Programmatic assessment for learning, meaningful personalisation of assessment

23d Thai Medical Education Conference

Flinders University

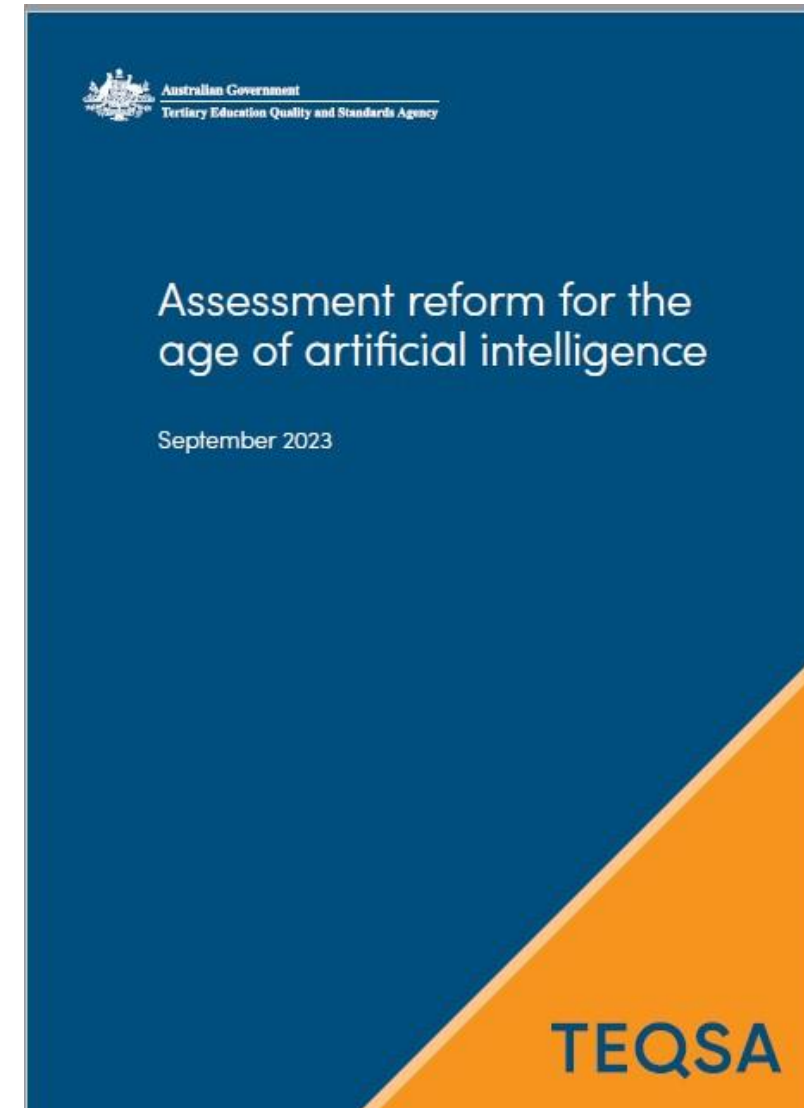**Prideaux**
Discipline of
Clinical Education

Acknowledgements

- Dr Gillian Kette
- Dr Nyoli Valentine
- Prof Cees van der Vleuten

Directions for assessment reform

All Australian higher education to submit **credible** and **concrete action** plans by June 2024
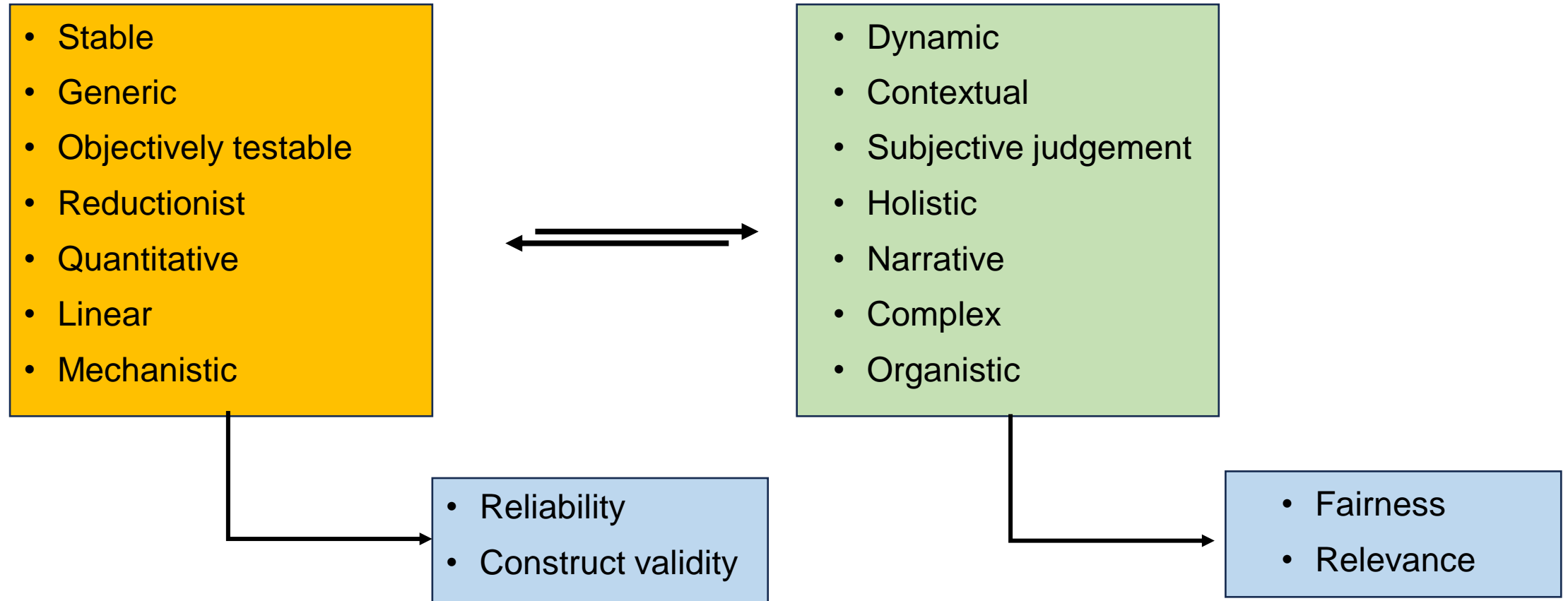
Why the urgency?

1 Changes in medical education

The nature of competence

- Stable
- Generic
- Objectively testable
- Reductionist
- Quantitative
- Linear
- Mechanistic

- Dynamic
- Contextual
- Subjective judgement
- Holistic
- Narrative
- Complex
- Organistic

- Reliability
- Construct validity

- Fairness
- Relevance

My study

PBL groups

Lectures

Practicums

Clerkships/clinical rotations

MCQs/OEQs

Oral examinations

OSCEs

Essays

Current students

TBL groups

(Recorded/streamed) lectures

Practicums

Clerkships/clinical rotations

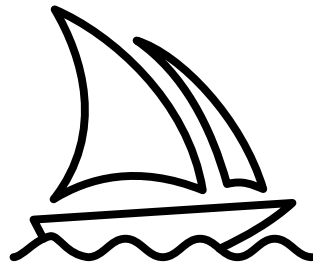MCQs/OEQs

Oral examinations

OSCEs

Essays

Portfolio

Prideaux
Discipline of
Clinical Education

My study

Current students

diary   clock   camera

messenger   movie editor

meteorologist photo album

scratch pad   games

task manager   library

hifi stereo

fitness coach atlas

And now....generative AI

How can I help you today?

SciSpace Copilot

wiseone

DALL·E 3

Prideaux
Discipline of
Clinical Education

2   Changes in technology

Prideaux
Discipline of
Clinical Education

Changes in the world

- Democratisation of knowledge

- Democratisation of transactions/trust

- Democratisation of teaching

- Democratisation of communication

- Democratisation of availability

- Democratisation of intelligence/cognition

How can I help you today?

Flinders University

Prideaux
Discipline of
Clinical Education

modern student 'affordances'

Communication/communities
Collaborations
Creations
Convergence

# 3	Problems with current assessment

The nature of competence

- Stable
- Generic
- Objectively testable
- Reductionist
- Quantitative
- Linear
- Mechanistic

- Dynamic
- Contextual
- Subjective judgement
- Holistic
- Narrative
- Complex
- Organistic

- Reliability
- Construct validity

- Fairness
- Relevance

Let's explore the current situation

- Traits are stable and generic characteristics
- Individual items are meaningless in themselves
- Only sum scores determine what the assessment measures
- Statistics are based on elimination of information
- One best assessment method for each trait

Traits are stable and generic characteristics

knowledge	skills	problem solving	attitude
TEST	TEST	TEST	TEST
TEST	TEST	TEST	TEST

Prideaux
Discipline of
Clinical Education

testing: theoretical situation

TEST

	1	2	3	4	R
A					4
B					3.2
C					2.4
D					1.6
E					0.8

PARALLEL TEST

	1'	2'	3'	4'	R'
A					4
B					3.2
C					2.4
D					1.6
E					0.8

Flinders University

Prideaux
Discipline of
Clinical Education

testing: real situation

TEST

	1	2	3	4	R
A					3.2
B					2.6
C					2.6
D					2.6
E					1.4

PARALLEL TEST

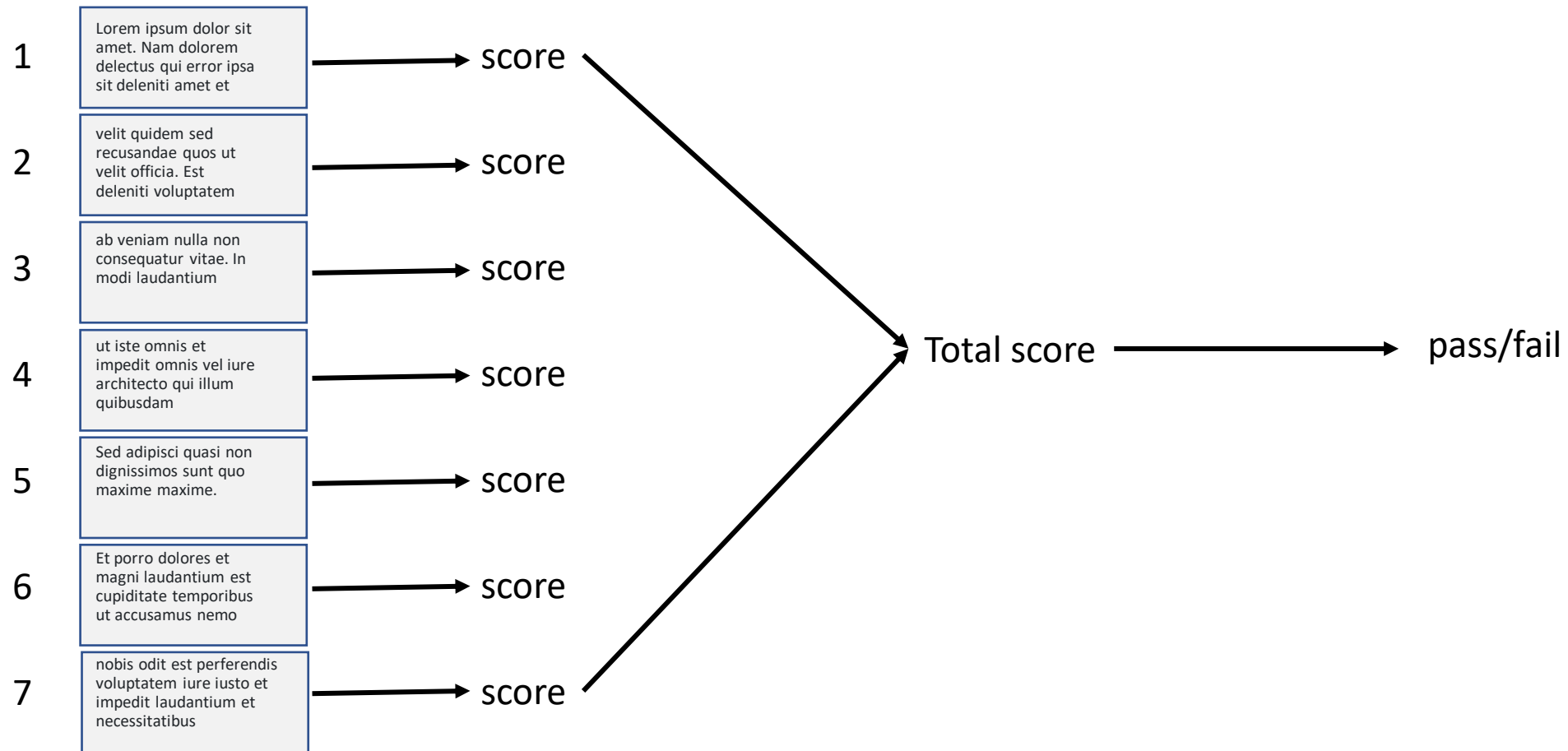|	1	2	3	4	R
A					2.6
B					3.2
C					2.8
D					2.2
E					2.4

Let's explore the current situation

- Traits are stable and generic characteristics
- Individual items are meaningless in themselves
- Only sum scores determine what the assessment measures
- Statistics are based on elimination of information
- One best assessment method for each trait

Flinders University

Prideaux
Discipline of
Clinical Education

Individual items are meaningless in themselves

Mr Jones is 35 years old. He visits his family physician because of chest pain. The pain comes and goes and can be improved by deep breathing or a change of posture. What is at this moment the most likely origin of his pain??

a   the chest wall;

b   the lungs;

c   the heart;

d   the oesophagus.

Individual items are meaningless in themselves

Ms. Smit is 72 years old. She has chest pains. Several times her blood pressure is taken and found to be 170/100 mmHg.
Which antihypertensive drug is most indicated for her??

a	captopril.
b	chlorthalidone.
c	metoprolol.

Individual items are meaningless in themselves

resuscitation
'station'
in an OSCE

Individual items are meaningless in themselves

communication
'station'
in an OSCE

Let's explore the current situation

- Traits are stable and generic characteristics
- Individual items are meaningless in themselves
- Only sum scores determine what the assessment measures
- Statistics are based on elimination of information
- One best assessment method for each trait

Total scores

Let's explore the current situation

- Traits are stable and generic characteristics
- Individual items are meaningless in themselves
- Only sum scores determine what the assessment measures
- Statistics are based on elimination of information
- One best assessment method for each trait

Flinders
University

Prideaux
Discipline of
Clinical Education

elimination of information

Let's explore the current situation

- Traits are stable and generic characteristics
- Individual items are meaningless in themselves
- Only sum scores determine what the assessment measures
- Statistics are based on elimination of information
- One best assessment method for each trait

old model of medical competence

knowledge	skils	problem solving	attitudes
TEST > TEST	TEST > TEST	TEST > TEST	TEST > TEST

Prideaux
Discipline of
Clinical Education

And now…..generative AI

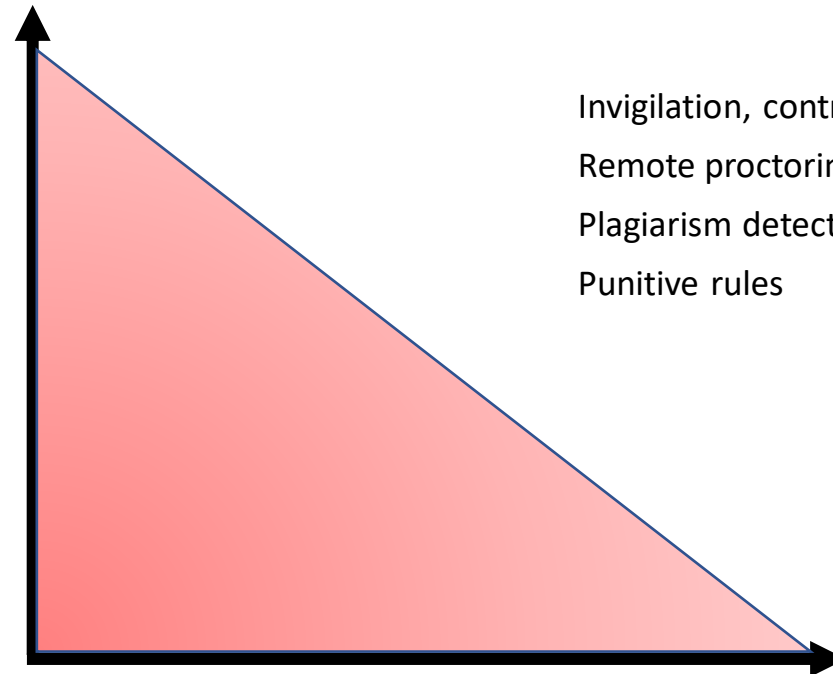# 4   Value proposition and process

Value proposition and processes

The crux of the problem?

Lifelong learning
Independent learning
Competence
Collaboration
Managing complexity
Managing 'not knowing'

Invigilation, control
Remote proctoring
Plagiarism detection
Punitive rules

education

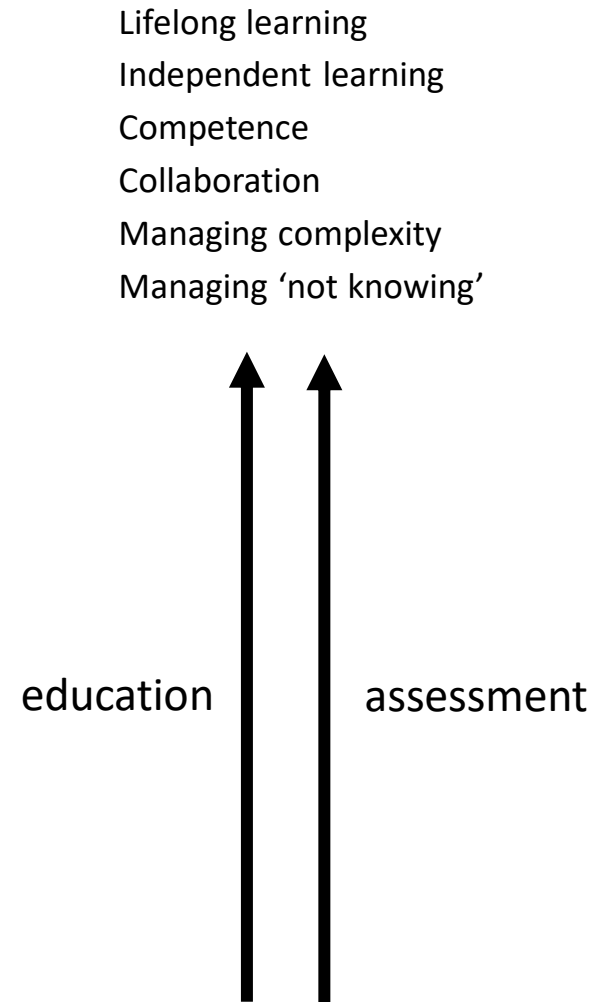assessment

Individual competence
Prescribed, predictable learning
Discriminating/distinguishing
Finite learning
Marks/rankings

The solution?

Lifelong learning

Independent learning

Competence

Collaboration

Managing complexity

Managing 'not knowing'

education   assessment

5 Assessment for learning

Assessment for learning ≠ just formative assessment

Feedback

Learner

Teacher

Organisation

Prideaux
Discipline of
Clinical Education
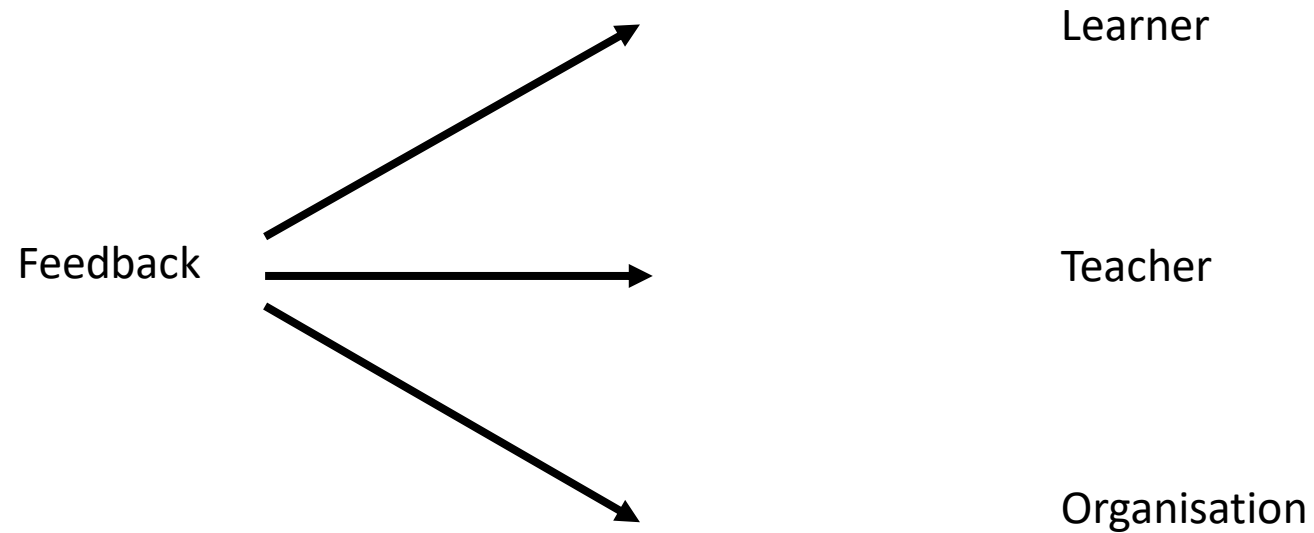
Assessment for learning
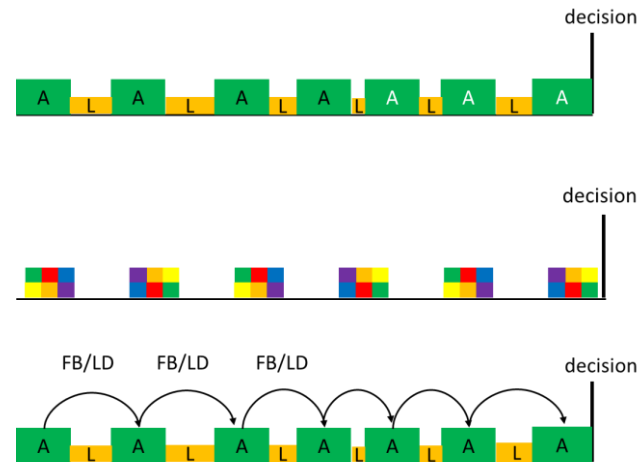
- Content

  - meaningful

  - active

  - collaborative

Assessment for learning

- Process:

  - Distributed

  - Interleaved
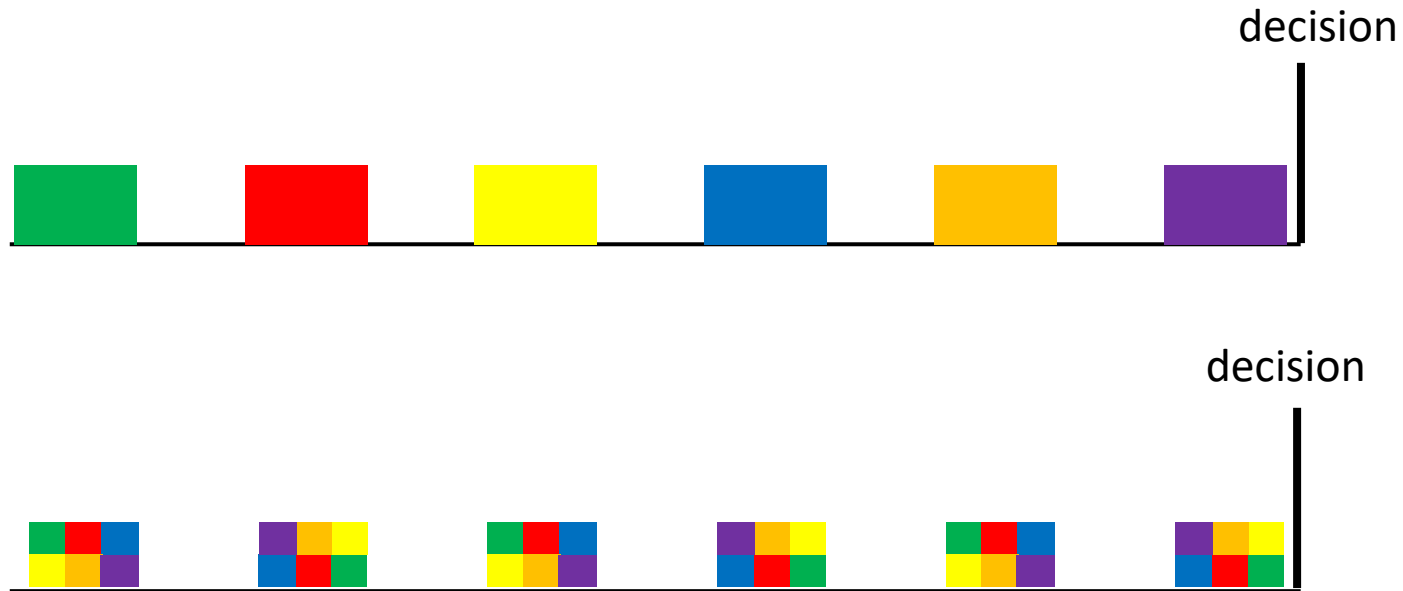
  - Connected

Assessment for learning: distributed

test

Assess-ment

Study time

SwotVac

decision

S A S A S A S A S A S A S
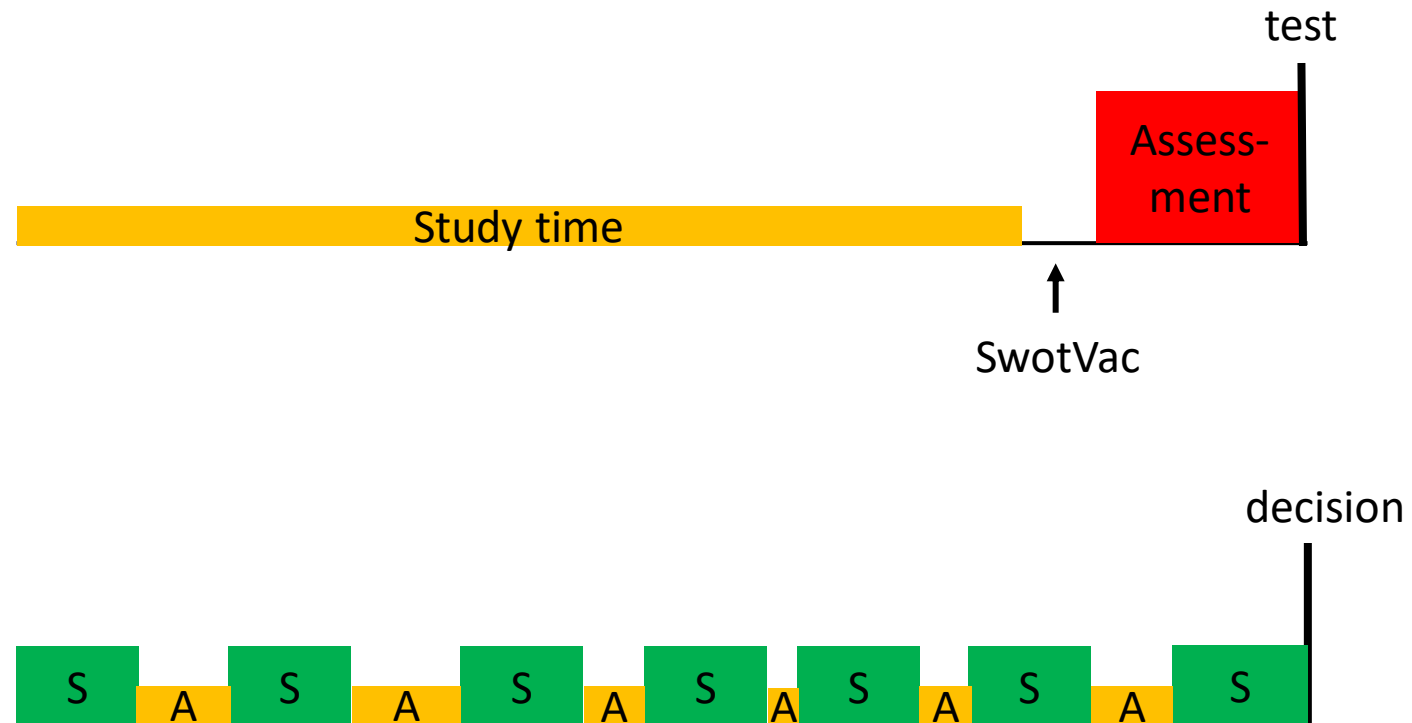
Assessment for learning: interleaved

decision

decision

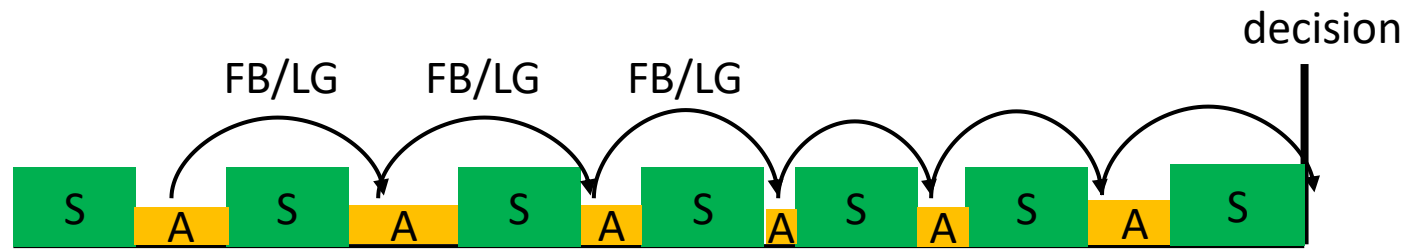# Assessment for learning: connected

Assessment *of* learning

Score 1 ← Mini-CEX 1 → *Feedback + LG 1*

Score 2 ← Mini-CEX 2 → *Feedback + LG 2*

Score 3 ← Mini-CEX 3 → *Feedback + LG 3*

Score 4 ← Mini-CEX 4 → *Feedback + LG 4*

Score 5 ← Mini-CEX 5 → *Feedback + LG 5*

Score 6 ← Mini-CEX 6 → *Feedback + LG 6*

Flinders University

Prideaux
Discipline of
Clinical Education

Assessment for learning

Score 1 ←	Mini-CEX 1 →	**Feedback + LG 1**
Score 2 ←	Mini-CEX 2 →	**Feedback + LG 2**
Score 3 ←	Mini-CEX 3 →	**Feedback + LG 3**
Score 4 ←	Mini-CEX 4 →	**Feedback + LG 4**
Score 5 ←	Mini-CEX 5 →	**Feedback + LG 5**
Score 6 ←	Mini-CEX 6 →	**Feedback + LG 6**
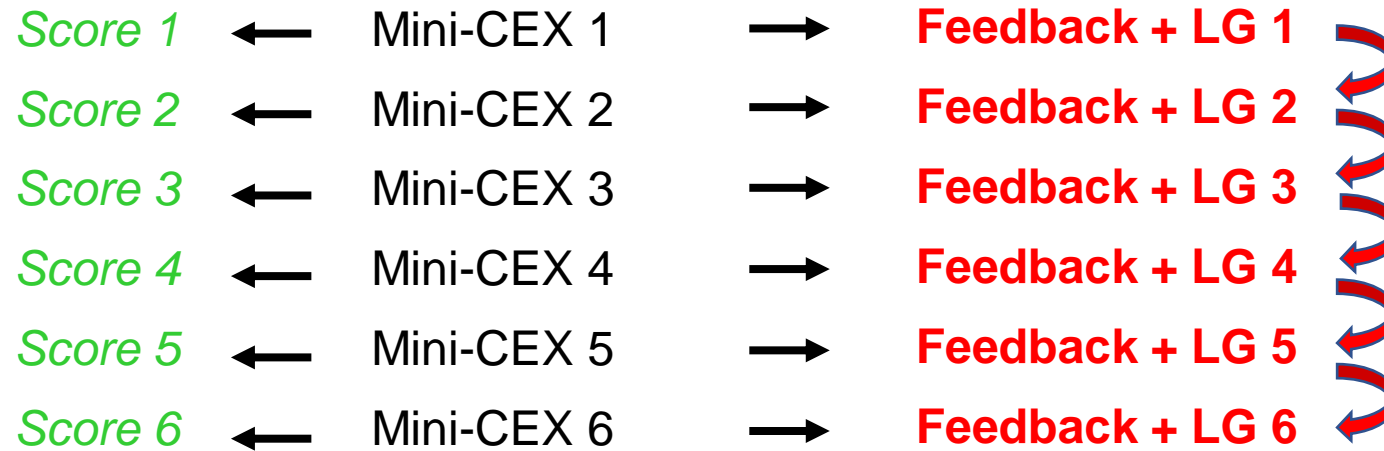
But why programmatic assessment?

6   Programmatic assessment

Programmatic assessment

- Multiple instruments with meaningful synthesis

Multiple instruments
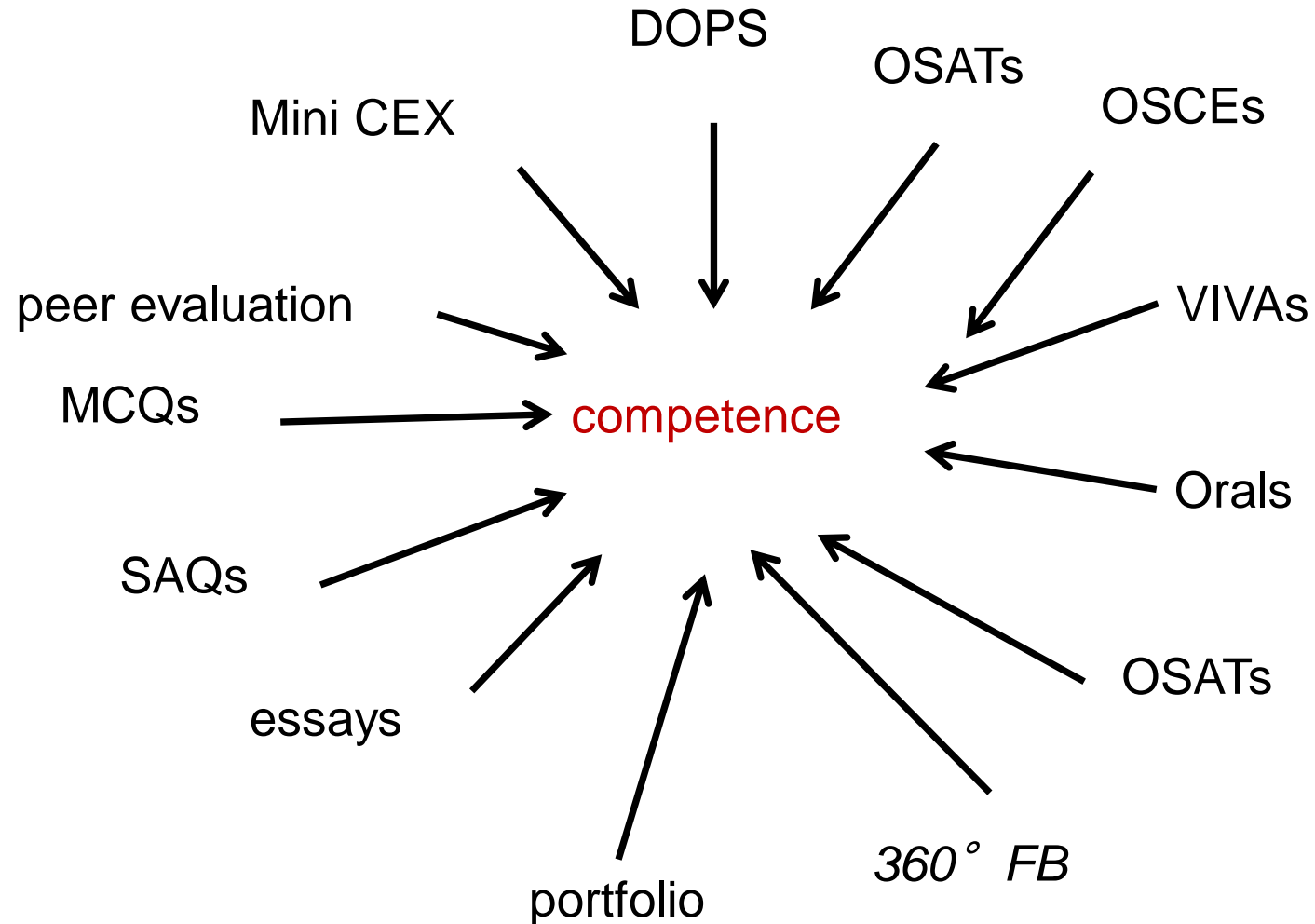
DOPS

OSATs

Mini CEX

OSCEs

peer evaluation

VIVAs

MCQs → competence

Orals

SAQs

essays

OSATs

portfolio

360° FB

Mulitple instruments

professional   communicator

Flinders University

Prideaux
Discipline of
Clinical Education
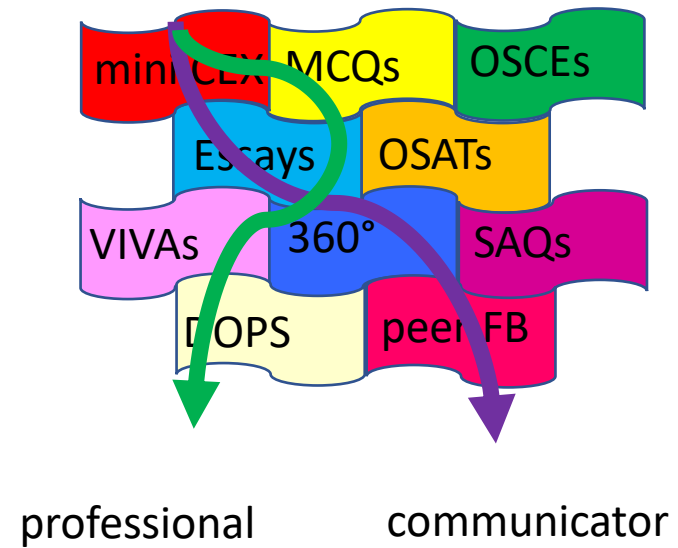
Programmatic assessment

- Multiple instruments with meaningful synthesis
  - Interleaving
  - Meaningfulness
  - complexity

Every assessment moment = decision moment

test + test + test = competent
decision · decision · decision

Prideaux
Discipline of
Clinical Education
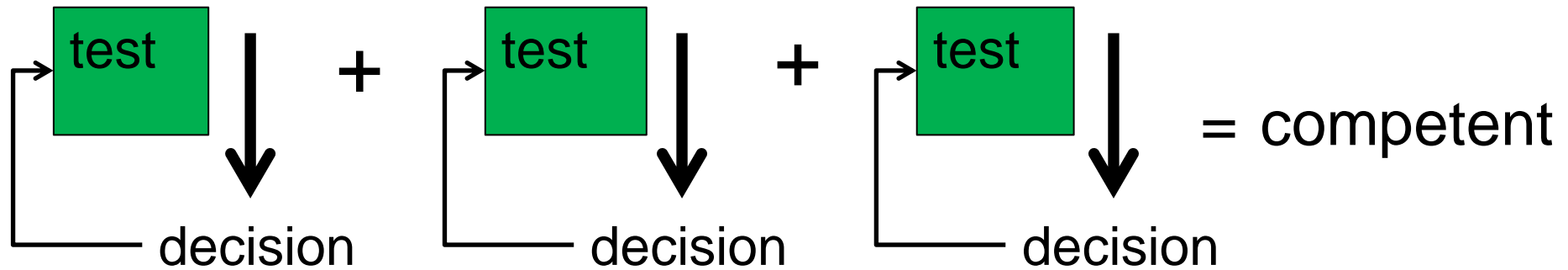
Longitudinal assessment
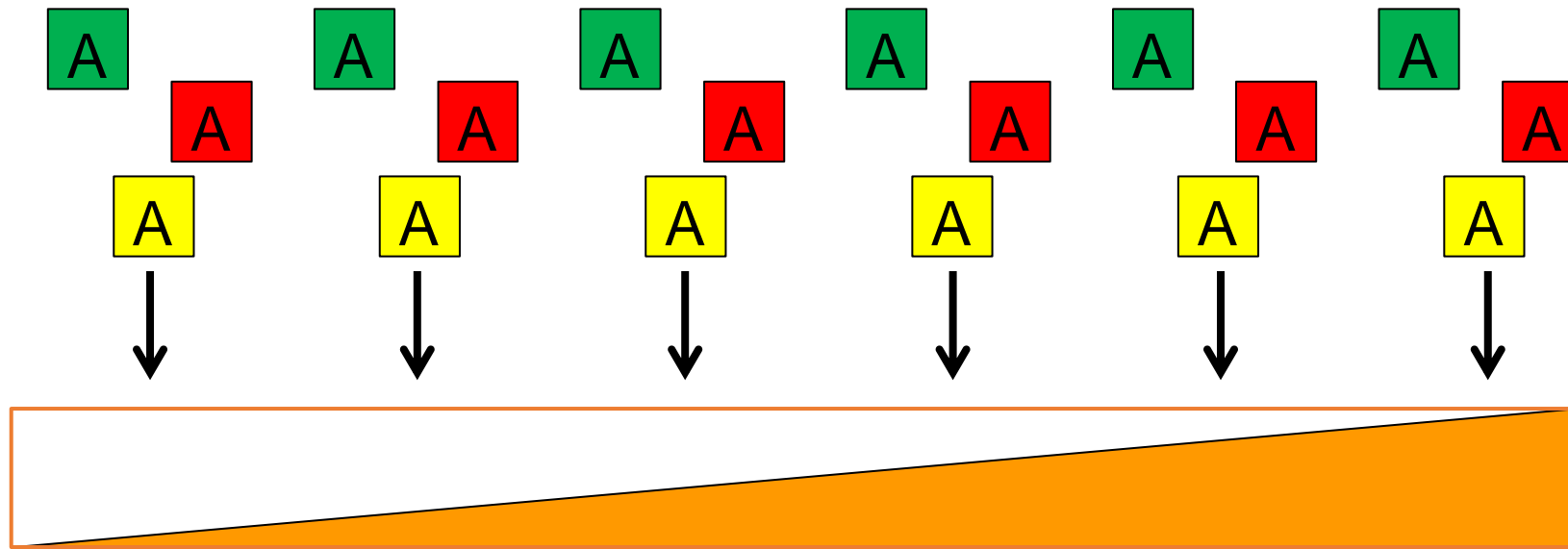
Programmatic assessment

- Continuous and longitudinal assessment
  - Distributed learning
  - Connected assessment

Proportionality

low stakes   medium stakes   high stakes

Prideaux
Discipline of
Clinical Education
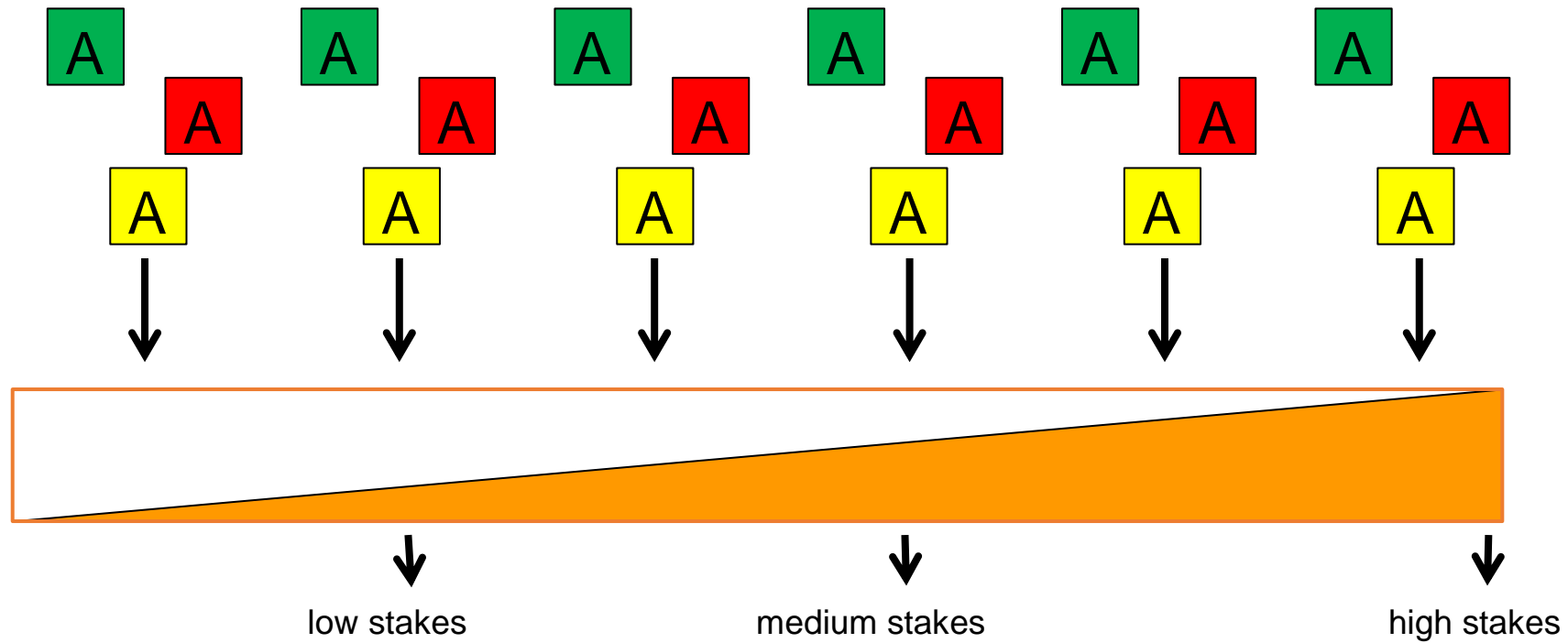
Programmatic assessment

- Proportionality of decisions
  - Goodhart's law
  - Fairness
  - Meaningfulness/Relevance
  - Prevention of leniency bias

Programmatic assessment

- Multiple instruments with meaningful synthesis
- Continuous and longitudinal assessment
- Proportionality of decisions
- Variety on quality approaches
- Assessment as a diagnostic process (instead of measurement)

Domains as lenses

professional

communicator

scholar

competence

collaborator

health advocate

manager

medical expert

Domains as lenses

socio-economic

nociceptive

pharmaco-
therapeutic

**Chronic benign
Low back pain**

neurological

orthopaedic

Occupational health

psychological

Flinders University

Prideaux
Discipline of
Clinical Education

Domains as lenses

professional

communicator

scholar

competence

collaborator

health advocate

manager

medical expert

Thank you

background reading

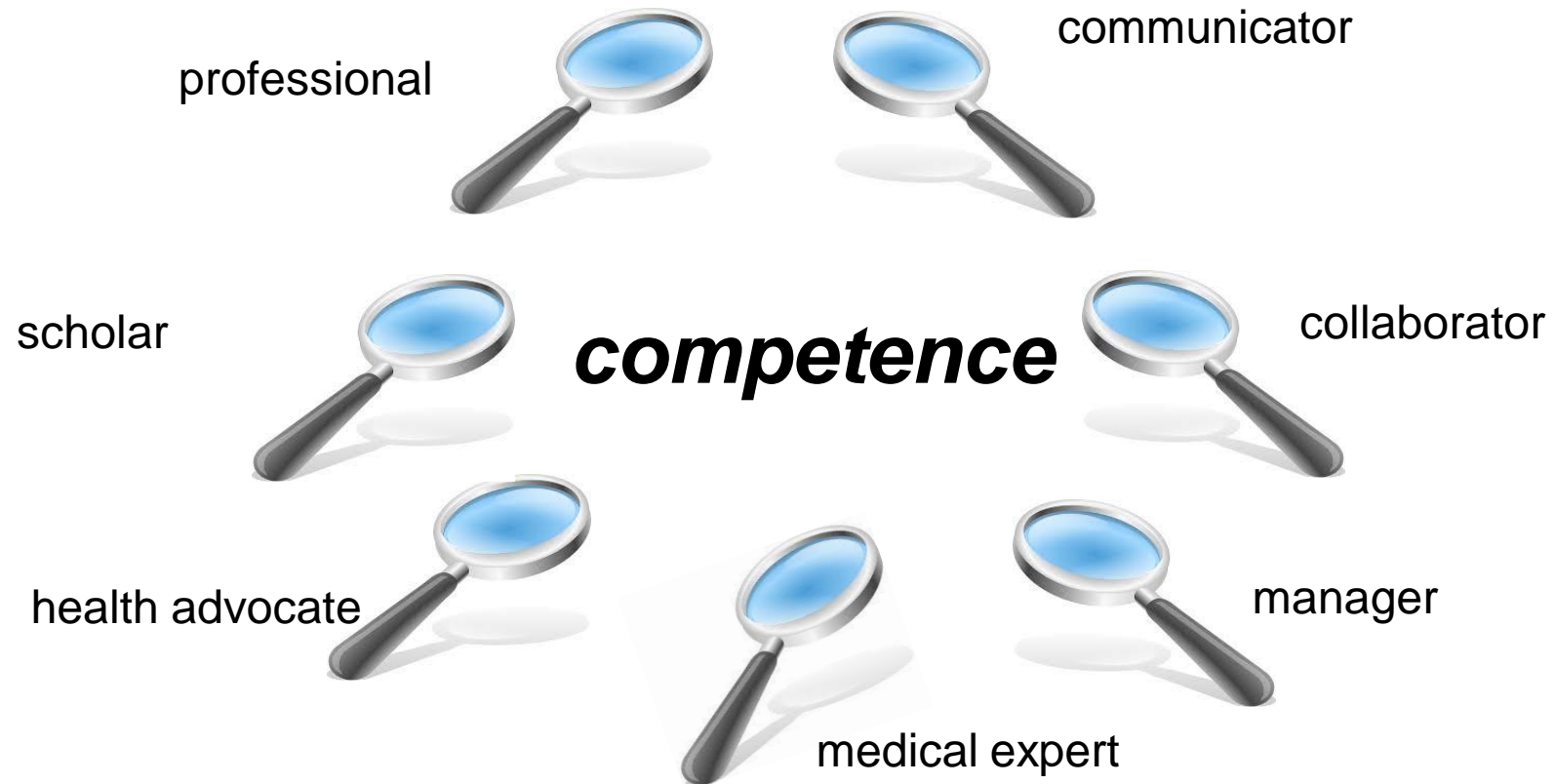1. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychological bulletin 1955;52(4):281 - 302.

2. Ebel RL. The practical validation of tests of ability. Educational Measurement: Issues and Practice 1983;2(2):7 - 10.

3. Cronbach LJ. What price simplicity? Educational Measurement: Issues and Practice 1983;2(2):11-12.

4. Kane MT. Validation. In: Brennan RL, ed. Educational Measurement. Westport: ACE/Praeger 2006:17 - 64.

5. Norman GR, Smith EKM, Powles AC, et al. Factors underlying performance on written tests of knowledge. Medical Education 1987;21:297-304.

6. Ward WC. A comparison of free-response and multiple-choice forms of verbal aptitude tests. Applied Psychological Measurement 1982;6(1):1-11.

7. Van der Vleuten CPM, Van Luyk SJ, Beckers HJM. A written test as an alternative to performance testing. Medical Education 1988;22:97-107.

8. Norman G, Swanson D, Case S. Conceptual and methodology issues in studies comparing assessment formats, issues in comparing item formats. Teaching and Learning in Medicine 1996;8(4):208-16.

9. Norman GR, Van der Vleuten CP, De Graaff E. Pitfalls in the pursuit of objectivity: issues of validity, efficiency and acceptability. Medical Education 1991;25(2):119-26.

10. Boud D. Assessment and the promotion of academic values. Studies in Higher Education 1990;15(1):101-11.

11. Govaerts MJB, Van der Vleuten CPM, Schuwirth LWT, et al. Broadening Perspectives on Clinical Performance Assessment: Rethinking the Nature of In-training Assessment. Advances in health sciences education 2007;12(2):239-60.

12. Canmeds. http://www.royalcollege.ca/portal/page/portal/rc/canmeds Ottawa2005 [accessed 26 August 2018.

13. ACGME. http://www.acgme.org/What-We-Do/Accreditation/Milestones/Overview Chicago2007 [accessed August 2018].

14. Albanese MA, Mejicano G, Mullan P, et al. Defining characteristics of educational competencies. Medical Education 2008;42(3):248-55.

Flinders University

Prideaux
Discipline of
Clinical Education

background reading

15. Van der Vleuten CPM. The assessment of Professional Competence: Developments, Research and Practical Implications. Advances in Health Science Education 1996;1(1):41-67.

16. Kane M. Current concerns in validity theory. Journal of Educational Measurement 2001;38(4):319-42.

17. Schuwirth LWT, Van der Vleuten CPM. A plea for new psychometrical models in educational assessment. Medical Education 2006;40(4):296-300.

18. Driessen E, Van der Vleuten CPM, Schuwirth LWT, et al. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. Medical Education 2005;39(2):214-20.

19. Van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. Medical Education 2005;39(3):309-17.

20. Dannefer E, Henson L. The portfolio approach to competency-based assessment at the Cleveland Clinic Lerner College of Medicine. Academic Medicine 2007;82(5):493-502.

21. Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. Medical Teacher 2012;34:205-14. doi: 10.3109/0142159X.2012.652239

22. Schuwirth LWT, Van der Vleuten CPM, Durning SJ. What programmatic assessment for learning in medical education can learn from healthcare. Perspectives on Medical Education 2017:1-5. doi: 10.1007/s40037-017-0345-1

23. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment and Kane's validity perspective. Medical Education 2012;46(1):38-48. doi: 10.1111/j.1365-2923.2011.04098.x

24. Schuwirth LWT, Van der Vleuten CPM. Assessing Competence: Extending the Approaches to Reliability. In: Hodges BD, Lingard L, eds. The Question of Competence. Ithaka New York US: Cornell University Press 2012.

25. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: from assessment of learning to assessment for learning. Medical Teacher 2011;33(6):478-85.

Flinders University

Prideaux
Discipline of
Clinical Education

background reading

26. Delandshere G, Petrosky AR. Assessment of Complex Performances: Limitations of Key Measurement Assumptions. Educational Researcher 1998;27(2):14-24.

27. Shute V. Focus on formative feedback. Review of educational research 2008;78(n):153-89.

28. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Medical Education* 2007;41:1124-30. doi: 10.1111/j.1365-2923.2007.02946.x

29. Cook DA, Kuper A, Hatala R, et al. When assessment data are words: validity evidence for qualitative educational assessments. Academic Medicine 2016;91(10):1359-69.

30. Popham WJ. Assessment Literacy for Teachers: Faddish or Fundamental? Theory Into Practice 2009;48:4-11.

31. Watling C, Driessen E, Van der Vleuten CPM, et al. Beyond individualism: professional culture and its influence on feedback. Medical Education 2013;47(6):585-94.

32. Harrison CJ, Könings KD, Schuwirth L, et al. Barriers to the uptake and use of feedback in the context of summative assessment. Advances in health sciences education 2015;20(1):229–45. doi: 10.1007/s10459-014-9524-6

33. Cilliers FJ, Schuwirth LWT, Adendorff HJ, et al. The mechanisms of impact of summative assessment on medical students' learning. Advances in health sciences education 2010;15:695-715. doi: 10.1007/s10459-010-9232-9

34. Valentine N, Durning S, Shanahan EM, et al. Fairness in human judgement in assessment: A hermeneutic literature review and conceptual framework. Advances in Health Sciences Education 2021;26(2):713-38.

35. Valentine N, Shanahan EM, Durning SJ, et al. Making it fair: Learners' and assessors' perspectives of the attributes of fair judgement. Medical Education 2021;55(9):1056-66. doi: https://doi.org/10.1111/medu.14574

36. Valentine N, Durning SJ, Shanahan EM, et al. The pursuit of fairness in assessment: Looking beyond the objective. Medical Teacher 2022:1-7.

Flinders University

Prideaux
Discipline of
Clinical Education