

Assessment in the Technology Age: Promising but not yet delivered

AiTELL Conference, online SISU, November 2021

Prof. Gavin T. L. Brown



**EDUCATION AND
SOCIAL WORK**

1



**EDUCATION AND
SOCIAL WORK**

Assessment Purposes

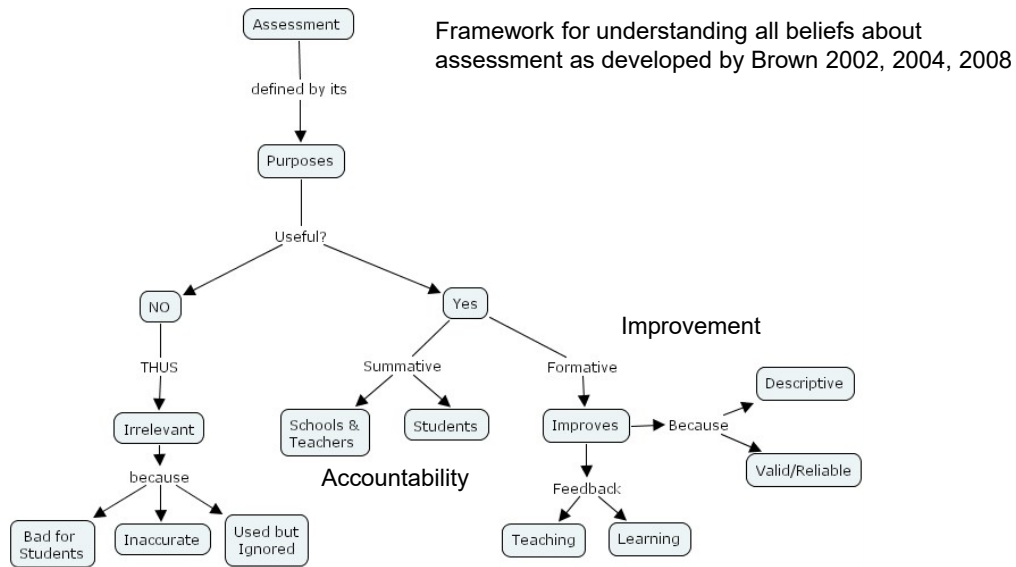
➤➤ Competing Tensions

3

Purposes of Assessment



EDUCATION AND
SOCIAL WORK

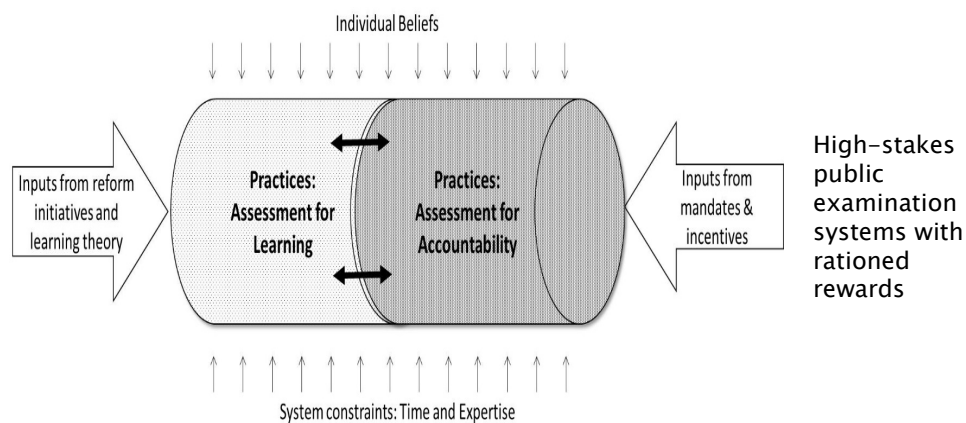


4

Assessment pushes in a single system

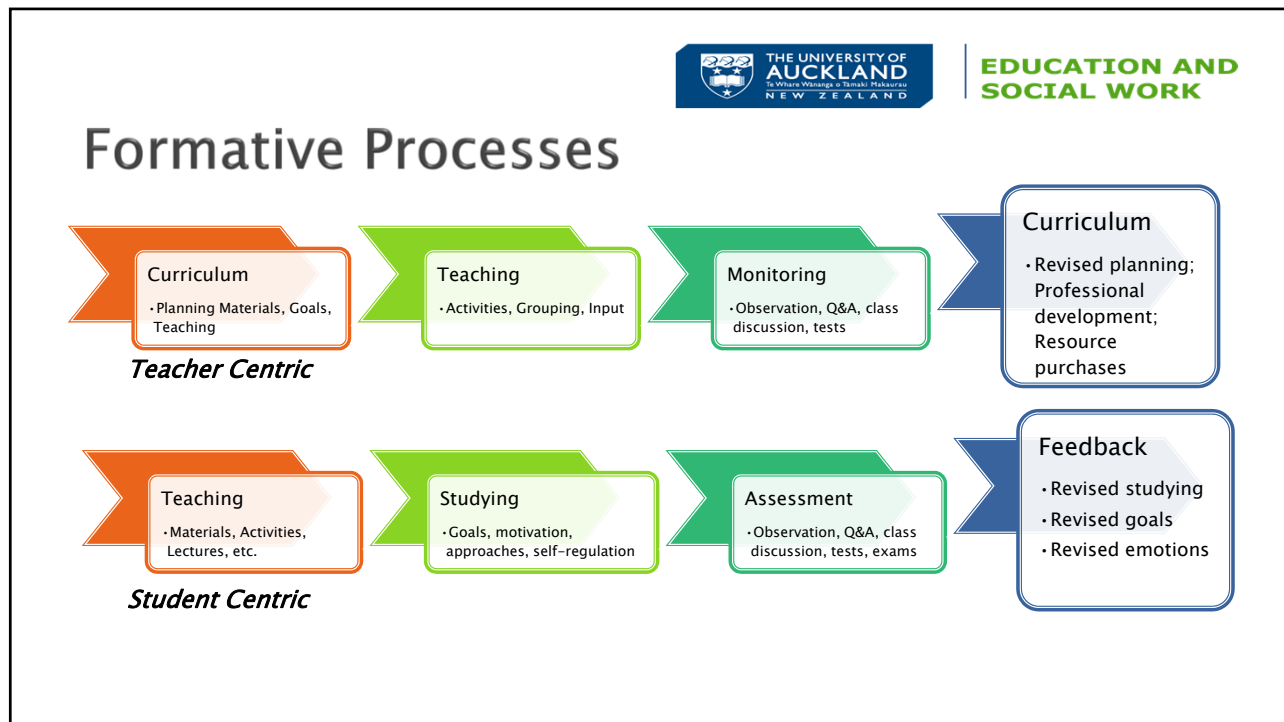


EDUCATION AND
SOCIAL WORK




Bonner, S. M. (2016). Teachers' perceptions about assessment: Competing narratives. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 21–39). New York: Routledge.

5

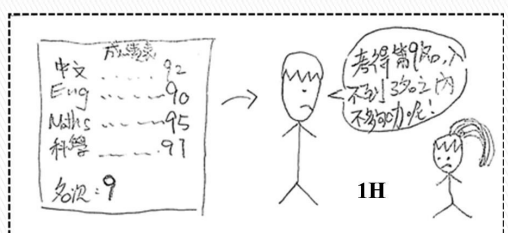


6


EDUCATION AND SOCIAL WORK

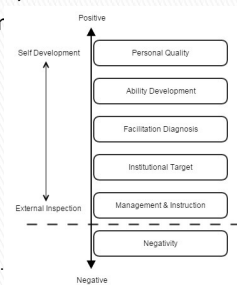
Even in China

- ▶ *zhong kao & gao kao*
- ▶ Social consequences for scores
- ▶ Moral imputation for scores
- ▶ Brown, G. T. L., & Wang, Z. (2013). Illustrating assessment: How Hong Kong university students conceive of the purposes of assessment. *Studies in Higher Education*, 38(7), 1037-1057. <https://doi.org/10.1080/03075079.2011.616955>



Summative Exams

- ▶ All-round development of good character and good person attributes
 - mid-1950s: the 'Three Goods' (i.e. ideology and morality, study and physical health)
 - 1980s: The 'Five Loves' (i.e. motherland, people, labour, science and socialism)
 - 2001: New Basic Curriculum reforms/*Integrated Quality Assessment*
 - Prof. Gao Lingbiao, South China Normal University
 - Brown, G. T. L., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, 2(1). <https://doi.org/10.1080/2331186X.2014.>



Formative Character Development

7

So what?

- ▶ You have to choose what your priority is and what you can live with
 - Help teachers do best possible job
 - Help students achieve best in the world
 - Identify the best and reward them
 - Identify and get rid of poor teachers & leaders
- ▶ Whatever you choose will impact your results
 - Best results: focus on IMPROVEMENT, not Evaluation/Punishment/Selection/Reward

8

Technologies

➤➤➤ Are they a good solution?

9

Psychometric Technologies



EDUCATION AND
SOCIAL WORK

- ▶ Rater effects
 - multi-facet Rasch modeling; generalizability theory; inter-rater reliability
- ▶ Sub-scores
 - Cognitive diagnostic models; testlet models; polytomous scoring
- ▶ Equivalence between groups
 - Differential item functioning; multigroup invariance testing
- ▶ Faster & more reliable marking
 - Online administration; Automated essay scoring
- ▶ More accurate ability estimation
 - Computer adaptive testing; Effort estimation; cheating detection

10

New Testing Technologies



EDUCATION AND
SOCIAL WORK

- Automatic Item Generation
- Technology-enhanced items cf. ETS
- Flexible test scheduling
 - Time zone control
- Adaptive testing
- “Objective” and “quick” scoring
- Automated text marking
 - Natural Language Processing
 - **GPT-3** new text writer—not marker
 - Your students can write good essays by machine –could you tell?



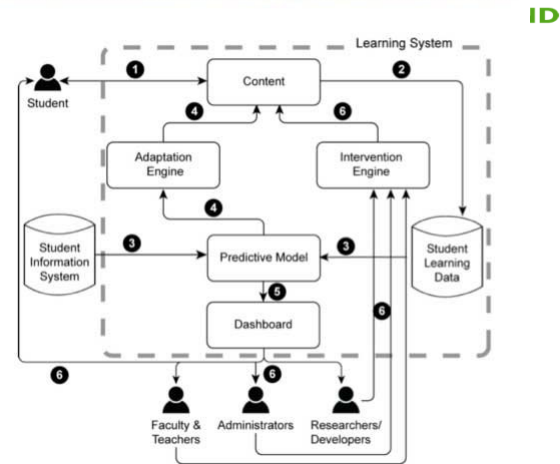
Katz, I. R., & Gorin, J. S. (2016). Computerising assessment: Impacts on education stakeholders. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (pp. 472-489). Routledge.

11

Learning Analytics

- ▶ Turning data into usable knowledge
 - linking data types into applications for improving processes and products
 - Using statistical engines and techniques
- ▶ But...predictive model & engines are a 'black box' of atheoretical(?) statistical analysis

The Components and Data Flow Through a Typical Adaptive Learning System



<https://stephenp.net/2012/04/15/enhancing-teaching-and-learning-through-educational-data-mining-and-learning-analytics-observations/>

12

BUT data do not explain themselves

- ▶ *causal questions can never be answered from data alone. They requires us to formulate a model of the process that generates the data, or at least some aspect of that process.*

- Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. New York: Hachette Book Group.

THE UNIVERSITY OF
AUCKLAND
SCHOOL OF EDUCATION
NEW ZEALAND

**EDUCATION AND
SOCIAL WORK**

CLIMATE CONTROL

You have no instructions for your new air conditioner. You need to work out how to use it.

You can change the top, central and bottom controls on the left by using the sliders (←→). The initial setting for each control is indicated by ▲.

By clicking 'APPLY', you will see any changes in the temperature and humidity of the room in the temperature and humidity graphs. The box to the left of each graph shows the current level of temperature or humidity.

Top Control

Central Control

Bottom Control

Temperature 33

Humidity 25

APPLY

RESET

optimal exploration strategy (i.e., vary-one-thing-at-a-time)

Question 1: CLIMATE CONTROL CP025Q01

Find whether each control influences temperature and humidity by changing the sliders. You can start again by clicking RESET.

Draw lines in the diagram on the right to show what each control influences. You can remove any line by clicking on it.

Top Control

Central Control

Bottom Control

Temperature

Humidity

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education*, 91, 92-105. <https://doi.org/10.1016/j.compedu.2015.10.018>

13

Role for AI?

- ▶ Efficiently identify strengths & weaknesses in performance based on rules learned from previously scored work
 - *Supervised machine learning algorithms*
- ▶ More quickly provide analysis to generate feedback
 - *Machines good at delivery, if we program them*
- ▶ Fast machine feedback better than slow human?
 - *Maybe yes, but tell me something I need to know*
- ▶ BUT
- ▶ Requires task to be on a computer
- ▶ Requires valid scoring model and theory of performance.

14

Online requires Security

- ▶ With online assessment that matter there is always the possibility of gaining higher marks in an illegitimate manner
- ▶ Solutions include:

Approach	Security feature	Intrusiveness	Example
Surveillance	Verify identity, monitor outsourcing & activity	HIGH	Remote proctored exams
Lockdown	Disable features on devices/networks	HIGH	Exam OS, lockdown browsers
Stylometrics	Compare writing style to previous work	Moderate	Text-matching vendors
Content matching	Compare to all previous work	Low	Text-matching (Turnitin), Image matching

What can you afford financially and in terms of impact on participant psychology? Can you be sure it can't be broken? Honesty is the goal.

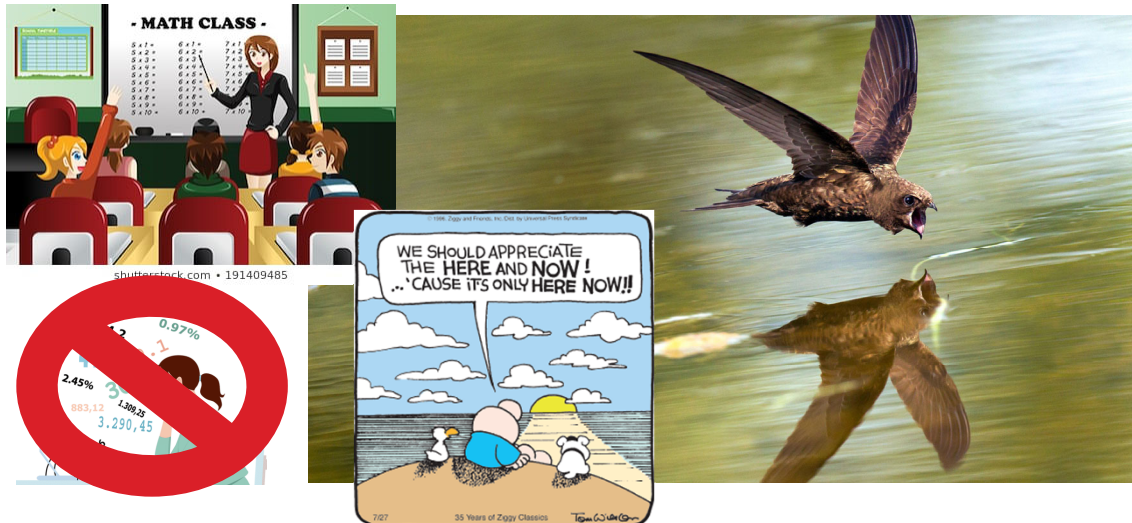
Dawson, P. (2021). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. London: Routledge. (p.27)

15

Classroom Assessment



EDUCATION AND
SOCIAL WORK



- *What should I do/teach next and to whom?*

16

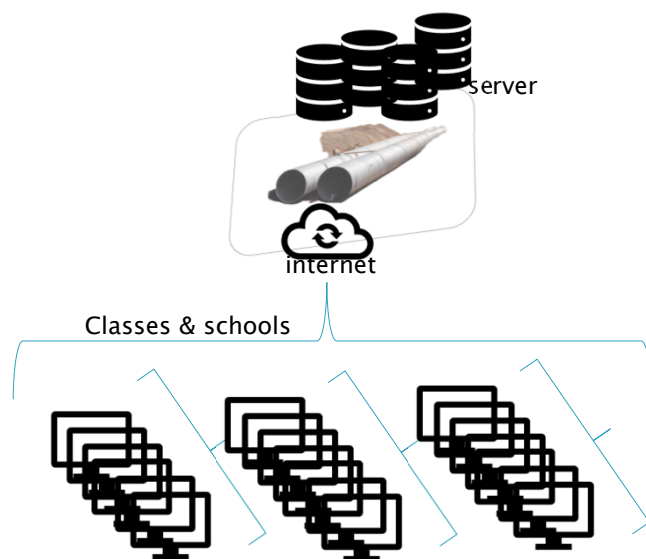
Infrastructure



EDUCATION AND
SOCIAL WORK

- ▶ Large screen computers
- ▶ Lots of machines for all
- ▶ High speed, high volume internet
- ▶ Robust powerful & secure servers
- ▶ Challenge
 - Who pays so that all have what the system requires?

Brown, G. T. (2019). Technologies and infrastructure: costs and obstacles in developing large-scale computer-based testing. *Education Inquiry*, 10(1), 4–20.
<https://doi.org/10.1080/20004508.2018.1529528>



17

The immediate future



EDUCATION AND
SOCIAL WORK

- ▶ Probably not computerised schooling
- ▶ But computers to help teachers and students in vivo
- ▶ Technologies that help in classroom assessment, not take over classrooms.
 - Complementary technologies from New Zealand/Aotearoa



18

Test weakness



EDUCATION AND
SOCIAL WORK

- ▶ Total score
 - No diagnostic information, everything is equally important
- ▶ Rank order score
 - No diagnostic information, what do the top/middle/low students need to do to improve?
- ▶ Curriculum/Content Alignment
 - General proficiency may not fit well with 'my class'
- ▶ Timing
 - Too late doesn't help now
- ▶ Teacher Communication
 - Requires ability to read statistics

19

NZ Solution: e-asTTle



EDUCATION AND
SOCIAL WORK

- ▶ Funded NZ Ministry of Education 2000–2008
- ▶ Incremental development
 - Local computers → local area networks → web application
 - Paper → screen → computer adaptive
 - Primary → Secondary
- ▶ Designed for teachers/school leaders to know
 - How are we doing?
 - Who needs what?
 - Who can go with whom?
 - What materials can help?

20

E-asTTle reports



EDUCATION AND
SOCIAL WORK

Console Report for Test: Entrance Test Eng 2004
Group: All Test Candidates
Date Tested: 11 November 2003

Curriculum Levels Report for Test: Entrance Test Eng 2004
Group: All Test Candidates
Date Tested: 11 November 2003

Interaction Effects

What Next Report for Test : help guide-customis
Group : All Test Candidates

Processes and Strategies Purposes and Audience

Level	Processes and Strategies	Purposes and Audience
6 Advanced		
6 Proficient		
6 Basic		
5 Advanced		
5 Proficient		
5 Basic		
4 Advanced		
4 Proficient		
4 Basic		
3 Advanced		
3 Proficient		
3 Basic		
2 Advanced		
2 Proficient		
2 Basic		

<http://esttle.org.nz/whatsnext/reading>

TKI TE Kete Hīkari

Communities Schools Search community

Search all of TKI

What Next search results

Keywords

Search within results Search all What Next resources

28 results were returned

Fact or opinion - WL3188

Assessment Resource Bank

This Level 2 Assessment Resource Bank task is about whether some statements are fact or an opinion. Students identify statements of fact from a range of statements.

<https://arba.nce.org.nz/resources/fact-or-opinion>

View full metadata record

Features of Text Forms: Instructions

Guidelines

This English Online page describes the features of instructional text. Includes the purpose, types of instructions, specific features, and language used. Provides a checklist.

http://enabonline.ki.org.nz/E_english-of-text-forms/Instructions

View full metadata record

Features of text forms: Descriptions

Guidelines

This English Online page describes the features of a description. Includes the purpose, types of descriptions, specific features, and language used. Provides a checklist.

http://enabonline.ki.org.nz/E_english-of-text-forms/Descriptions

View full metadata record

Features of text forms: Explanations

Guidelines

This English Online page describes the features of an explanation. Includes the purpose, types of explanations, specific features, and language used. Provides a checklist.

http://enabonline.ki.org.nz/E_english-of-text-forms/Explanations

View full metadata record

Applied filters

Learning area: English [X]

Curriculum level: 2 [X]

Strand: Listening, reading and viewing [X]

Sub-strand: Purposes and audiences [X]

Clear all [X]

Year level

Type of resource

Brown, G. T. L., O'Leary, T. M., & Hattie, J. a. C. (2019). Effective reporting for formative assessment: The asTTle case example. In D. Zapata-Rivera (Ed.), *Score reporting research and applications* (pp. 107-125). Routledge.

21

Involving students



EDUCATION AND
SOCIAL WORK



<https://peerwise.cs.auckland.ac.nz/>

- ▶ Self-testing helps learning
- ▶ PeerWise (Paul Denny, U. Auckland)
 - Free
 - Students write questions
 - Students answer each other's questions
 - Students evaluate each other's questions
 - Students who do more questions tend to learn more
 - Students can learn from peer feedback
 - But only MCQ—but that might not be a problem?

22

E-portfolios



EDUCATION AND
SOCIAL WORK



Collect and Select for a
purpose

Breadth
Depth
Development



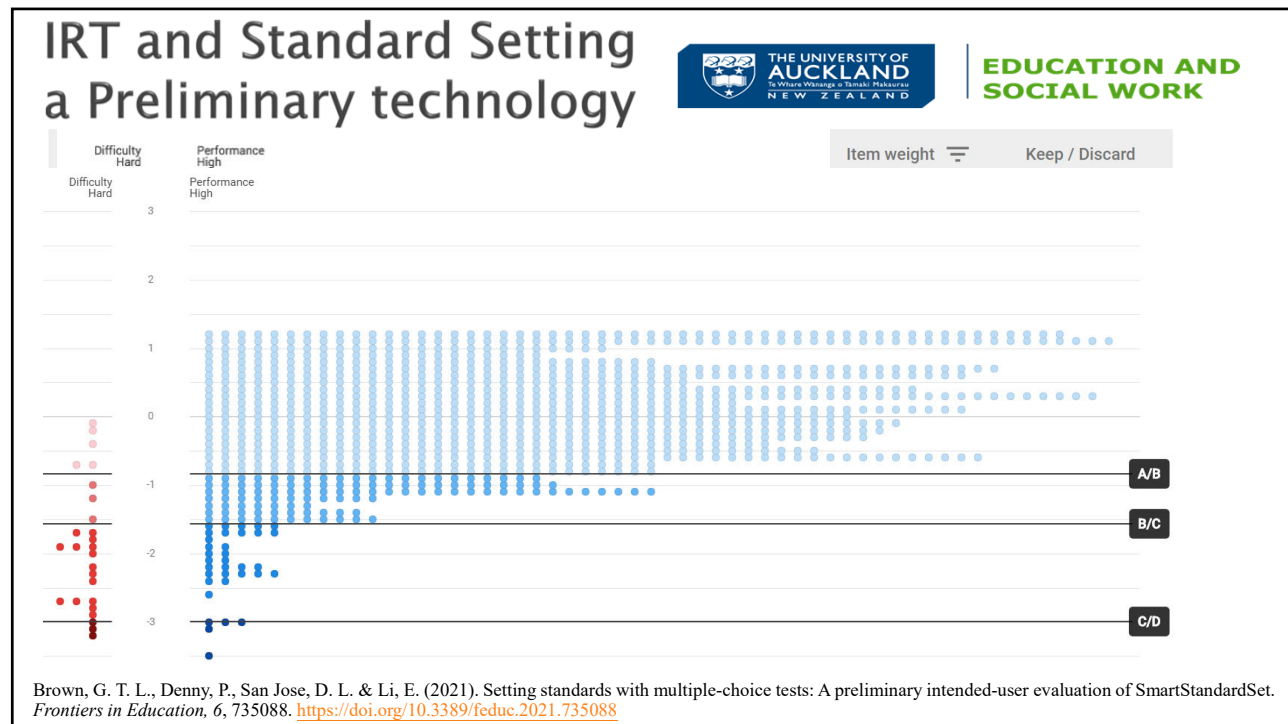
Require students to write a reflection that shows how the
material in the portfolio achieves those 3 purposes



Unfortunately,

little support for feedback (peers, instructors),
little insight as to quality

23



24

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

EDUCATION AND SOCIAL WORK

Lessons for Higher Education

»»

25



EDUCATION AND
SOCIAL WORK

Clarity of Objectives for Assessment

- ▶ What do you want?
 - High quality, high-stakes, secure terminal evaluations?
 - Evaluations that mimic authentically work beyond university gates?
 - Assessments that
 - give students useful feedback so they improve?
 - inform instructors as to what was learned or not?
 - are validly aligned with intended learning outcomes?
 - don't rely on just one format?
 - are easy to administer and give accurate scores?
 - elicit higher order or deep cognitive abilities?
 - involve students in making judgments about quality?

Impossible to have everything in one test or system?

26



EDUCATION AND
SOCIAL WORK

Valuable, difficult, time-consuming PD

- | | |
|---|---|
| <ul style="list-style-type: none"> ▶ How to design assessments for <ul style="list-style-type: none"> ◦ validity ◦ authenticity ◦ reliability ▶ How to score assessments ▶ How to use assessments formatively ▶ How to give feedback ▶ How to involve students in assessment appropriately | <ul style="list-style-type: none"> ▶ How to see the learning behind the assessment performance ▶ How to receive and use feedback ▶ How to give and receive peer evaluative judgments ▶ How to evaluate their own work realistically (no self-deception) |
|---|---|

Instructor professional
development

Student development

27

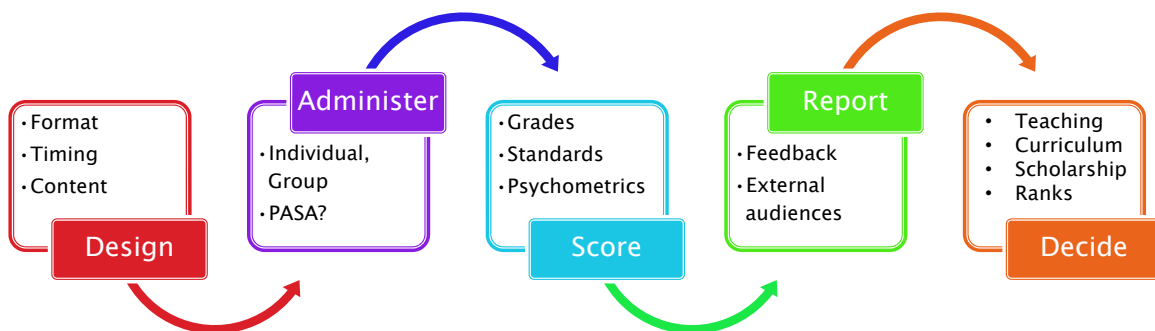
Technology Development Principles

Valuable Information	Develop/select technologies that inform the teacher/tutor/instructor concerning <i>who needs what</i> and learners learn <i>what next</i>
Access & Equity	Ensure low-cost access to sufficient appropriate technologies for ALL • Technology depends on infrastructure of equal opportunity prior to entry to Higher Education
Incremental	Start with low-hanging fruit that have proven benefit, but aim high • easy to use technologies: e.g., PeerWise

Hattie, J. A., & Brown, G. T. L. (2008). Technology for school-based assessment and assessment for learning: Development principles from New Zealand. *Journal of Educational Technology Systems*, 36(2), 189-201. doi:10.2190/ET.36.2.g

28

The whole process matters



Machines & Technology still can't do all of this. You need skilled humans. Well-designed system can have positive impact.

29

Assessment in the Technology Age: Promising but not yet delivered

AiTELL Conference, online SISU, November 2021

Prof. Gavin T. L. Brown



**EDUCATION AND
SOCIAL WORK**