Alignment between Assessment and Learning analytics

Sandra Milligan
Assessment, innovation & learning analytics

- **Reporting & Credentialing**
- **Integrity Monitoring**
- **Adaptive Testing**
- **Feedback on Learning**
- **Automated Scoring**
- **Items & Task Development**
- **Wider Range of Evidence**
- **Administrative Efficiency**

**What's Assessed**

- Stylometry
- Identity checkers
- Wide range of metrics
- Advisor apps
- Chatbots
- Explainer apps
- Dashboards

**How It's Assessed**

- Block chain
- Digital testamers
- Profiles
- Item generators
- Simulations
- Agents
- Text analysis
- MC
- Short answer
- Process data
- Big data
- Videos
- Responses to agents
- Beautiful portfolios

**Administrative Efficiency**

- Peer assessment, self assessment
- LMS systems
- Remote Proctoring

**Item Generators & Task Development**

- Item generators
- Simulations
- Agents
Assessment innovation & analytics

REPORTING & CREDENTIALING

INTEGRITY MONITORING

ITEMS & TASK DEVELOPMENT

ITEM generators
Simulations
Agents

AUTOMATED SCORING

Text analysis
MC
Short answer

WIDER RANGE OF EVIDENCE

Process data
Big data
Videos
Responses to agents
Beautiful portfolios

WHAT'S ASSESSED

FEEDBACK ON LEARNING

Wide range of metrics
Advisor apps
Chatbots
Explainer apps
Dashboards

How it's assessed

ADAPTIVE TESTING

Stylometry
Identity checkers

ADMINISTRATIVE EFFICIENCY

Peer assessment, self assessment
LMS systems
Remote Proctoring

Block chain
Digital testamers
Profiles

Block chain
Digital testamers
Profiles
‘First-movers’

• Define ‘success’ of learners broadly, encompassing the whole person (schools) or deep competence (disciplines & professions)

• Allow passion or interest driven learning,, context-specific, non-standardised approaches to teaching and assessment

• Want to see changes to the currency of recognition, and metrics for accountability, to sustain reforms

• Seek to ‘re-form’, not just to continuously improve the organisation of learning and teaching

• Want it all scalable and practical
1. Contemporary ambitions for learning

Learner agency:
Capacity to chart their own learning towards expertise in a chosen path, taking risks, investing in learning to attain their purposes, harnessing interests, and taking responsibility for the results attained.

Connectivity:
Capacity to create and sustain valuable connections to support induction into wider communities in which they will participate and contribute as workers, community members, and citizens.

Knowledge and knowhow:
Mastery of the terminology, concepts, theories, structures and processes that make up the disciplines, vocations and cultural domains and application of knowledge through building, designing, providing services, performing, growing or creating things of value.

Learning staples:
Including critical analysis and evaluation, teamwork and collaboration, problem solving, creativity, capacity for resilience and self-care, intercultural capability, and entrepreneurial skill.

Basic literacies:
Literacy, numeracy, and the digital literacies.
Learning ambitions frameworks
2. New kinds of credentials and reports: standards-referenced profiling
3. Assessment of complexity: not just adding things up

Performance in a complex domain arises from the integration of knowledge, knowhow, attitudes, values and beliefs: a uni-dimensional scale. It’s the integration we are after, not the addition.
3. Good quality scales: qualitative, not quantitative shifts

Consumes content from teaching staff, seeks generic knowledge

Level 1 Reader

Level 2 Consumer

Level 3 Self-regulator

Level 4 Collaborator

Level 5 Reciprocal teacher

Independent learner, seeks expertise in a domain, systematic, persistent, self-evaluates, reflective, practices, uses automated teaching agents, monitors peers, may share

Collaborative learner, seeks practical wisdom in a domain, wide attention span, engaged, dialogic, risk-taker, independent-minded, critical consumer, evaluates peer input, seeks other perspectives, produces, writes, creates, comments, teaches, supports & mentors others

Copyright s.milligan@unimelb.edu.au
Each ‘X’ represents 22.3 cases.
Assessment of individual mentation*
using common, invigilated standard performance tasks
freed from purpose, personal interest, language, social and cultural context
uncomplicated by real-world conditions
scored to remove individual judgements by assessors
to position the learner
on a numerical scale
to separate people by degree of proficiency
or to select candidates for next steps.
Assessment of complex competencies
By gathering and analysing diverse evidence
(what learners do, say, make or write)
during complex authentic performances
embedded with purpose, personal interest, language, social and cultural context
to support a judgment
referenced to agreed standards
trusted by those concerned
to position the learner
on a scale of competence
from less expert to more
showing what learners know & can do & who they are
and what they need to learn next
to guide further learning
Dominant paradigm: at its best
Assessment of individual mentation*
using common, invigilated standard performance tasks
freed from purpose, personal interest, language, social and cultural context
uncomplicated by real-world conditions
scored to remove individual judgements by assessors
to position the learner
on a numerical scale
to separate people by degree of proficiency
or to select candidates for next steps.

Frontier paradigm
Assessment of complex competencies
By gathering and analysing diverse evidence
(what learners do, say, make or write)
during complex authentic performances
embedded with purpose, personal interest, language, social and cultural context
to support a judgment
referenced to agreed standards
trusted by those concerned
to position the learner
on a scale of competence
from less expert to more
showing what learners know & can do & who they are
and what they need to learn next
to guide further learning
or to select candidates for next steps.
Plausibility checks

1. What’s measured is what’s intended?  
   Depth/competence vs recall/comprehension

2. Higher score mean greater proficiency?  
   \[ S_{\text{Total}} = S_{\text{Sought}} + S_{\text{Irrelevant}} + \text{Error} \]

3. Learning represented as change?  
   Learning = growth, development of proficiency

4. Engagement is used as a proxy for proficiency?  
   Proxies can drive the wrong behaviour

5. Feedback is about actual learning?  
   Only the ‘right feedback is good”

6. More data: is it better?  
   Best = the right data, single scale

7. High correlations are represented as good?  
   High correlations: redundancy, lack of impact

8. Normal distributions should not be normal  
   Good teaching destroys normal curves
Key frontiers:
How to generate and represent warrantable assessments of complex competencies

Reliable and valid assessment of complex competencies (low error, right learning)
Combining different kinds of evidence to capture proficiency
Representing this kind of learning on scales that demonstrate qualitative shifts
Supporting the scalable application of human judgment
Establishing standards inherent in complex performances and products
Context responsiveness while retaining validity and reliability
Thank you