>Selection</th>
<th>Time</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Item selection; pre-determined</td>
<td>No faculty time, no faculty skill if task rubrics included</td>
<td>Excellent for accreditation and accountability</td>
</tr>
<tr>
<td>Medium</td>
<td>Students have some choice, typically from a list of options</td>
<td>Faculty decide if students made the right choices and assess the composite collection</td>
<td>Tricky for accreditation, especially if no task rubrics included</td>
</tr>
<tr>
<td>High</td>
<td>Students have wide choices</td>
<td>Faculty reassess what was already assessed</td>
<td>Learning and formal evaluation. Not useful for accreditation</td>
</tr>
</tbody>
</table>

(Wilkerson & Lang, 2004)