S Y M P O S I U M

Assessment: A Constructive Companion in Higher Education – Embracing Programmatic Assessment for learning

Sylvia Heeneman[#]
Walther van Mook[#^]
Lambert Schuwirth[*]

[#]Maastricht University, [^]Maastricht University Medical Centre, the Netherlands
**Flinders University, Adelaide, Australia*

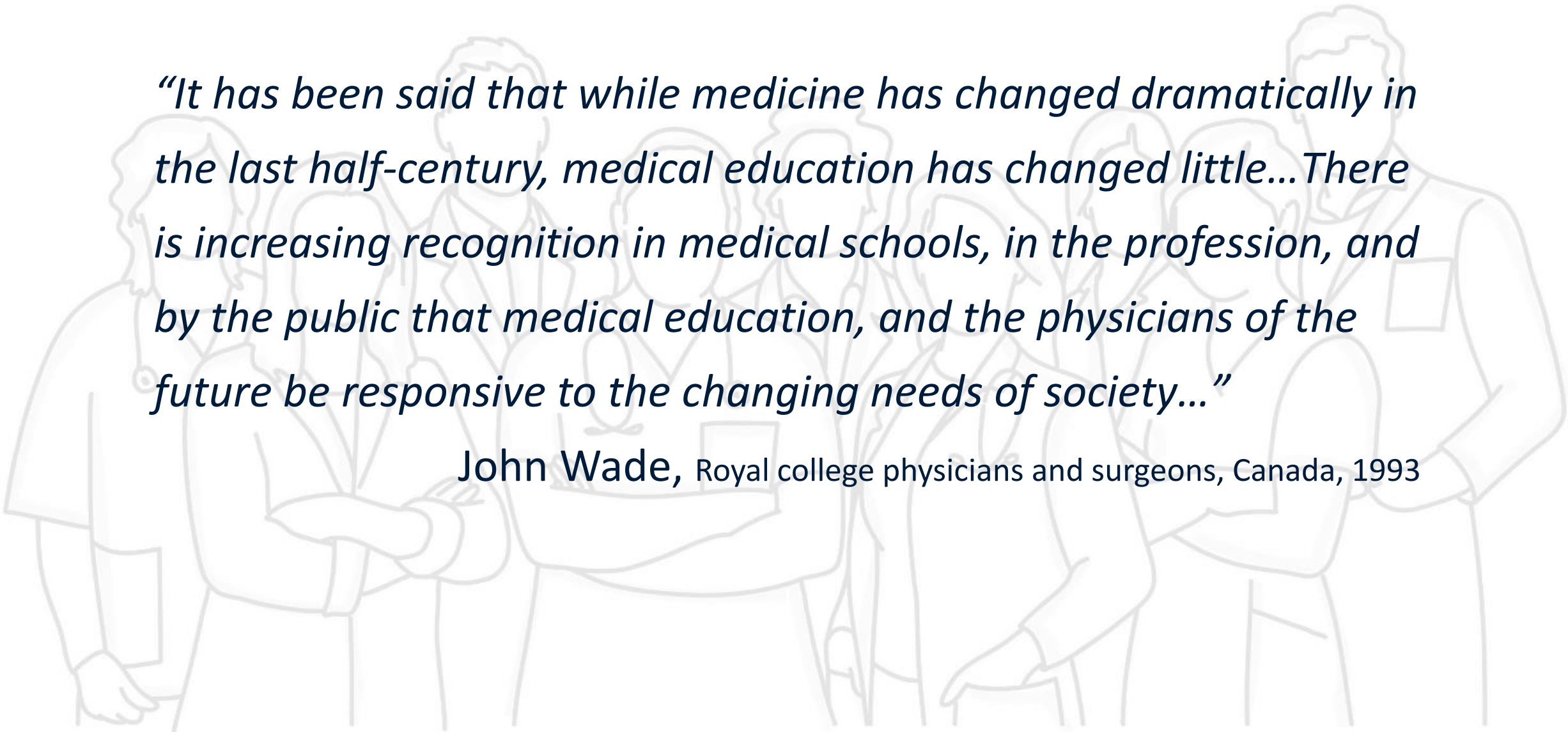
Feb 7th, 2024

HEALTH CARE PROFESSIONS

HEALTH CARE PROFESSIONS
COMPETENCY–BASED EDUCATION

"It has been said that while medicine has changed dramatically in the last half-century, medical education has changed little…There is increasing recognition in medical schools, in the profession, and by the public that medical education, and the physicians of the future be responsive to the changing needs of society…"

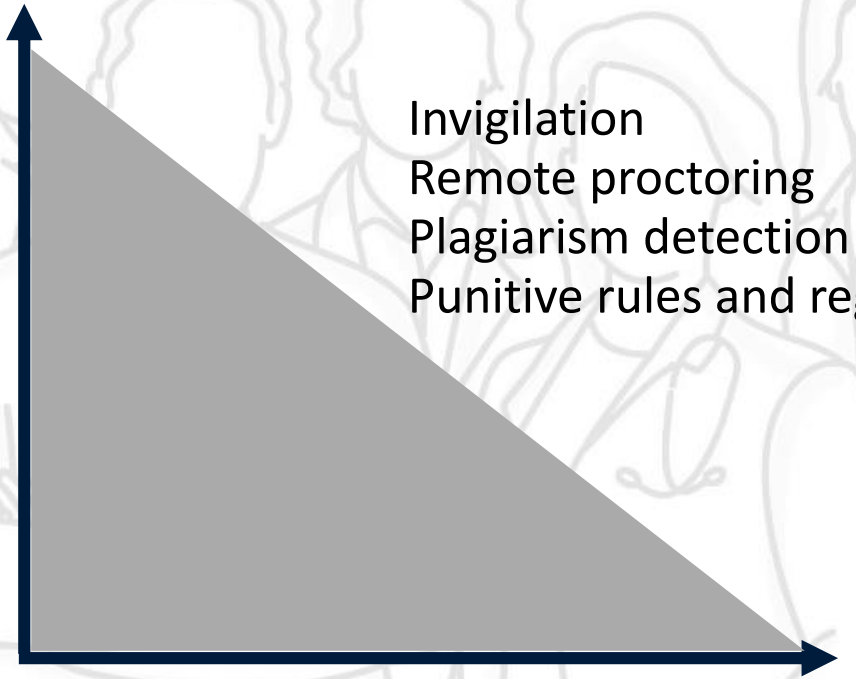
John Wade, Royal college physicians and surgeons, Canada, 1993



HEALTH CARE PROFESSIONS
COMPETENCY–BASED ASSESSMENT & EDUCATION

Learning
Competence
Collaboration
Managing uncertainty

education



assessment

Invigilation
Remote proctoring
Plagiarism detection
Punitive rules and regulations

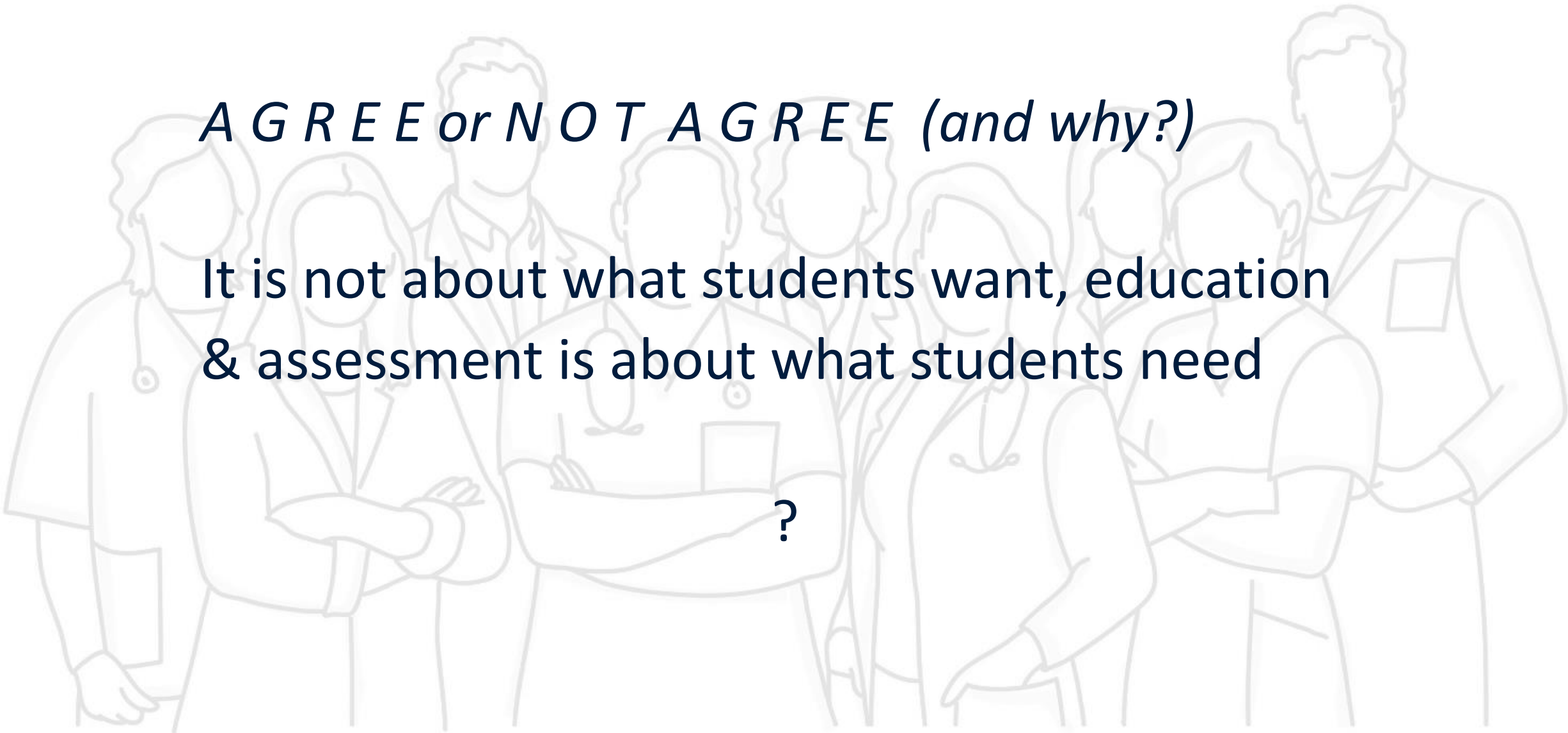
Individual competence
Discriminating
Ranking
Grading

C O M P E T E N C Y – B A S E D A S S E S S M E N T

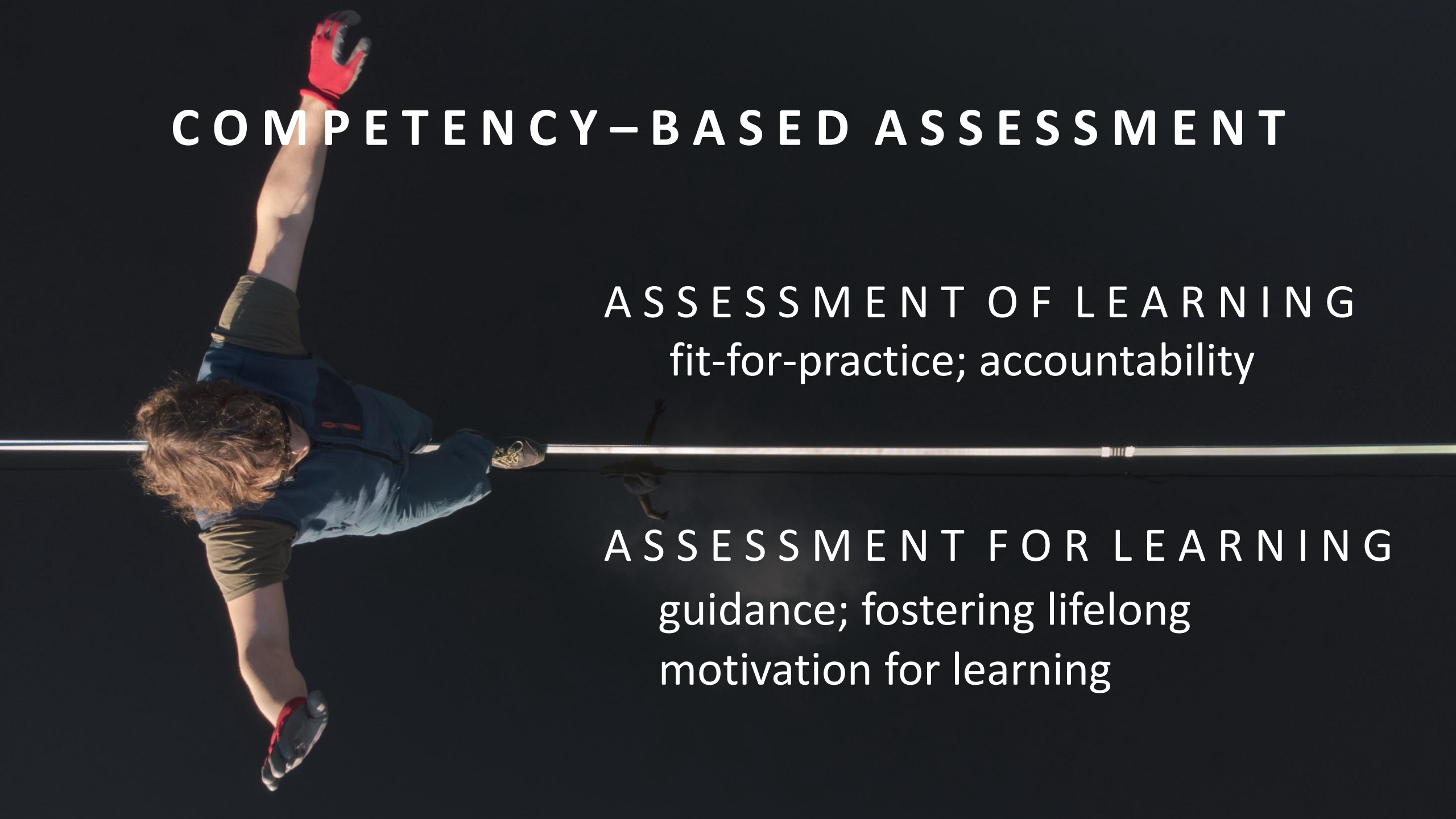
A G R E E or N O T A G R E E (and why?)

It is not about what students want, education & assessment is about what students need

?



COMPETENCY–BASED ASSESSMENT



ASSESSMENT OF LEARNING

fit-for-practice; accountability

ASSESSMENT FOR LEARNING

guidance; fostering lifelong motivation for learning

COMPETENCY–BASED ASSESSMENT

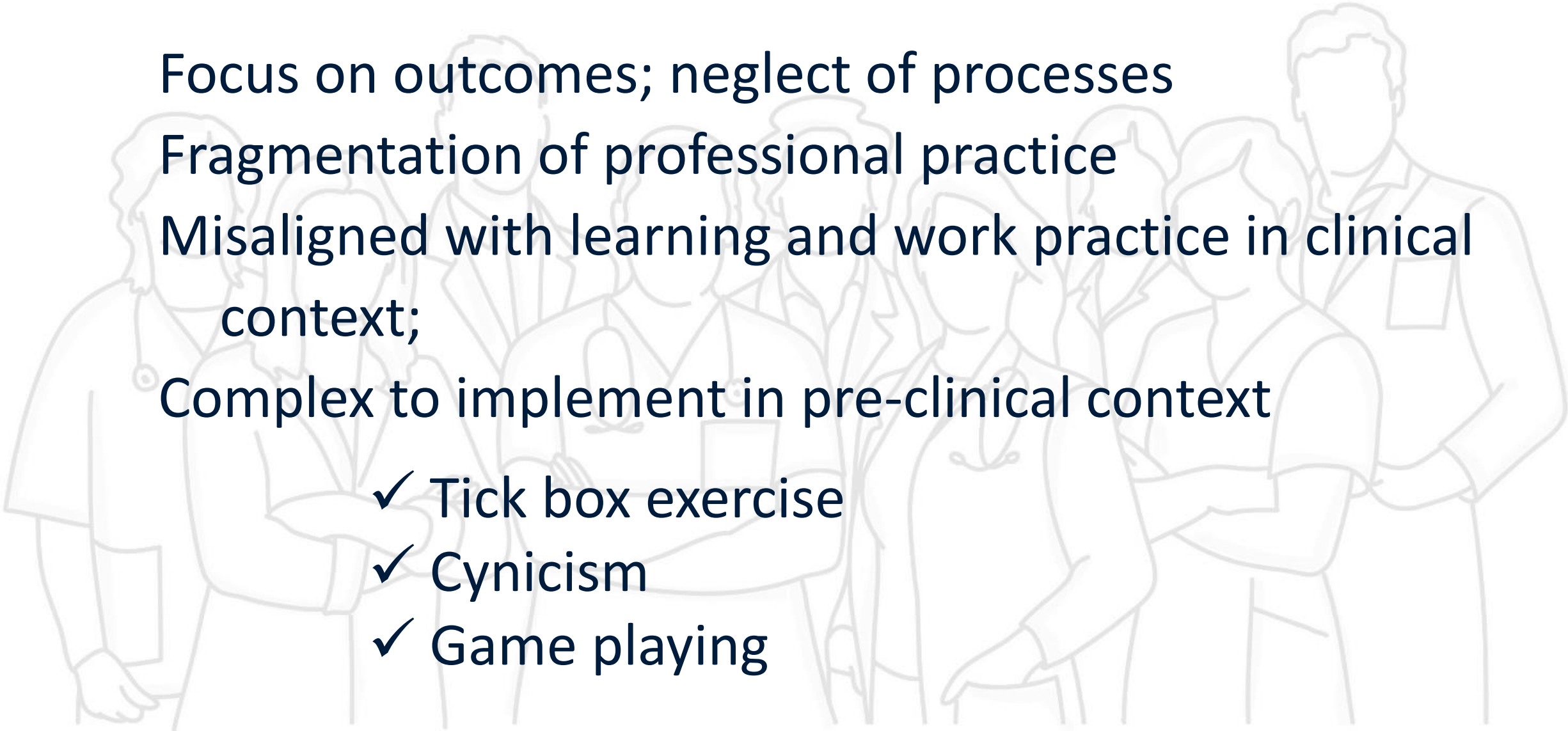
Focus on outcomes; neglect of processes

Fragmentation of professional practice

Misaligned with learning and work practice in clinical context;

Complex to implement in pre-clinical context

- ✓ Tick box exercise
- ✓ Cynicism
- ✓ Game playing



COMPETENCY–BASED ASSESSMENT

FIT FOR PURPOSE ASSESSMENT THEORY ?!

FORMATIVE ASSESSMENT

ASSESSMENT FOR LEARNING

PROGRAM–MATIC ASSESSMENT

ASSESSMENT FOR LEARNING

FORMATIVE ASSESSMENT

F --> S

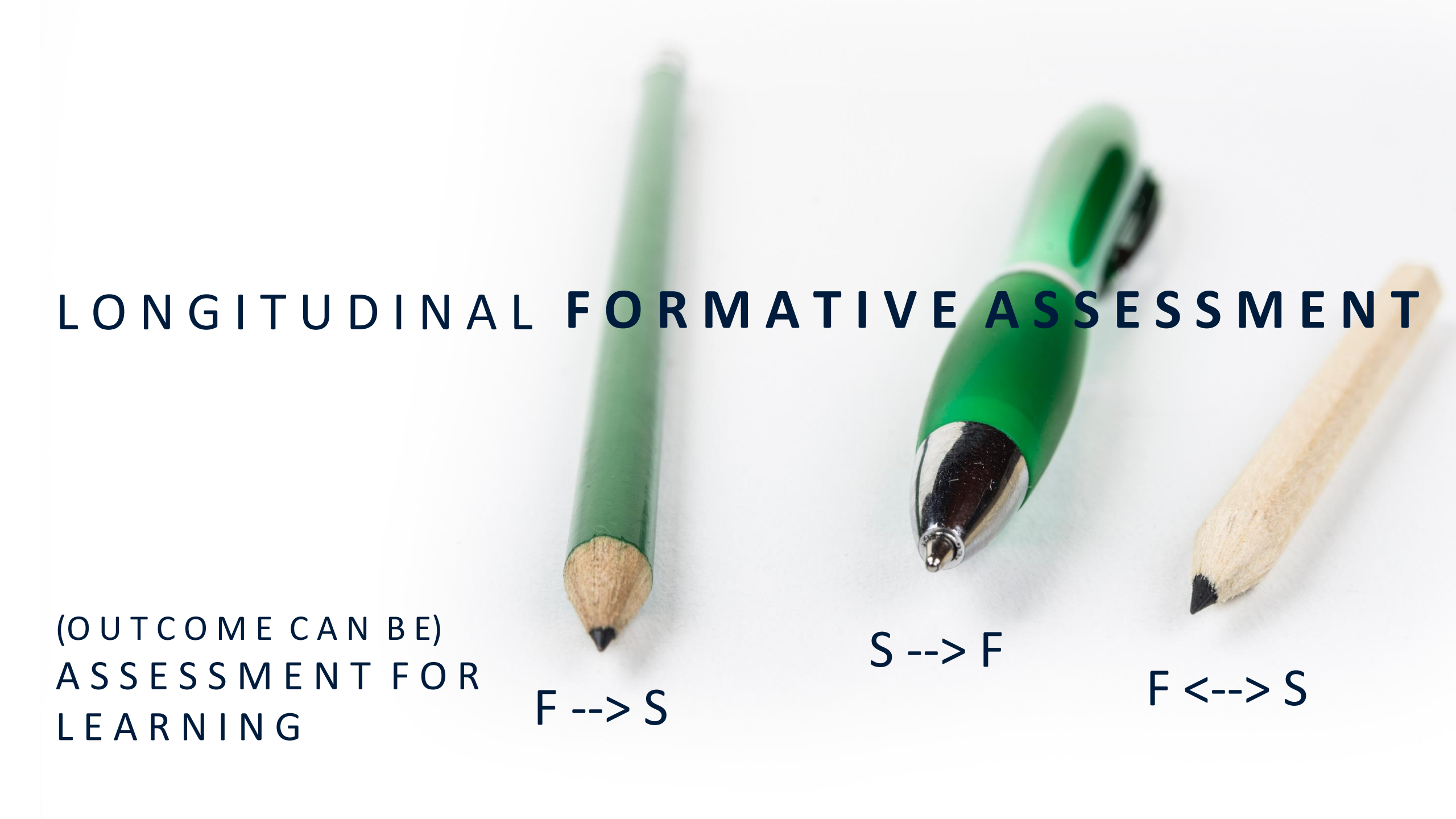
F <-- S

F <--> S

SUMMATIVE ASSESSMENT

Linda Allal, 2011, 2019

LONGITUDINAL **FORMATIVE ASSESSMENT**



(OUTCOME CAN BE)
ASSESSMENT FOR
LEARNING

F --> S

S --> F

F <--> S

FORMATIVE ASSESSMENT

'SINGLE’ USE

DIFFERENT OUTCOME

ASSESSMENT FOR LEARNING
(possible) PROCESSES

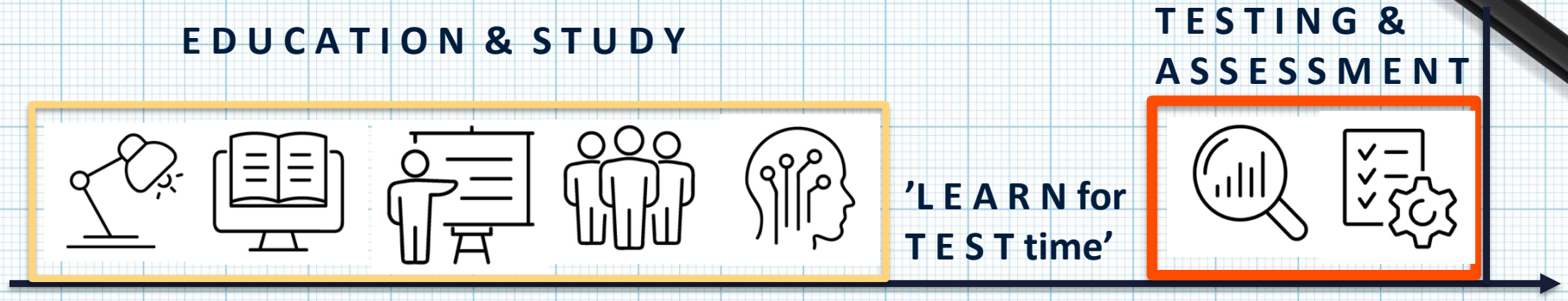
DISTRIBUTED

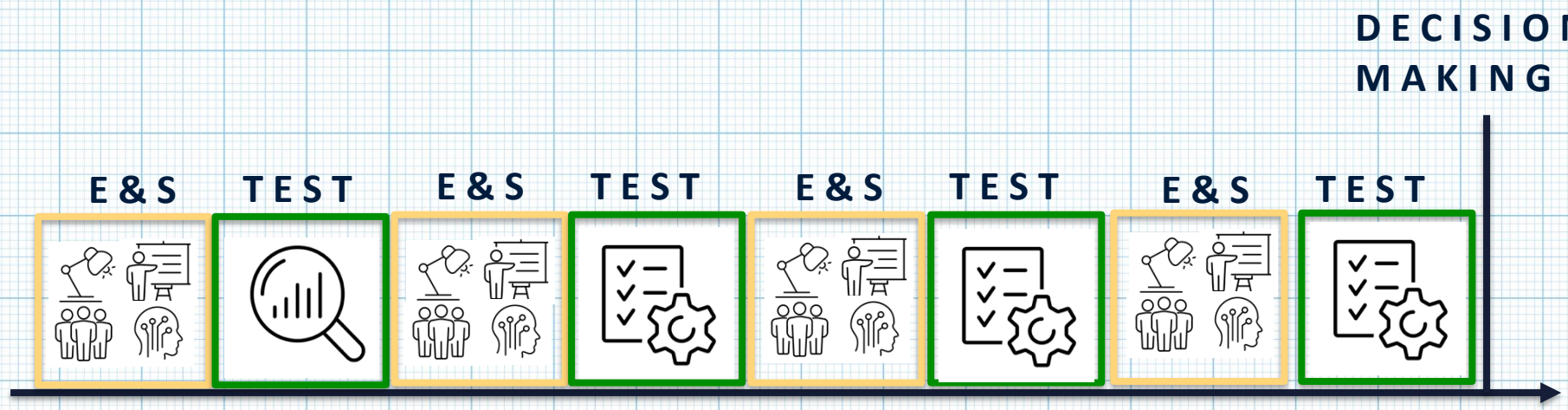
INTERLEAVED

CONNECTED

ASSESSMENT FOR LEARNING

DISTRIBUTED

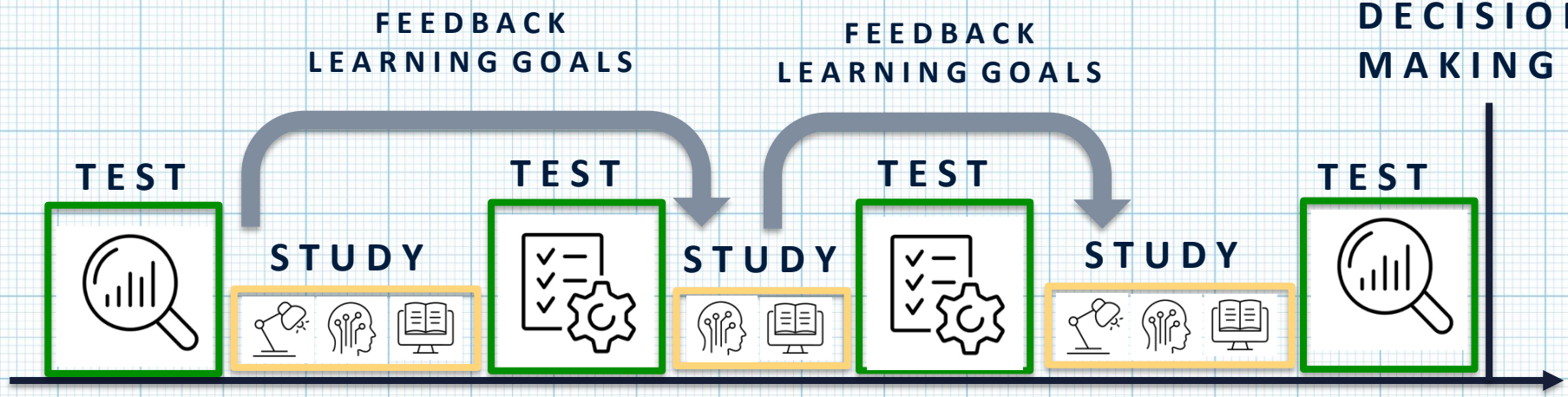




Similarity F --> S

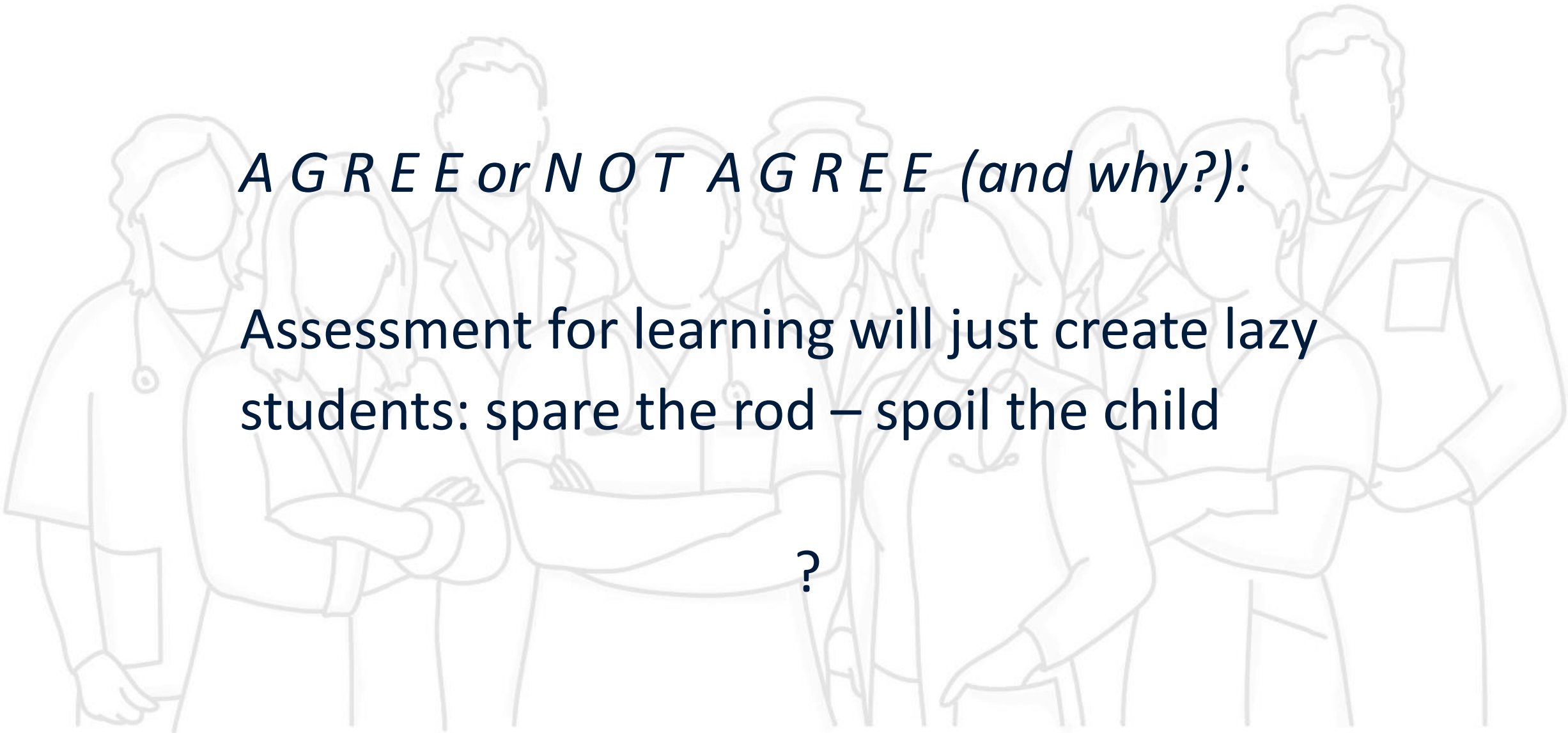
ASSESSMENT FOR LEARNING

DISTRIBUTED



Similarity F <--> S

A S S E S S M E N T F O R L E A R N I N G



A G R E E or N O T A G R E E (and why?):

Assessment for learning will just create lazy students: spare the rod – spoil the child

?

PROGRAMMATIC ASSESSMENT

IMPORTANT STEPPING STONES

OPTIMIZE ASSESSMENT AS PROGRAM (UTILITY)

DATAPOINTS

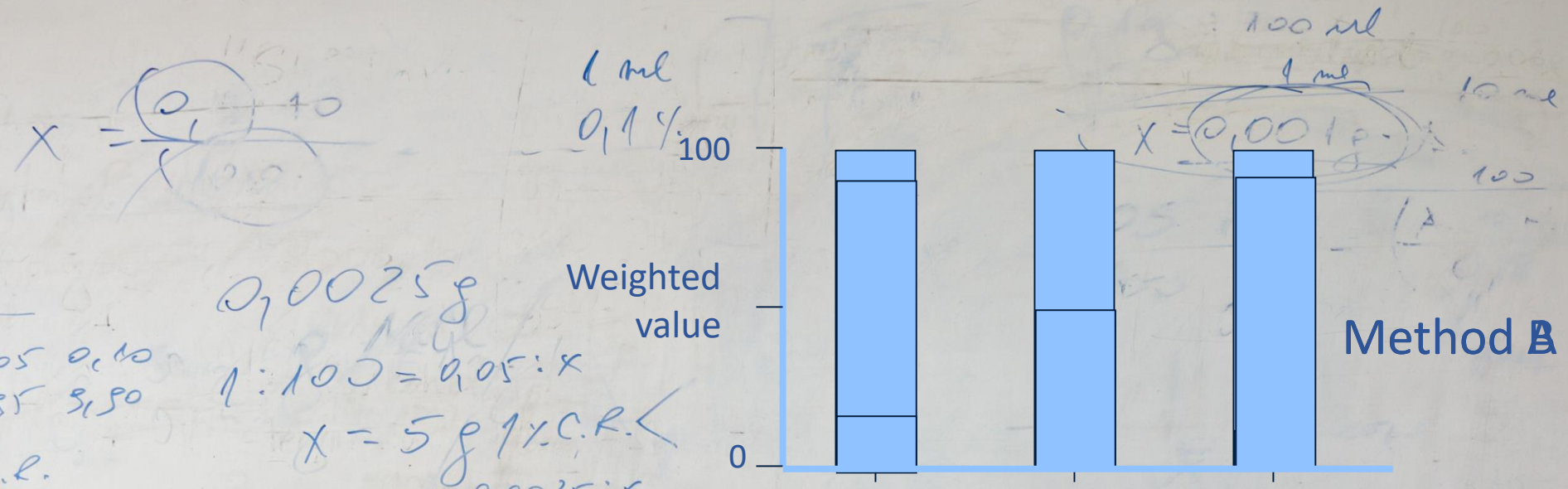
NARRATIVE FEEDBACK

The Assessment of Professional Competence: Developments, Research and Practical Implications

C.P.M. VAN DER VLEUTEN
P.O. Box 616, 6200 ML Maastricht, The Netherlands

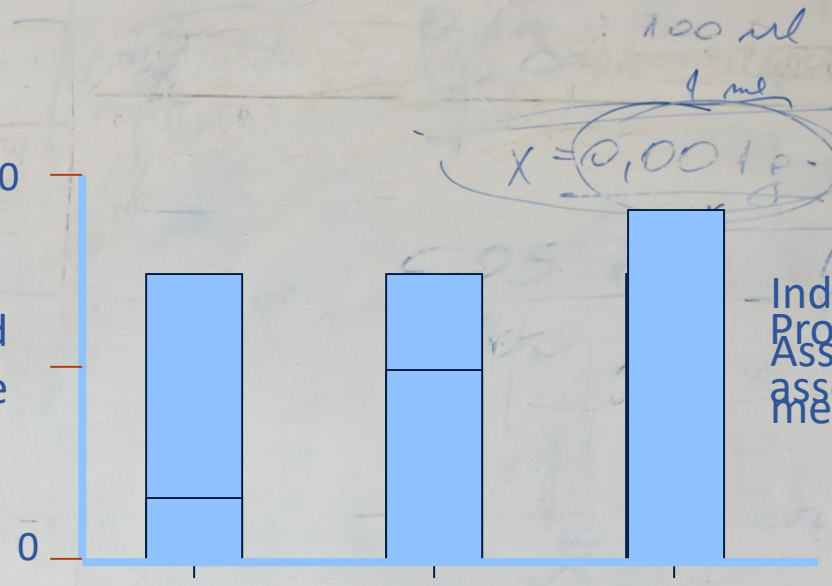
Introduction

Educational achievement testing is an area of turmoil in the health sciences. Examinations are a constant source of problems for many teachers, curriculum designers and educationalists. The evaluation of student achievement is continuously debated at educational meetings, conferences and workshops. It is an area in which tradi-



$$U = {}_wR \times {}_wV \times {}_wEI$$

- U = 'Utility'
- w = weigthed
- R = Reliability
- V = Validity
- EI = Educational Impact



$$U = {}_wR \times {}_wV \times {}_wEI$$

- U = 'Utility'
- w = weigthed
- R = Reliability
- V = Validity
- EI = Educational Impact

assessment

Assessing professional competence: from methods to programmes

Cees P M van der Vleuten & Lambert W T Schuwirth

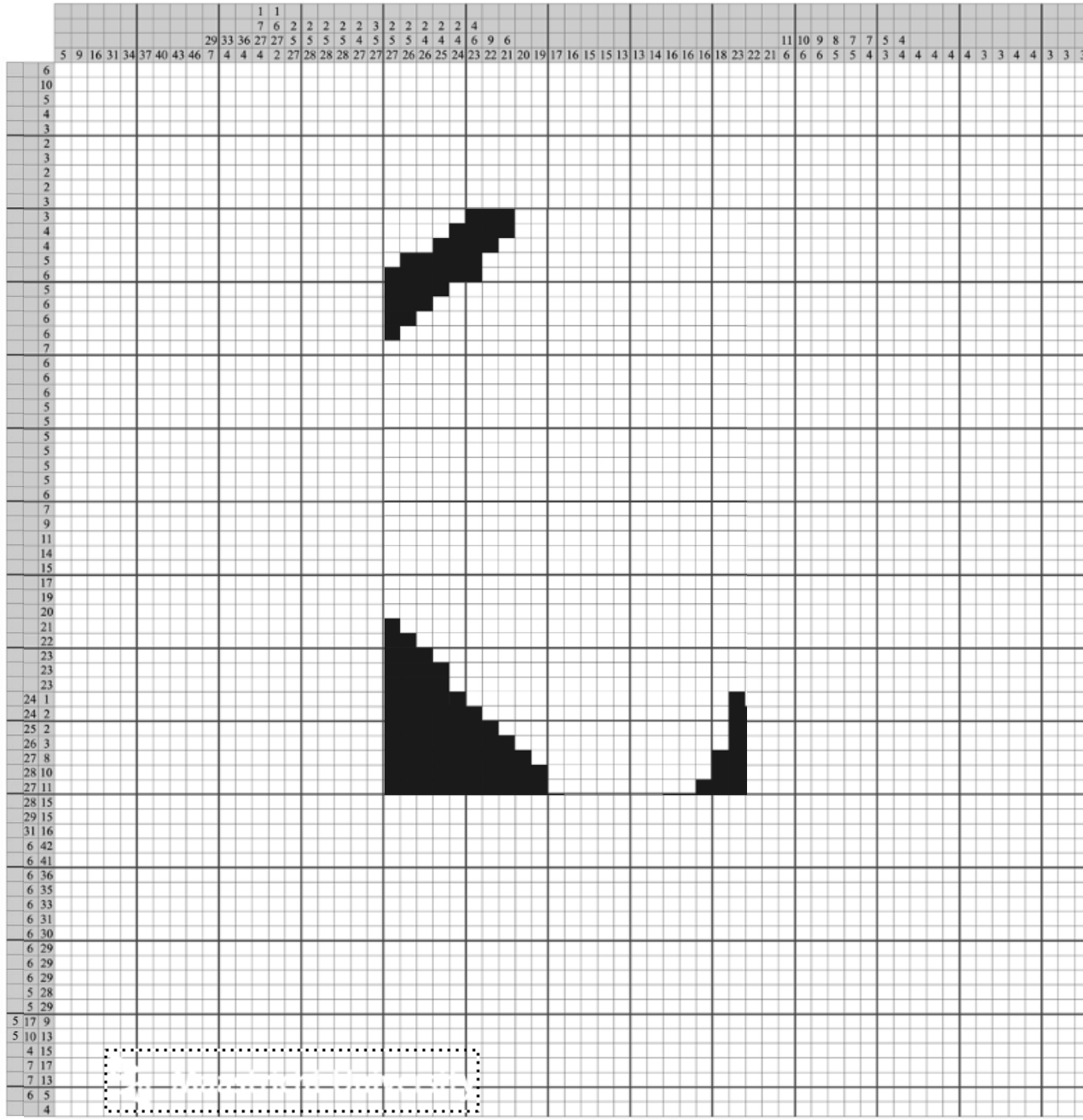
INTRODUCTION We use a utility model to illustrate that, firstly, selecting an assessment method involves context-dependent compromises, and secondly, that assessment is not a measurement problem but an instructional design problem, comprising educational, implementation and resource aspects. In the model, assessment characteristics are differently

Therefore, we need an instructional design perspective.

IMPLICATIONS FOR DEVELOPMENT AND RESEARCH Programmatic instructional design hinges on a careful description and motivation of choices, whose effectiveness should be measured against

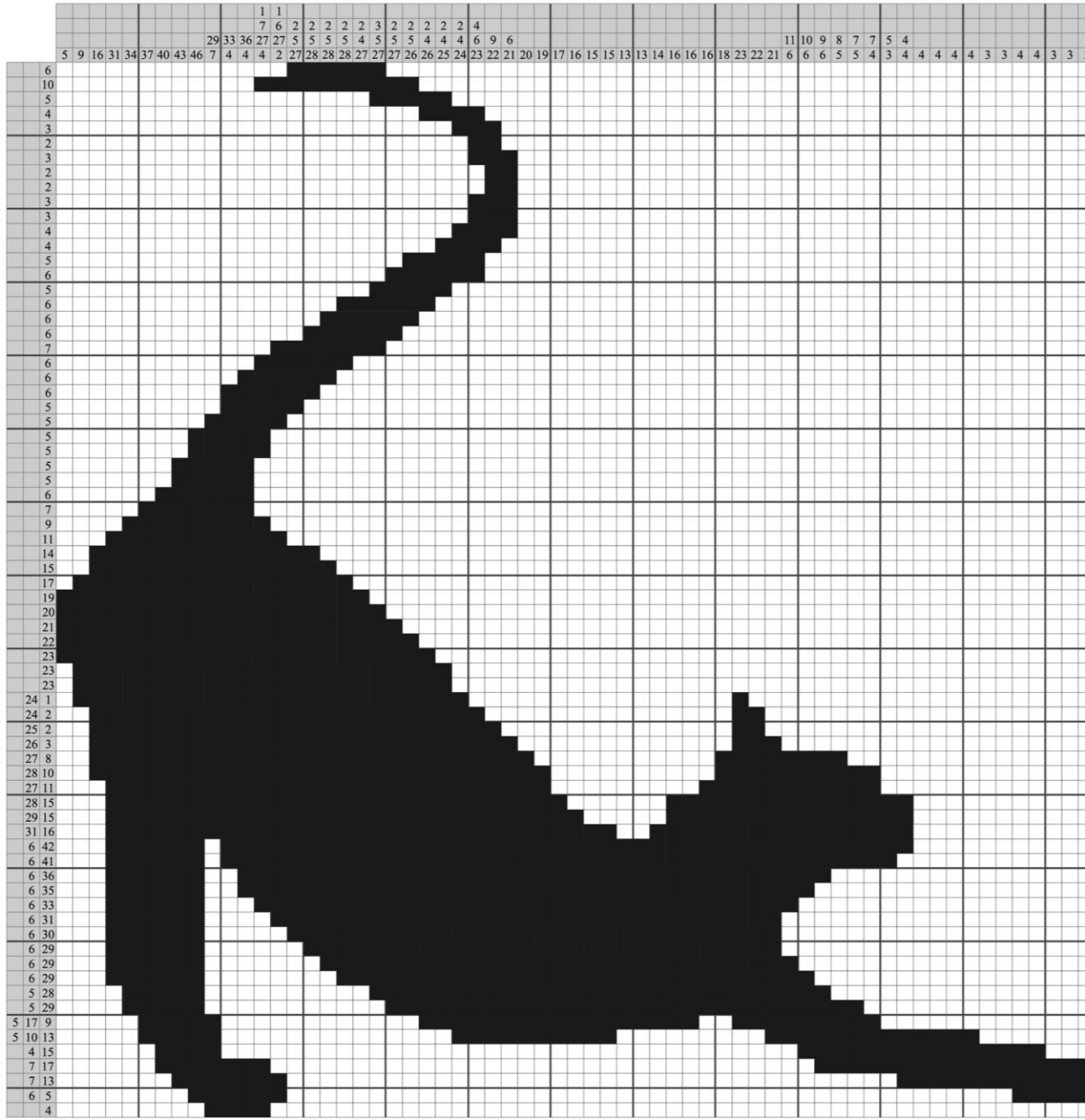
Reliability, as a function of the sample size

Testing time	Multiple choice Q.	Oral exam	Mini-CEX
1 hour	0.62	0.60	0.54
2 hours	0.76	0.75	0.69
4 hours	0.93	0.86	0.82
8 hours	0.93	0.90	0.90



Reliability, as a function of the sample size

Testing time	Multiple choice Q.	Oral exam	Mini-CEX
1 hour	0.62	0.60	0.54
2 hours	0.76	0.75	0.69
4 hours	0.93	0.86	0.82
8 hours	0.93	0.90	0.90



2011; 33: 478–485 MEDICAL TEACHER

Programmatic assessment: From assessment of learning to assessment for learning

LAMBERT W. T. SCHUWIRTH & CEES P. M. VAN DER VLEUTEN
Maastricht University, The Netherlands

Abstract

In assessment a considerable shift in thinking has occurred from assessment of learning to assessment for learning. This has important implications for the conceptual framework from which to approach the issue of assessment, but also with respect to the research agenda. The main conceptual changes pertain to programmes of assessment. This has led to a broadened perspective on the types of construct assessment tries to capture, the way information from various sources is collected and collated, the role of human judgement and the variety of psychometric methods to determine the quality of the assessment. Research into the quality of assessment programmes, how assessment influences learning and teaching, new psychometric models and the role of human judgement is much needed.

Assessment leads to information Datapoints

C O N T I N U U M of stakes

Datapoints

High stake

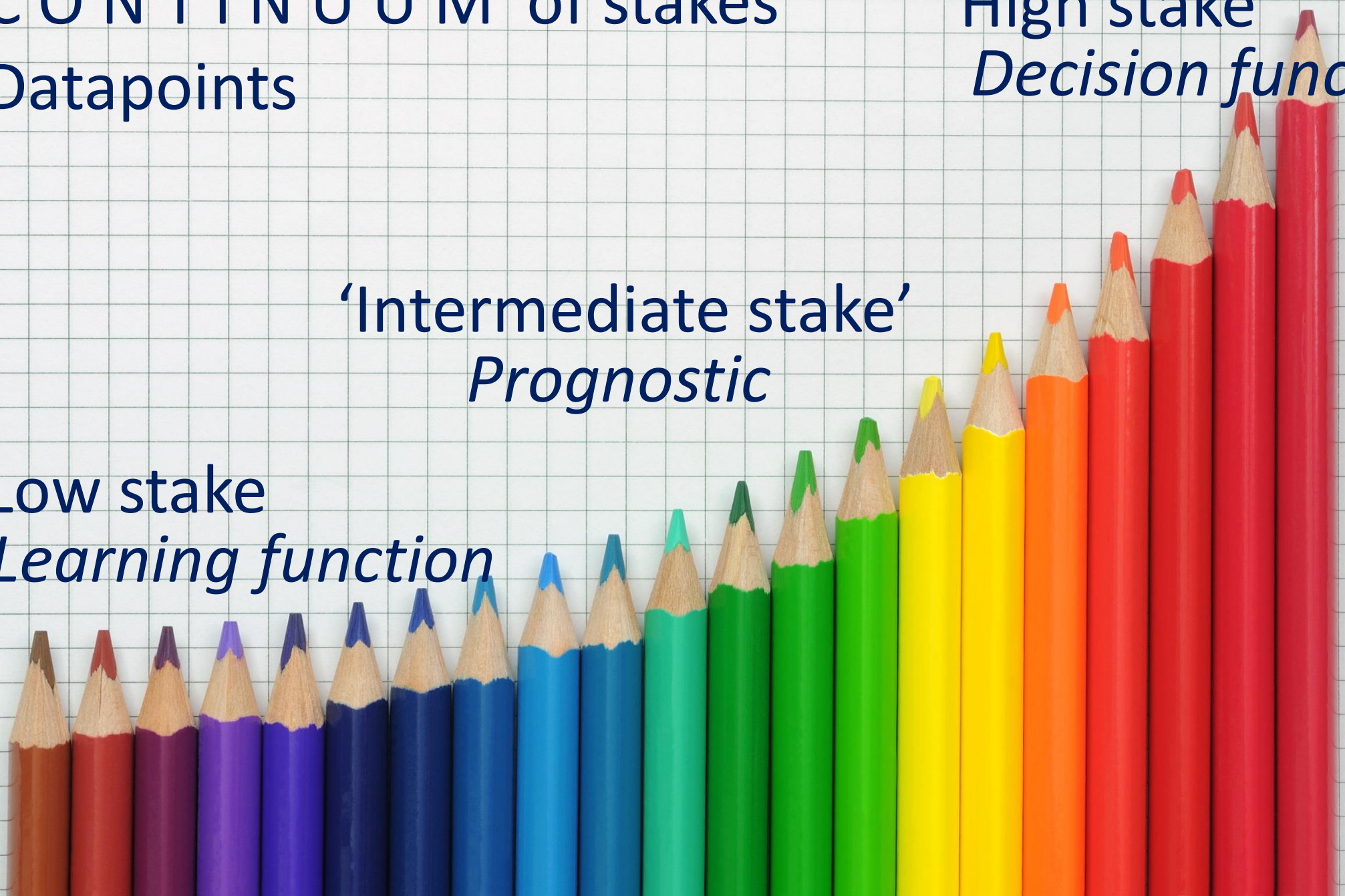
Decision function

‘Intermediate stake’

Prognostic

Low stake

Learning function



DATAPOINTS

Mix/ diversity of formats

1 Datapoint: information on several competencies

Datapoint serves the learning and decision-function

Aggregation

- For high-stake decision making
- For learning

‘N A R R A T I V E’

Feedback dialogue
Follow-up of feedback

Guidance of students
- Self-regulation of learning
- needs ‘regulation’

Informative for
- Learning function
- Decision-making

2015, 37: 641–646 MEDICAL TEACHER

TWELVE TIPS

Twelve Tips for programmatic assessment

C.P.M. VAN DER VLEUTEN[1], L.W.T. SCHUWIRTH[2], E.W. DRIESSEN[1], M.J.B. GOVAERTS[1] & S. HEENEMAN[1]

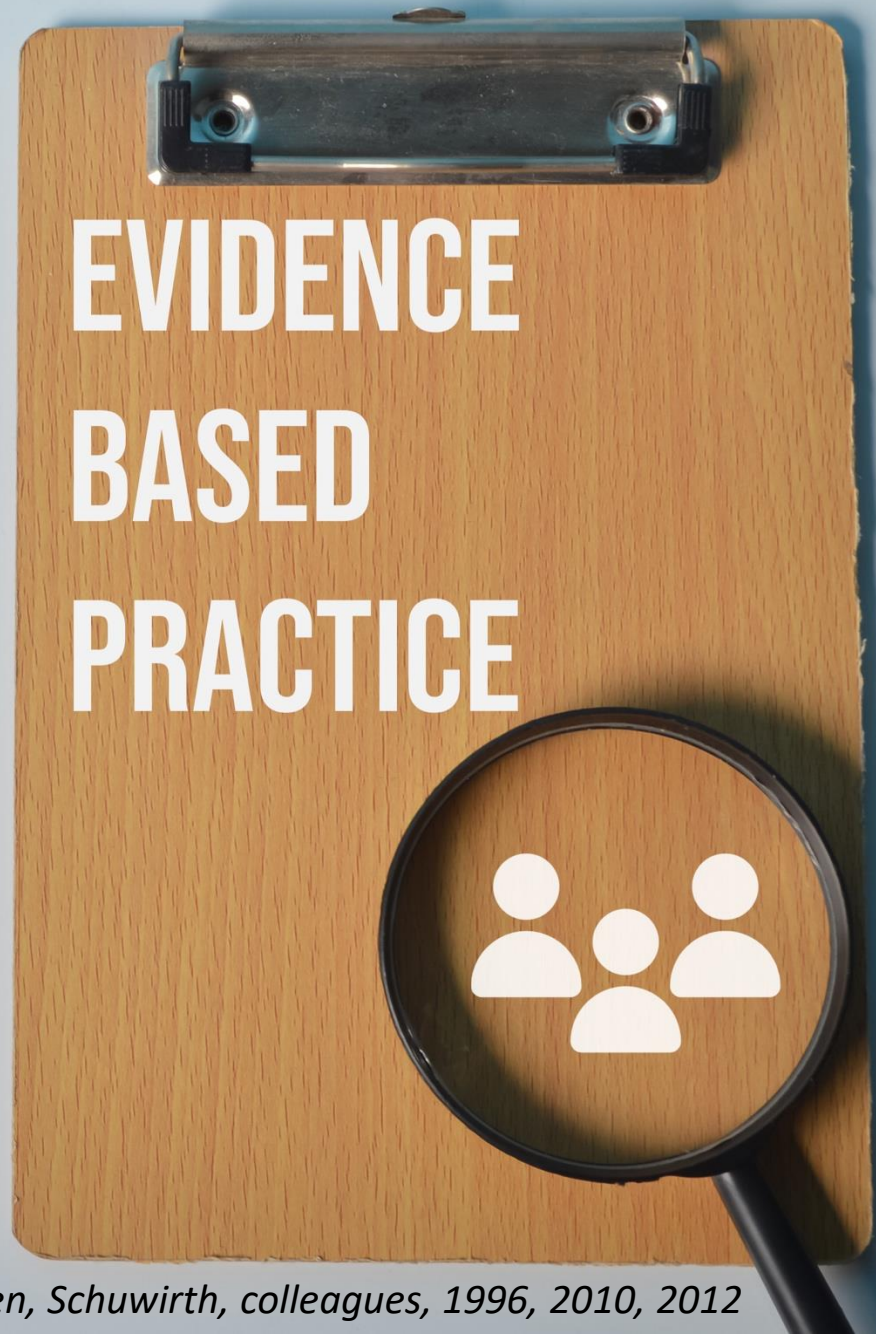
[1]Maastricht University, Maastricht, The Netherlands, [2]Flinders University, Adelaide, Australia

Abstract

Programmatic assessment is an integral approach to the design of an assessment program with the intent to optimise its learning function, its decision-making function and its curriculum quality-assurance function. Individual methods of assessment, purposefully chosen for their alignment with the curriculum outcomes and their information value for the learner, the teacher and the organisation, are seen as individual data points. The information value of these individual data points is maximised by giving feedback to the learner. There is a decoupling of assessment moment and decision moment. Intermediate and high-stakes decisions are based on multiple data points after a meaningful aggregation of information and supported by rigorous organisational procedures to ensure their dependability. Self-regulation of learning, through analysis of the assessment information and the attainment of the ensuing learning goals, is scaffolded by a mentoring system. Programmatic assessment-for-learning can be applied to any part of the training continuum, provided that the underlying learning conception is constructivist. This paper provides concrete recommendations for implementation of programmatic assessment.

Design principles – Integrate learning and decision function of assessment

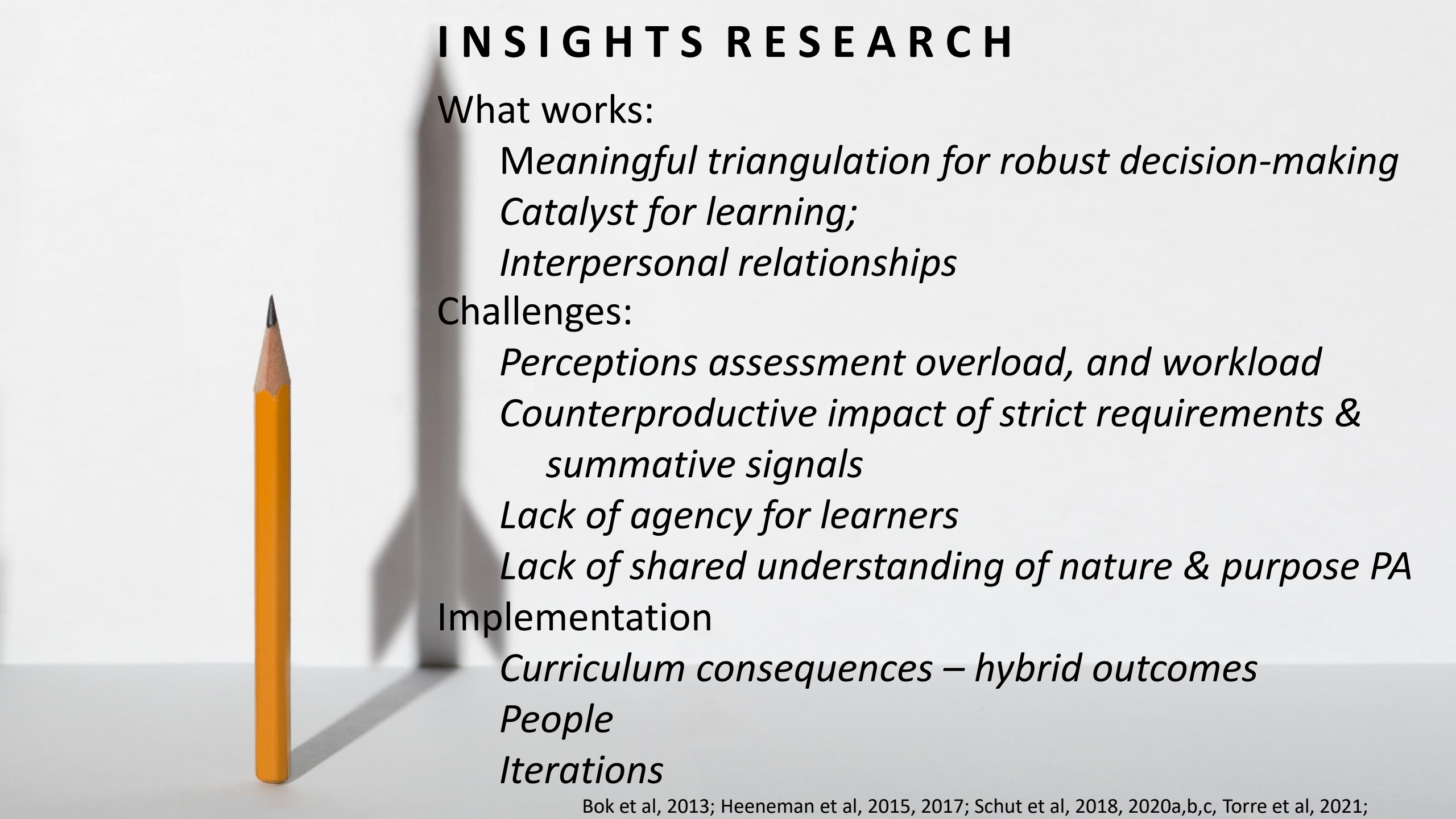
1. A mix of different assessment and feedback formats (data points) provide insight in student development and level of performance
2. Each data point is feedback-oriented and has no fail/pass decisions
3. The learning outcomes or competency framework is the backbone of the program of assessment
4. There is a constant dialogue between the learner and a coach on the use of feedback and self-direction
5. The number of data points is proportionally related to the 'stake' (importance) of the decision
6. The high stake decision (final judgement) is based on human judgement (committee)



van der Vleuten, Schuwirth, colleagues, 1996, 2010, 2012

CONTEXT MATTERS

CONTEXT
MATTERS



INSIGHTS RESEARCH

What works:

- *Meaningful triangulation for robust decision-making*
- *Catalyst for learning;*
- *Interpersonal relationships*

Challenges:

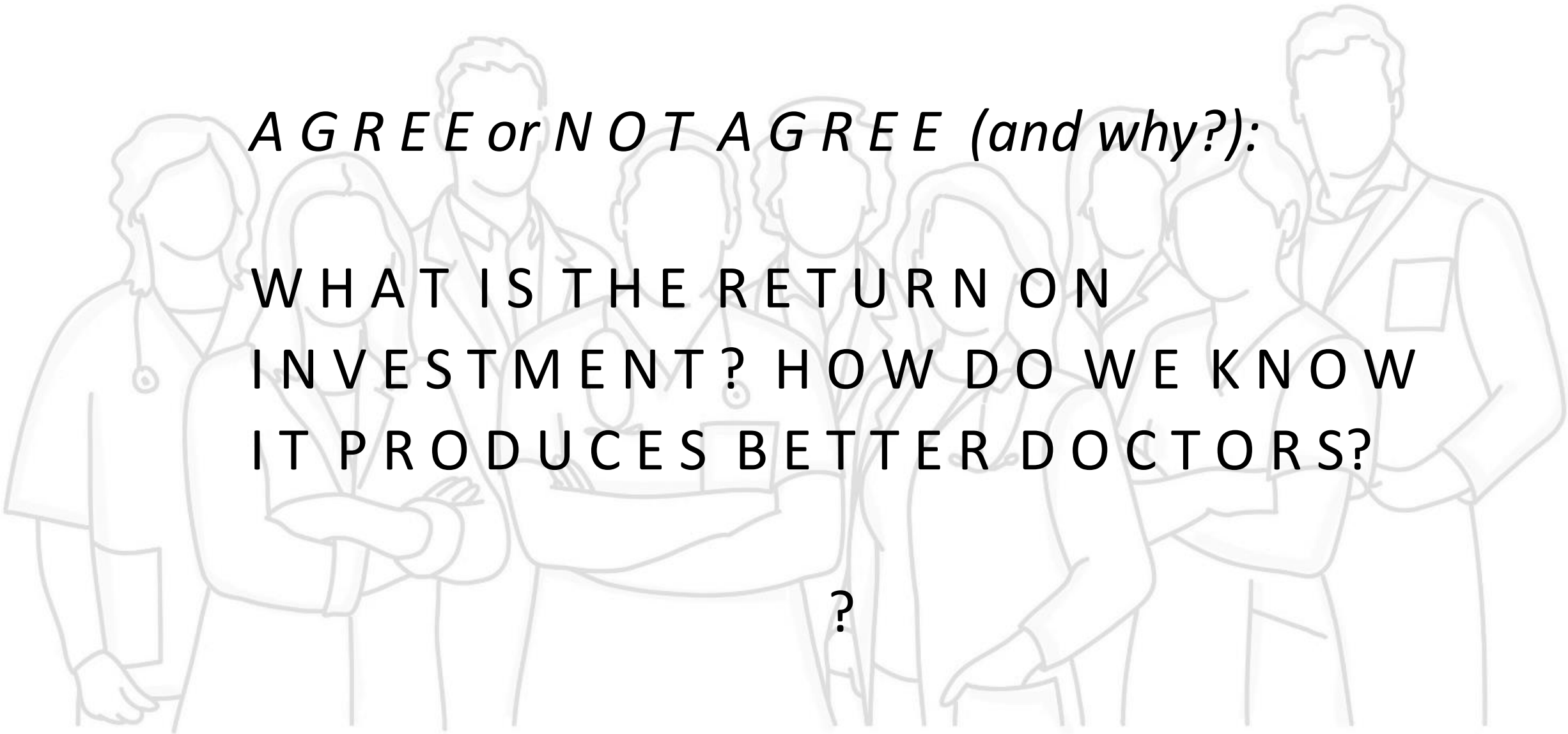
- *Perceptions assessment overload, and workload*
- *Counterproductive impact of strict requirements & summative signals*
- *Lack of agency for learners*
- *Lack of shared understanding of nature & purpose PA*

Implementation

- *Curriculum consequences – hybrid outcomes*
- *People*
- *Iterations*

Bok et al, 2013; Heeneman et al, 2015, 2017; Schut et al, 2018, 2020a,b,c, Torre et al, 2021;

PROGRAMMATIC ASSESSMENT



AGREE or NOT AGREE (and why?):

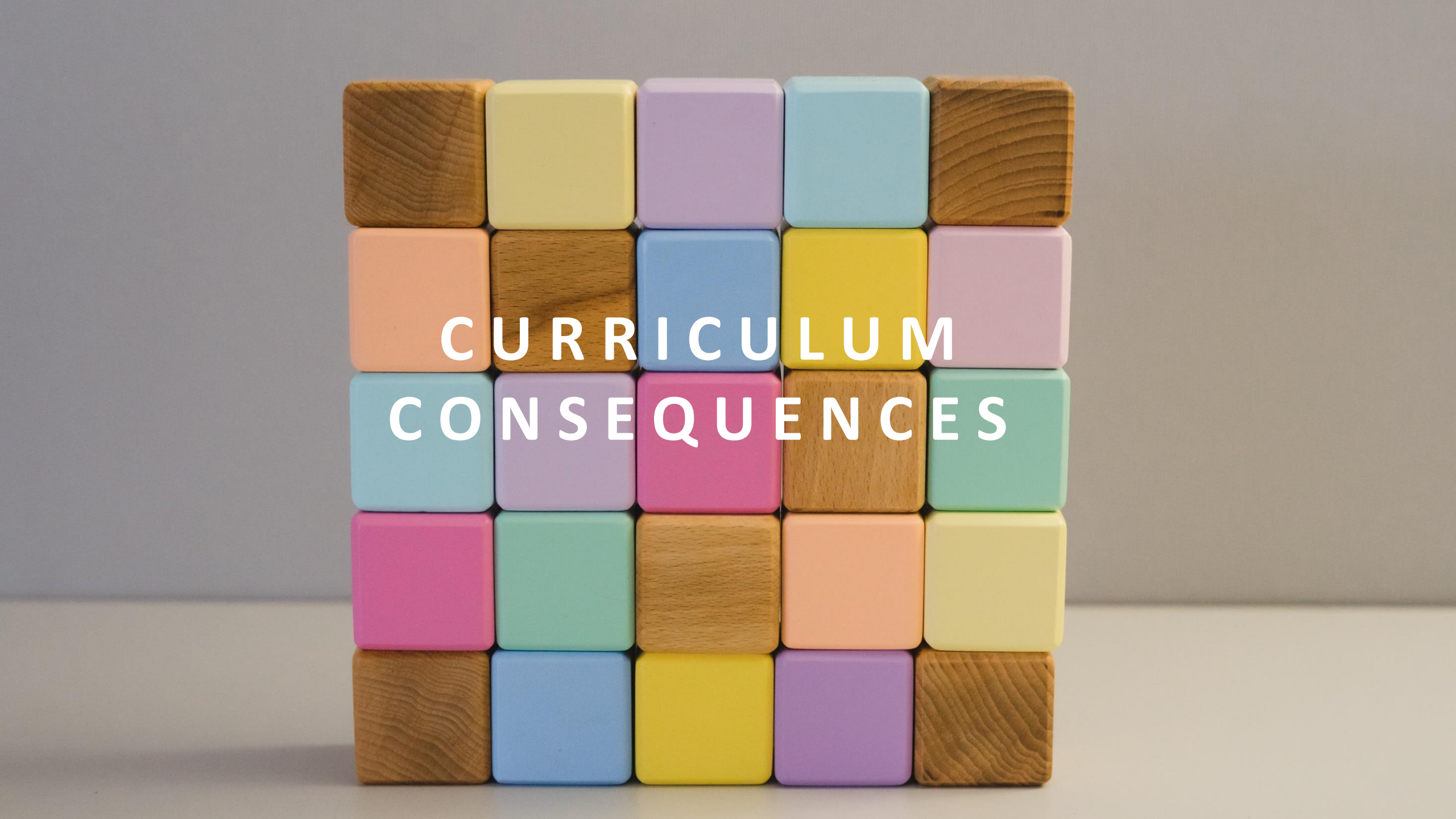
WHAT IS THE RETURN ON INVESTMENT? HOW DO WE KNOW IT PRODUCES BETTER DOCTORS?

?

CONSIDERATIONS & SOME EXAMPLES

DESIGN & IMPLEMENTATION
PROGRAMMATIC ASSESSMENT

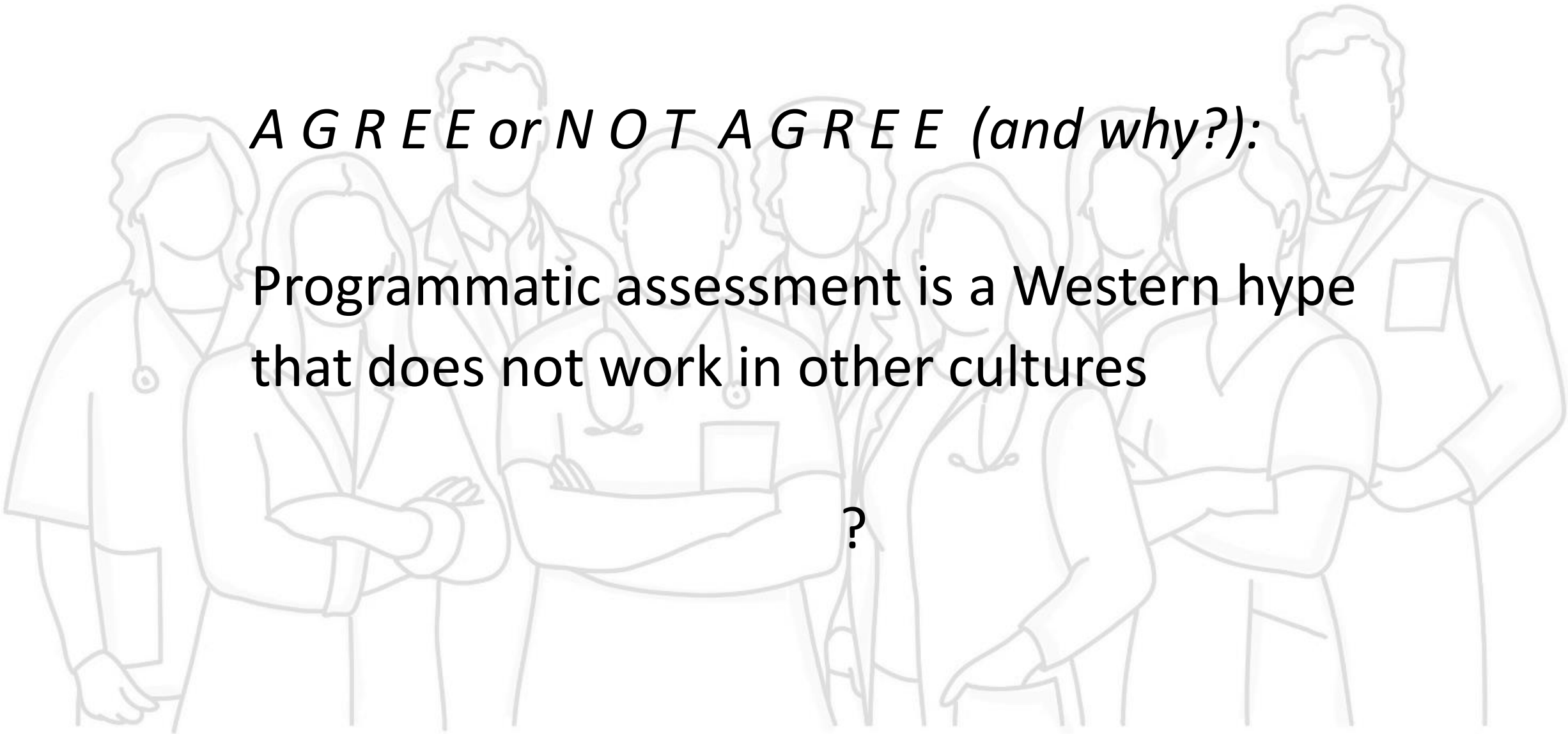
PRINCIPLES
PROGRAMMATIC ASSESMENT


CURRICULUM
CONSEQUENCES

CURRICULUM CONSEQUENCES

HYBRID

P R O G R A M M A T I C A S S E S S M E N T



A G R E E or N O T A G R E E (and why?):

Programmatic assessment is a Western hype that does not work in other cultures

?

WORDS MATTER

LOW STAKE

HIGH STAKE

ASSESSMENT

DECISION

SATURATION

AGGREGATION

AT STAKE

(or mistake?)

DESIGN-ed as low stake datapoint, **perceived** as high stake '***assessment'*** with consequences, due to e.g.

- *scheduled resits (summative signals, use of words)*
- *limited agency for learners*
- *counterproductive impact of strict requirements*
- *internal motivation (learner)*

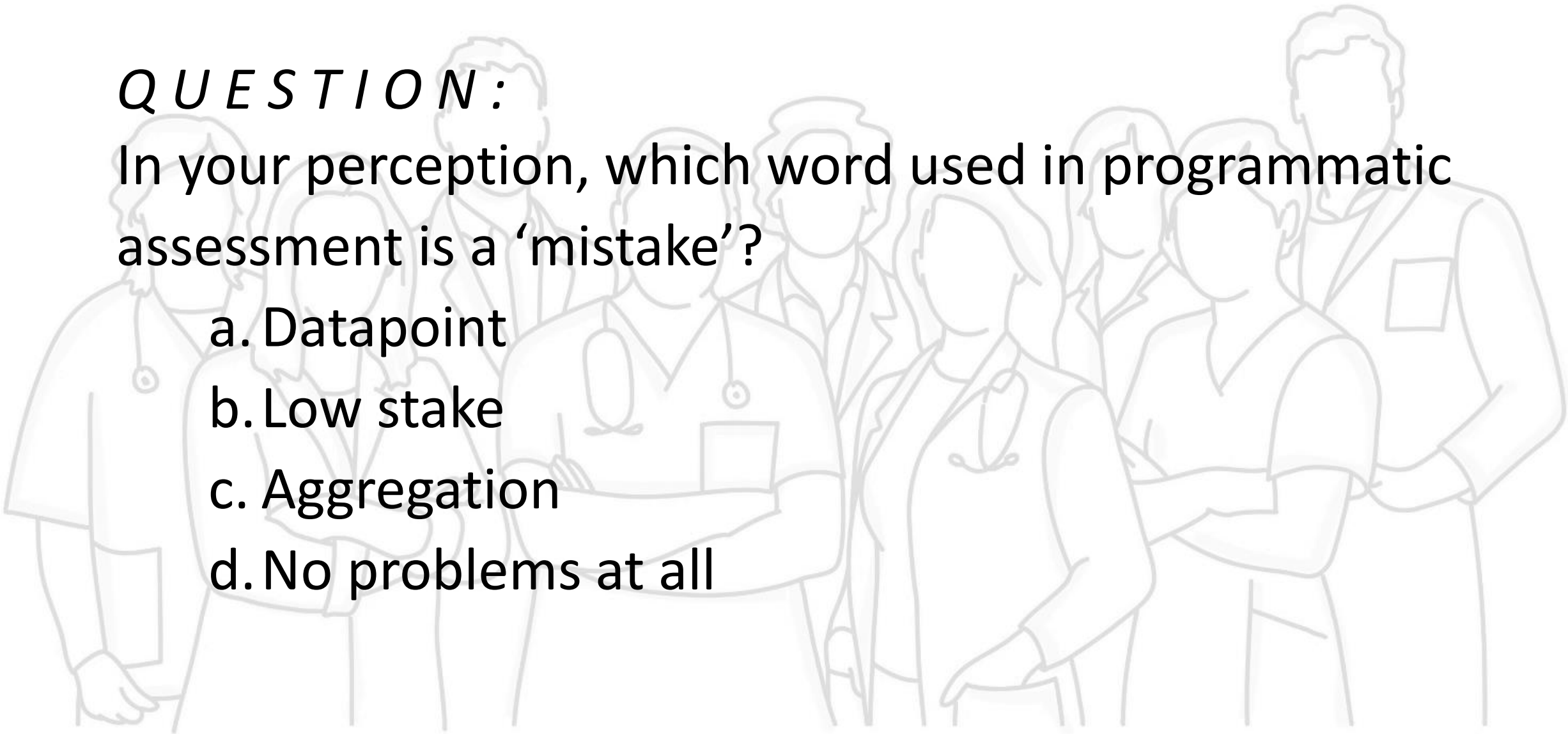
Bok et al, 2013; Heeneman et al, 2015, 2017; Schut et al, 2018, 2020a,b,c, Torre et al, 2021;

P R O G R A M M A T I C A S S E S S M E N T

Q U E S T I O N :

In your perception, which word used in programmatic assessment is a ‘mistake’?

a. Datapoint
b. Low stake
c. Aggregation
d. No problems at all



(s o m e) E X P E R I E N C E S
W O R D S M A T T E R

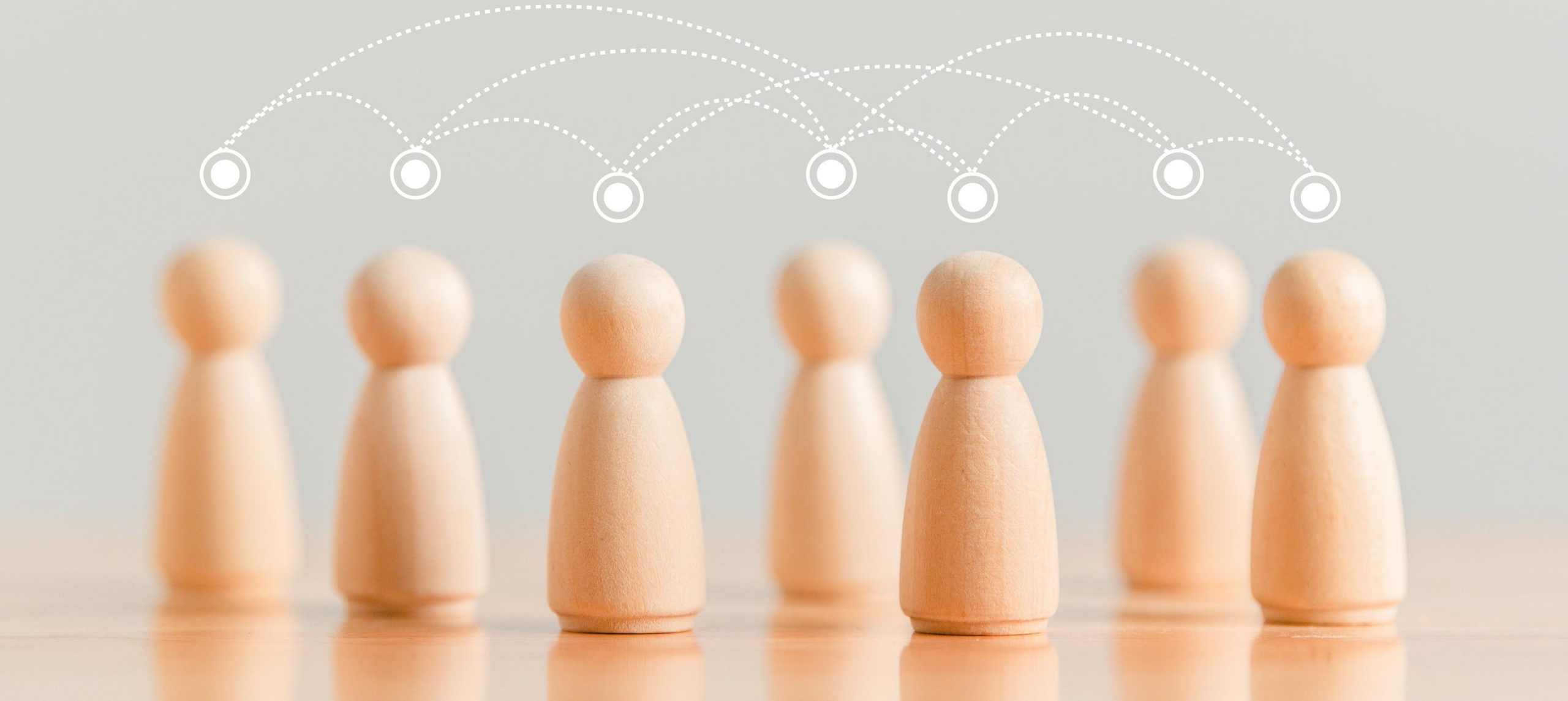
Using words such as

‘Performance information’
‘Competency exam’
‘Self-evaluation test’
‘High-stake decision’ (making)

PEOPLE

ASSUMPTIONS ASSESSMENT

PEOPLE–
SHARED UNDERSTANDING



ASSESSMENT CULTURE

Programmatic assessment: shift towards an assessment culture in which daily feedback was normalized and in which learning and self-reflection could thrive

Schut et al, Persp. Med. Educ. 2020

PEOPLE–
SHARED UNDERSTANDING

Clearly introduce the nature and purpose of PA, for individual datapoints and the system as a whole;

Normalize daily feedback, observation, and follow-up, as well as reflection and continuous improvement

Address learners' and teachers' assessment beliefs and the implications of a learner-led assessment approach

Ensure a supportive infrastructure

Offer leadership in times of change

Bok et al, 2013; Heeneman et al, 2015, 2017; Schut et al, 2018, 2020a,b,c, Torre et al, 2021;

(s o m e) E X P E R I E N C E S

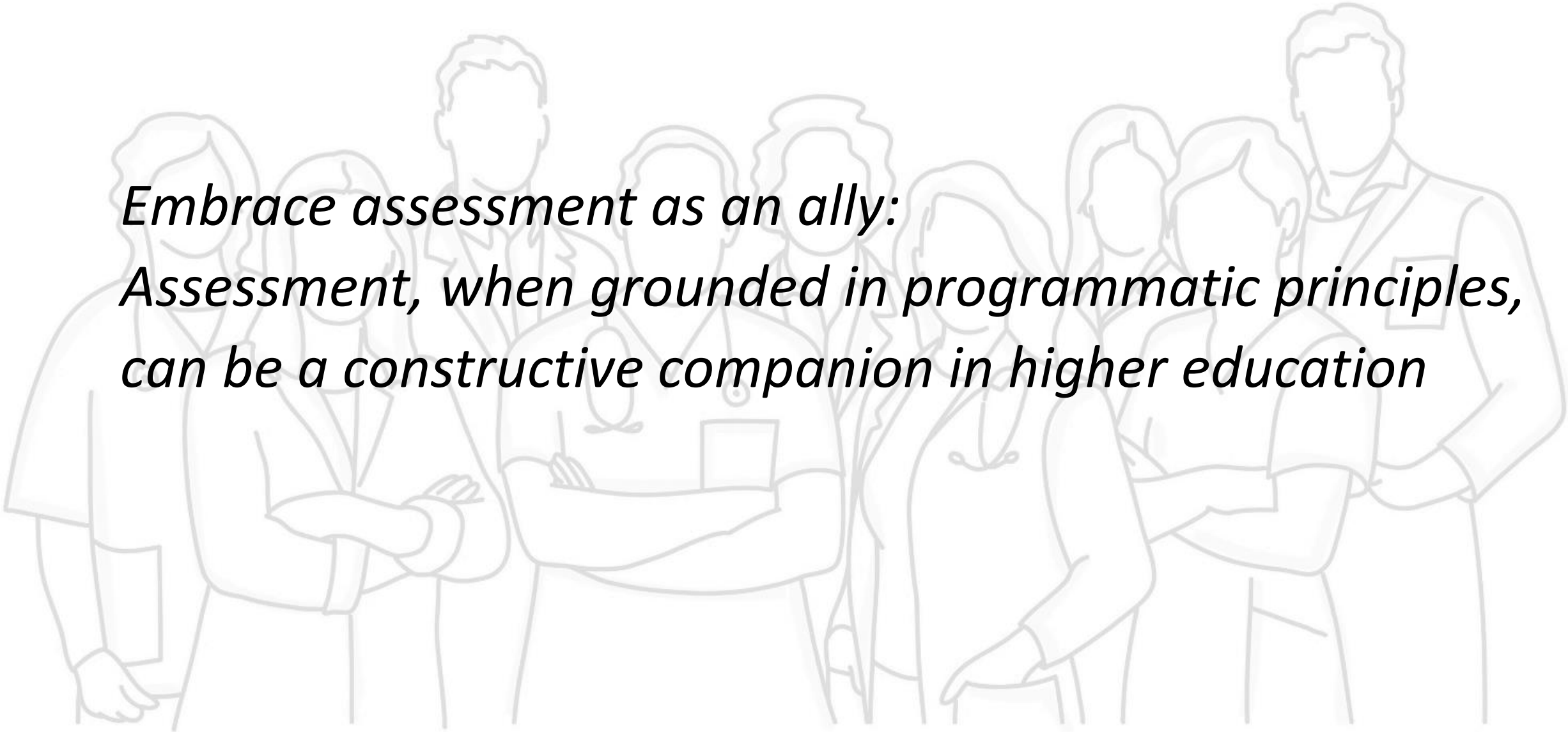
P E O P L E

During design, (always) involve stakeholders, co-design

Implementation in small(er) group or part of curriculum

Students as partners – showing to staff (and new students) what happens with feedback, for learning/ high stake decision-making

P R O G R A M M A T I C A S S E S S M E N T



Embrace assessment as an ally:
Assessment, when grounded in programmatic principles, can be a constructive companion in higher education

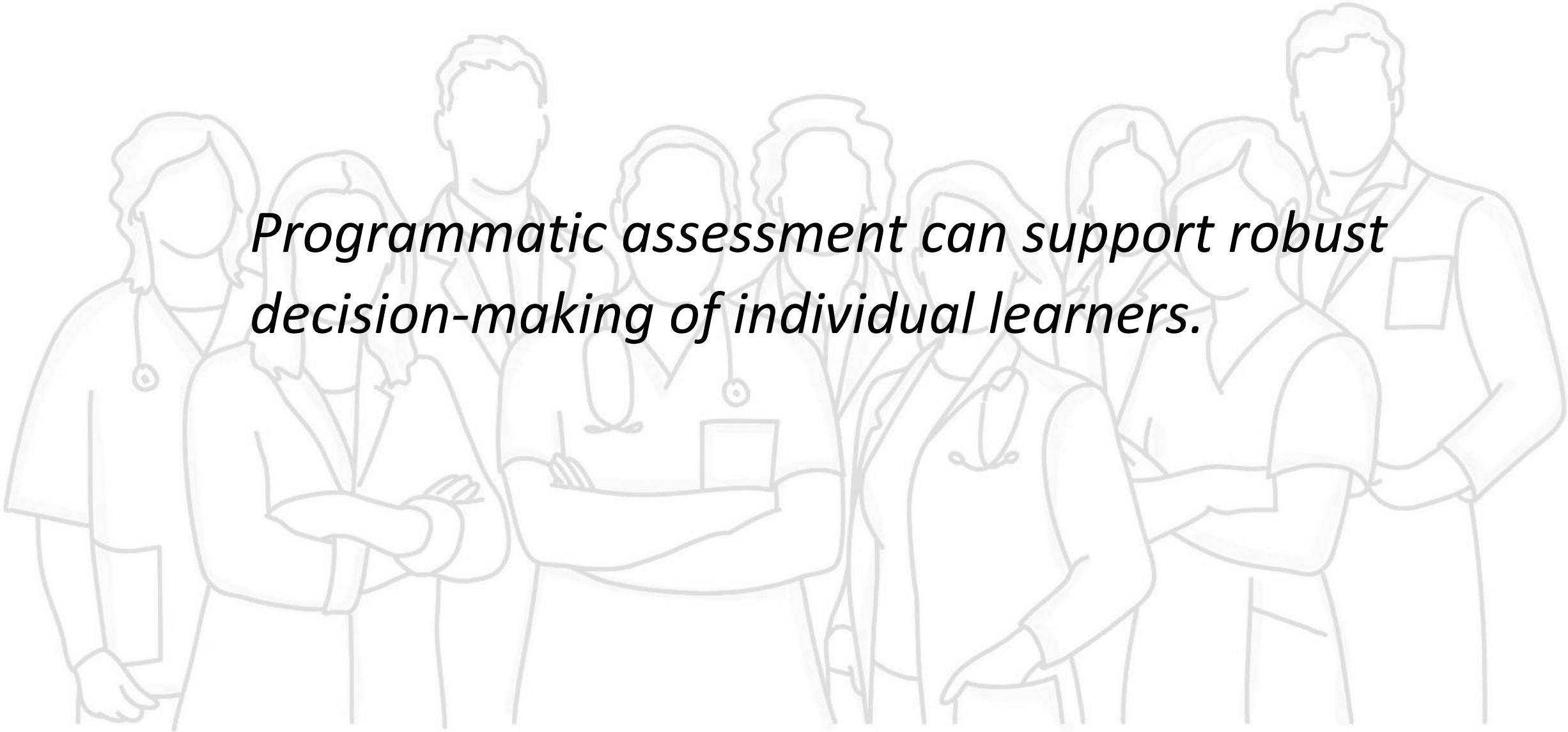
E X A M P L E :

Embrace assessment as an ally: Assessment, when grounded in programmatic principles, can be a constructive companion in higher education

Use of progress test data in the coaching of students

Flinders University

P R O G R A M M A T I C A S S E S S M E N T



Programmatic assessment can support robust decision-making of individual learners.

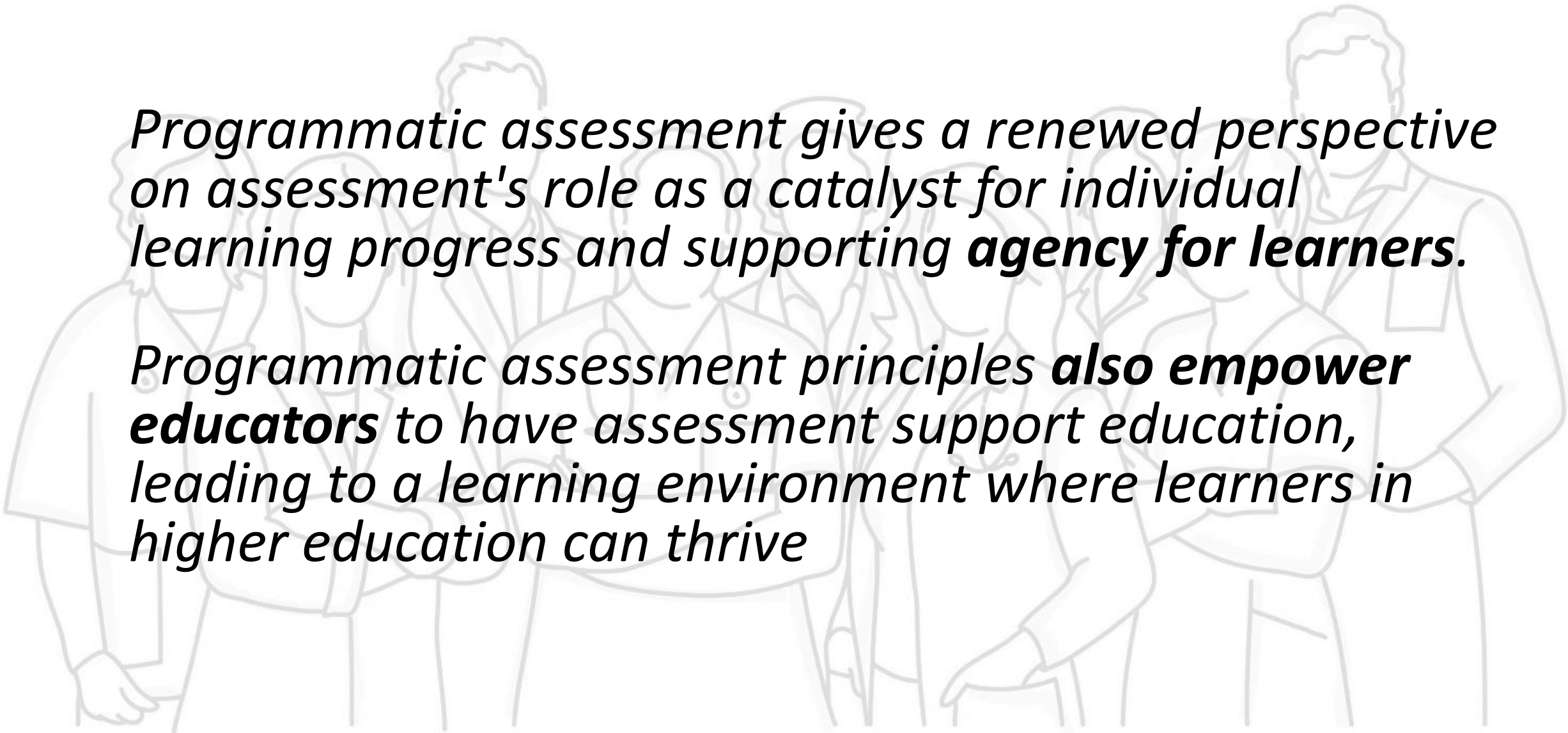
E X A M P L E :

Programmatic assessment can support robust decision-making of individual learners.

Maastricht University

Collection of feedback and information on professionalism embedded in undergraduate medical programme and decision-making on (possible) sanctions (if needed) by committee, using variety of information and a dialogue

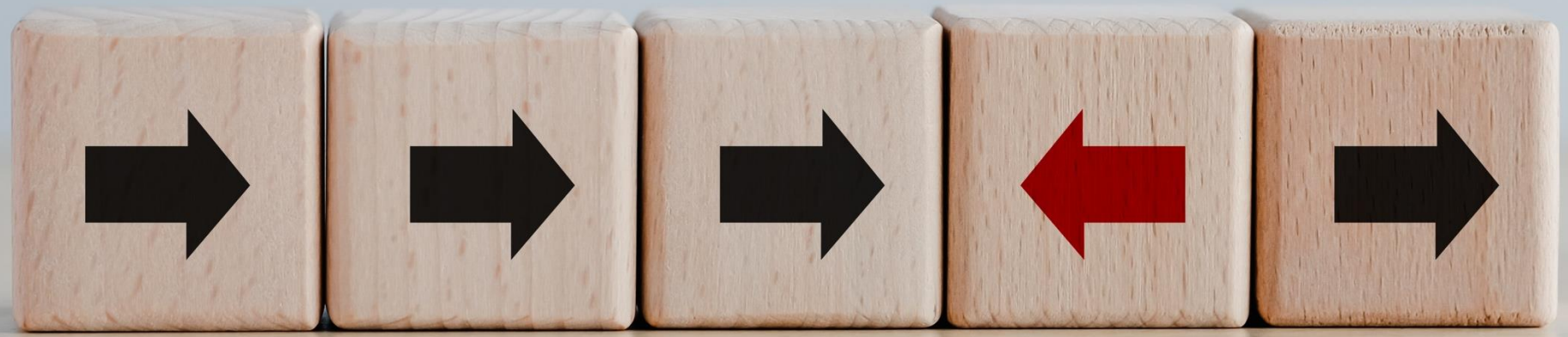
P R O G R A M M A T I C A S S E S S M E N T



Programmatic assessment gives a renewed perspective on assessment's role as a catalyst for individual learning progress and supporting ***agency for learners****.*

Programmatic assessment principles ***also empower educators*** *to have assessment support education, leading to a learning environment where learners in higher education can thrive*

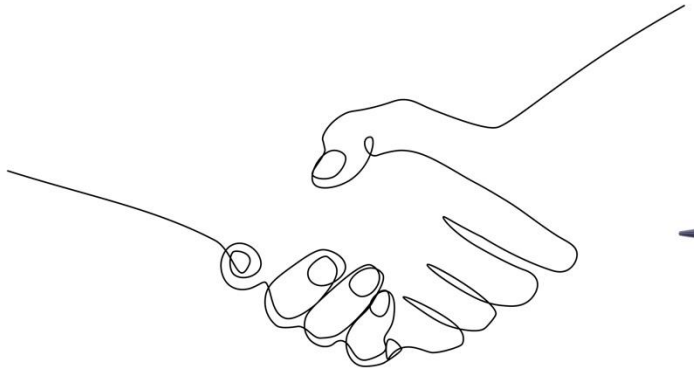
“The intentional actions that constitute learners’ participation in the social experience of learning”



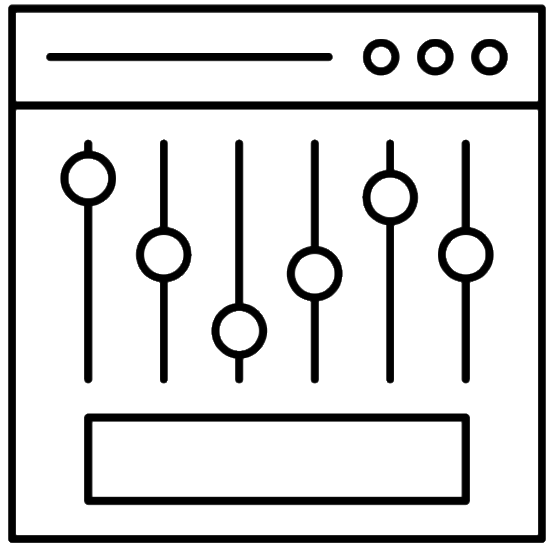
AGENCY

Watling et al. Med Educ, 2021
Billett. Br J Educ Studies. 2008

IN THE CONTEXT OF ASSESSMENT, IS YOUR MINDSET MORE ON.. OR ..?



TRUST



CONTROL

AGENCY IN PROGRAMMATIC ASSESSMENT

Design factors (structure)

- providing or hindering learners' opportunities to exercise control over their assessment experience
- opportunities to
  - collect evidence and improve (offered within programme)

Personal attributes

- Interplay between learner experiences with assessment and confidence in PA

Role of teachers

- learner-teacher relationships

Heeneman et al, 2015; Schut et al, 2020a

E X A M P L E

Bachelor of medicine

Maastricht University

- Self-evaluation testing without invigilation (students decides when & where)
- Student deciding which performance information (evidence) to present to high stake decision Committee

S Y M P O S I U M

Assessment: A Constructive Companion in Higher Education – Embracing Programmatic Assessment for learning

Q U E S T I O N :

What is your main take home message on assessment as a constructive companion?

S Y M P O S I U M

Assessment: A Constructive Companion in Higher Education – Embracing Programmatic Assessment for learning

T H A N K Y O U F O R Y O U R A T T E N T I O N !