

Formative student assessment in virtual and distance education



Dr Saeideh Ghaffarifar,
M.D, M.M.Ed, D.H.E.P

Professor at Medical Education department
Tabriz University of Medical Sciences

ghaffarifars@tbzmed.ac.ir

sa.ghafarifar@gmail.com

<https://edc.tbzmed.ac.ir/Page-247>



Your expectations from today's workshop?



Learning objectives

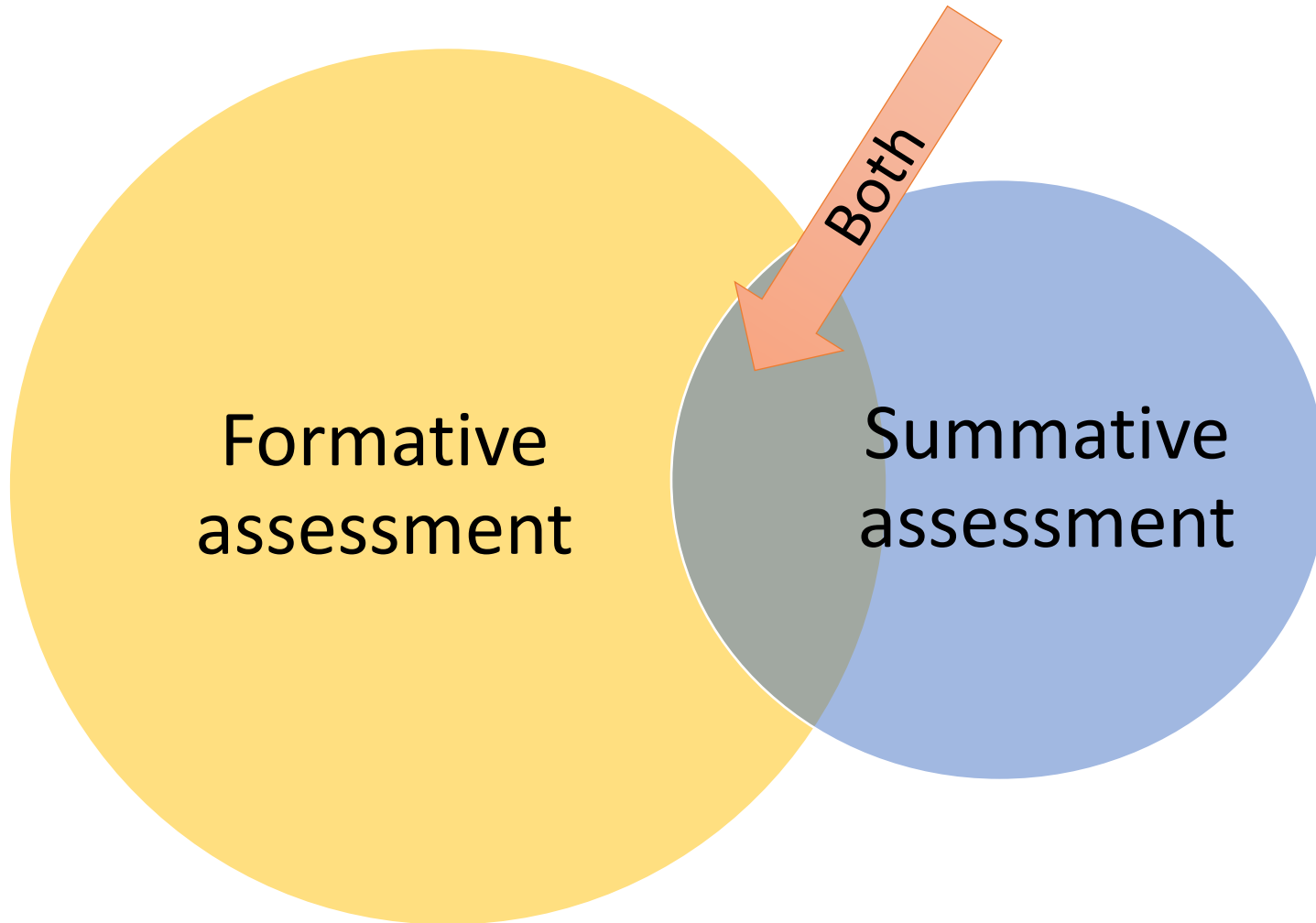
After this presentation, we will be able to:

1. Review fundamentals of student assessment
2. List challenges of virtual formative examinations
3. Criticize actions to buffer challenges of virtual formative student assessments
4. Describe assessment clock as a model to prioritize students' formative assessment in different contexts and situations

Let us know



Approaches to student assessment



Let us know!

- 1) How close to reality were the exam scores you estimated for students under normal circumstances?
- 2) What are the first words that come to your mind when it comes to the formative student assessment?
- 3) Do you agree with holding online structured viva voce examinations?
- 4) Are you in favor of AI-based and live monitoring systems in online student assessments?



Some challenges of formative assessments

Heavy workloads, especially in large class sizes

Necessity of multiple assessments in case of some methods to be credible and trustworthy

Time-consuming data analysis

Feedback providing burden

Leakage!

Quality feedback(qualitative rather quantitatively)

Some challenges of online examinations

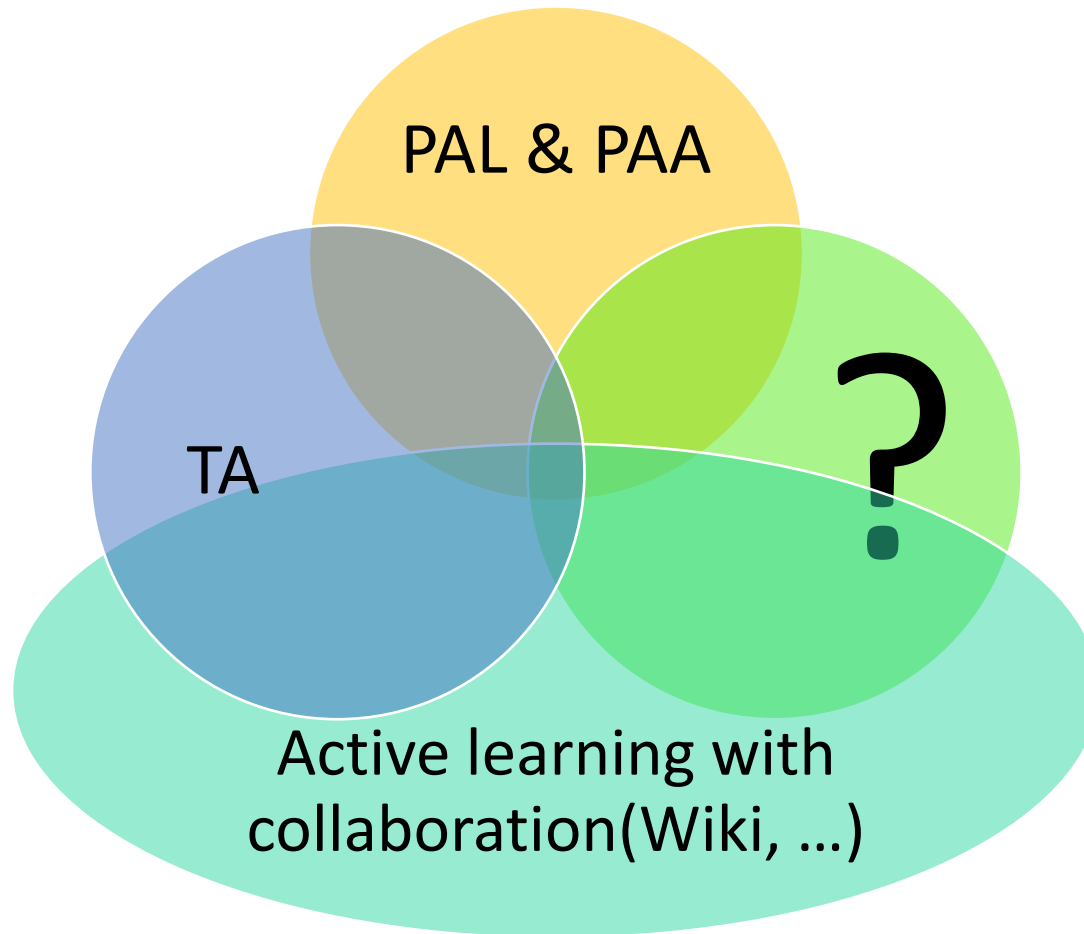
Demanding robust digital infrastructure

Proctoring arrangements

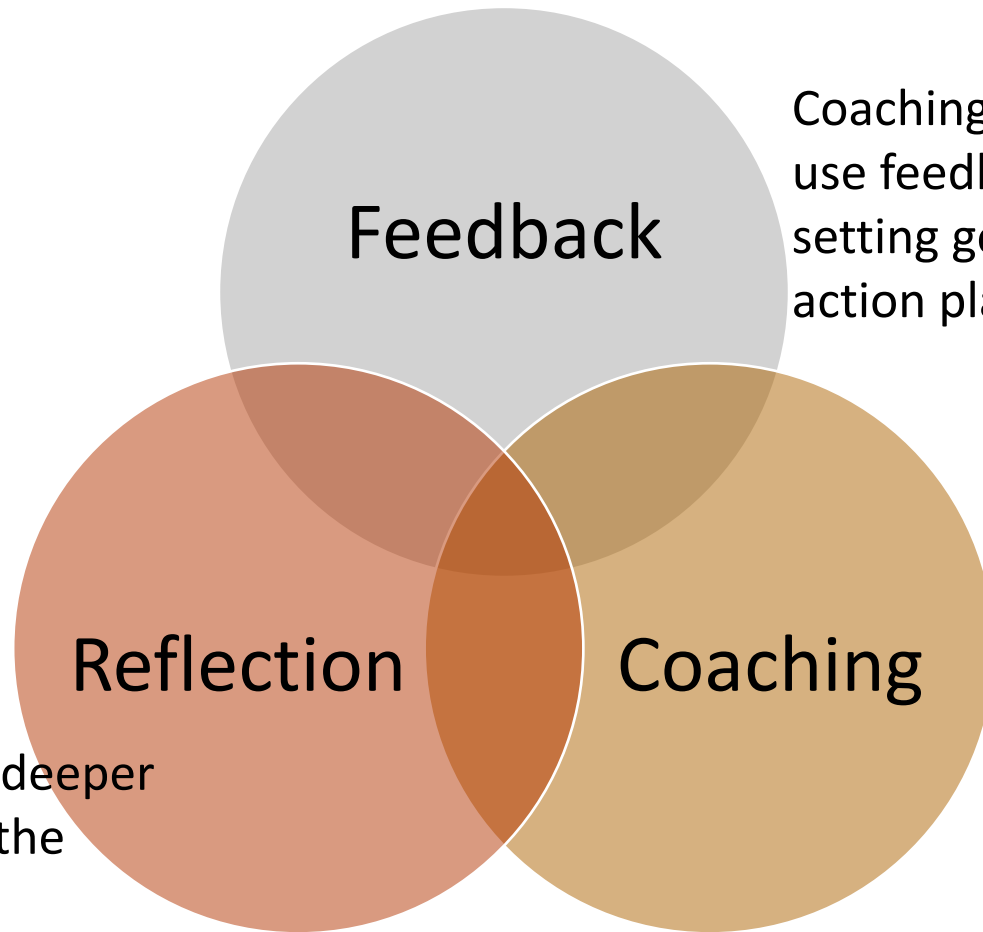
Technical support

Digital divides

Actions to buffer challenges



Actions to buffer challenges



Coaching enables learners to use feedback effectively by setting goals and developing action plans

Reflection fosters deeper understanding of the feedback

Actions to buffer challenges

Lending laptops and Wi-Fi hotspots

Offering asynchronous options

Distributing paper-based assignments in low-resource settings

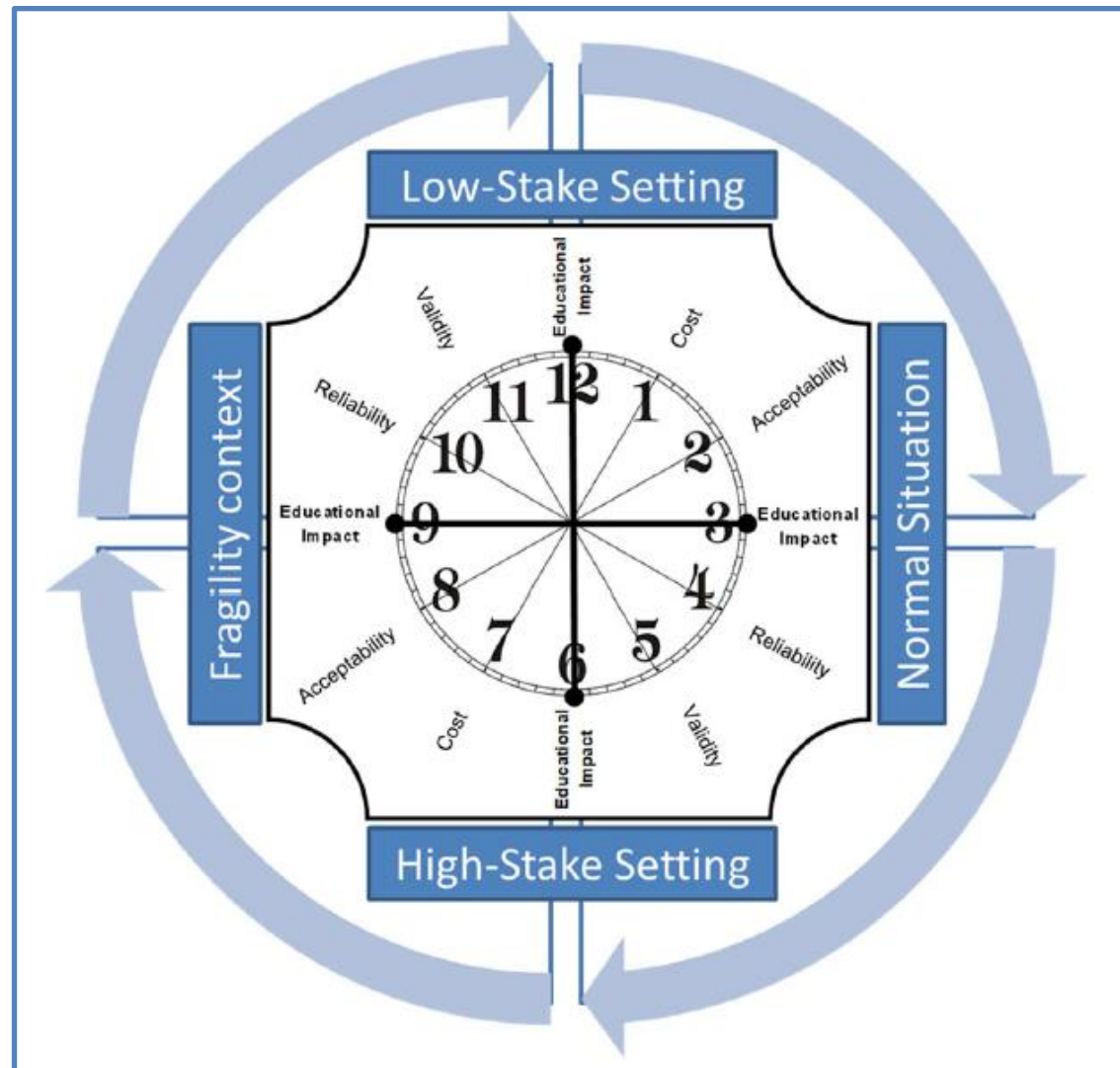
Completing assessments while sheltering in fragile settings

Remote proctoring (Controversial)

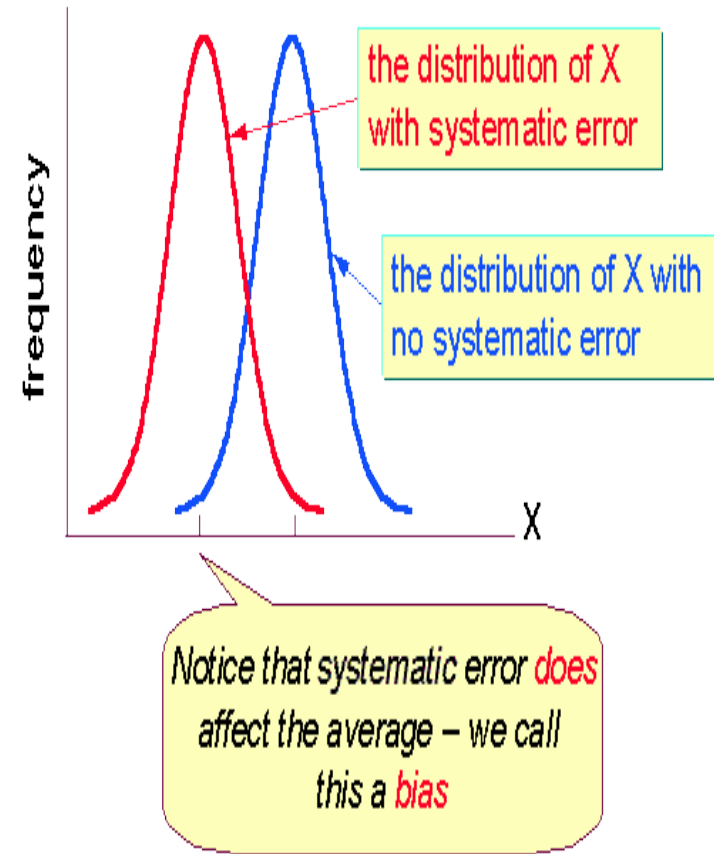
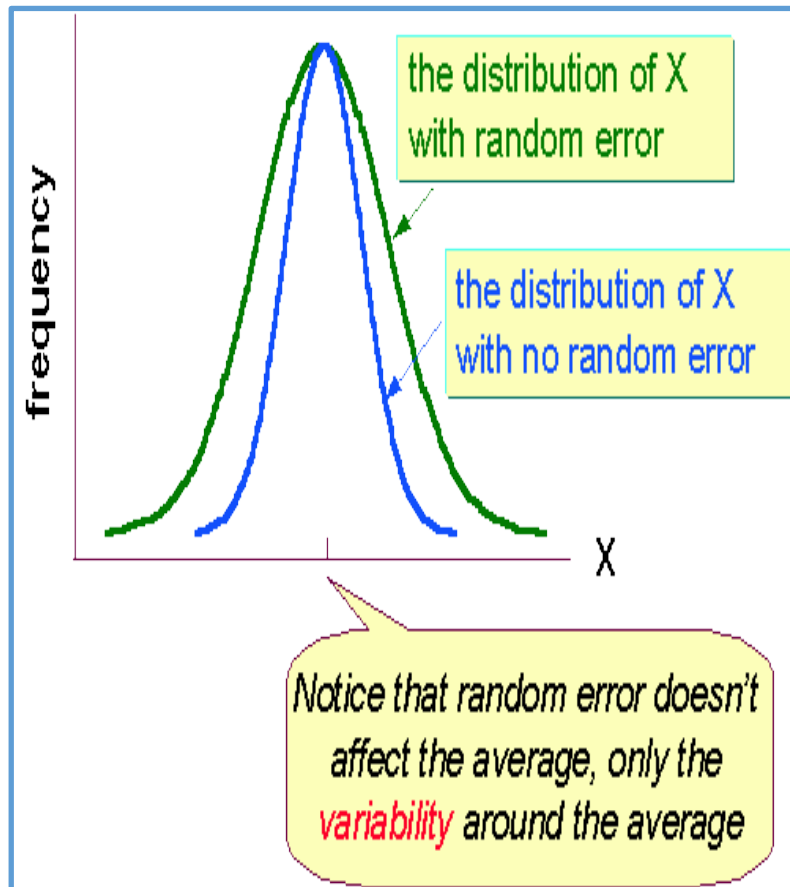
AI-based and live monitoring systems (concerns related to privacy, technical instability, and unequal access)



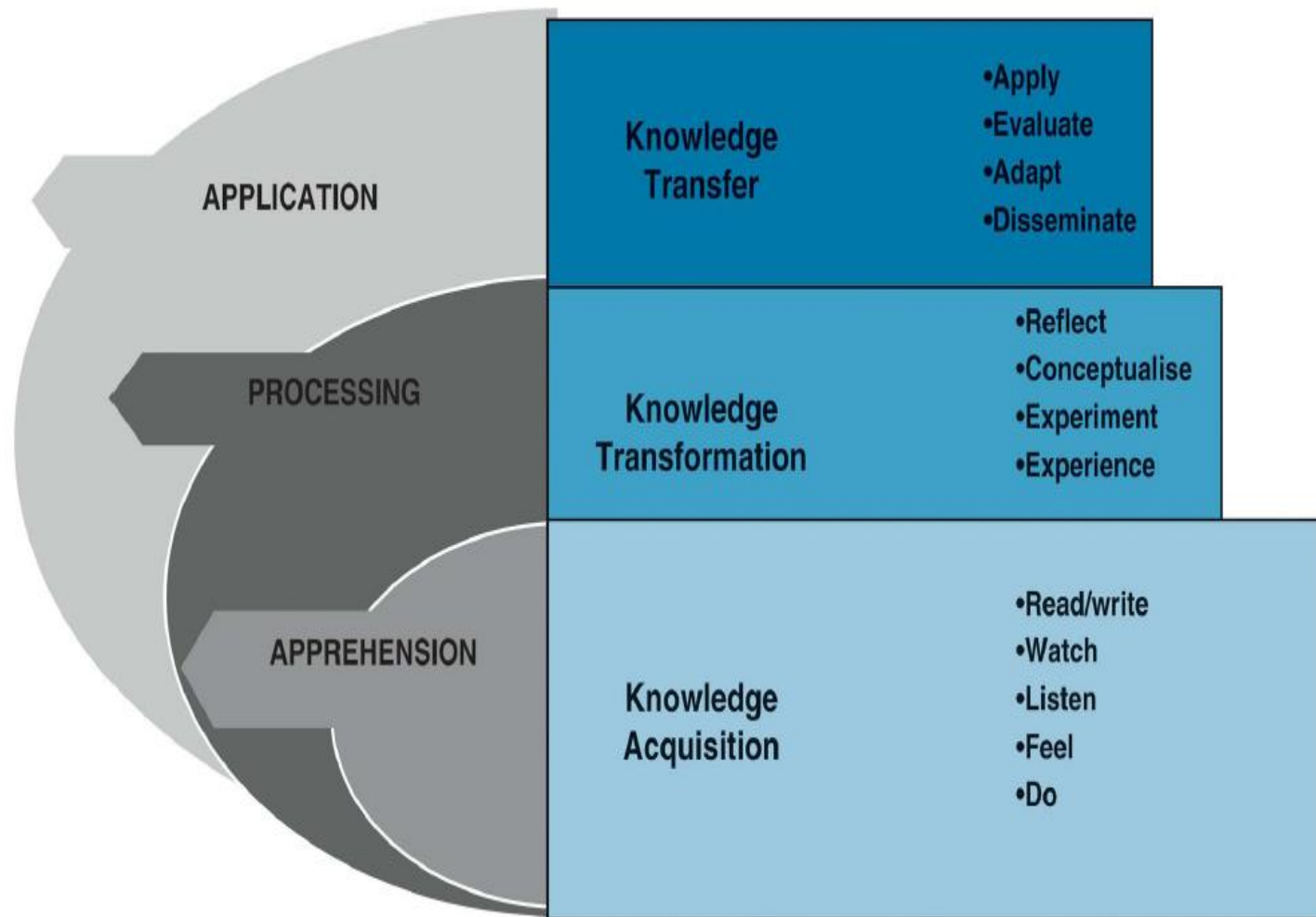
Assessment clock as a model to prioritize students' assessment in fragile contexts



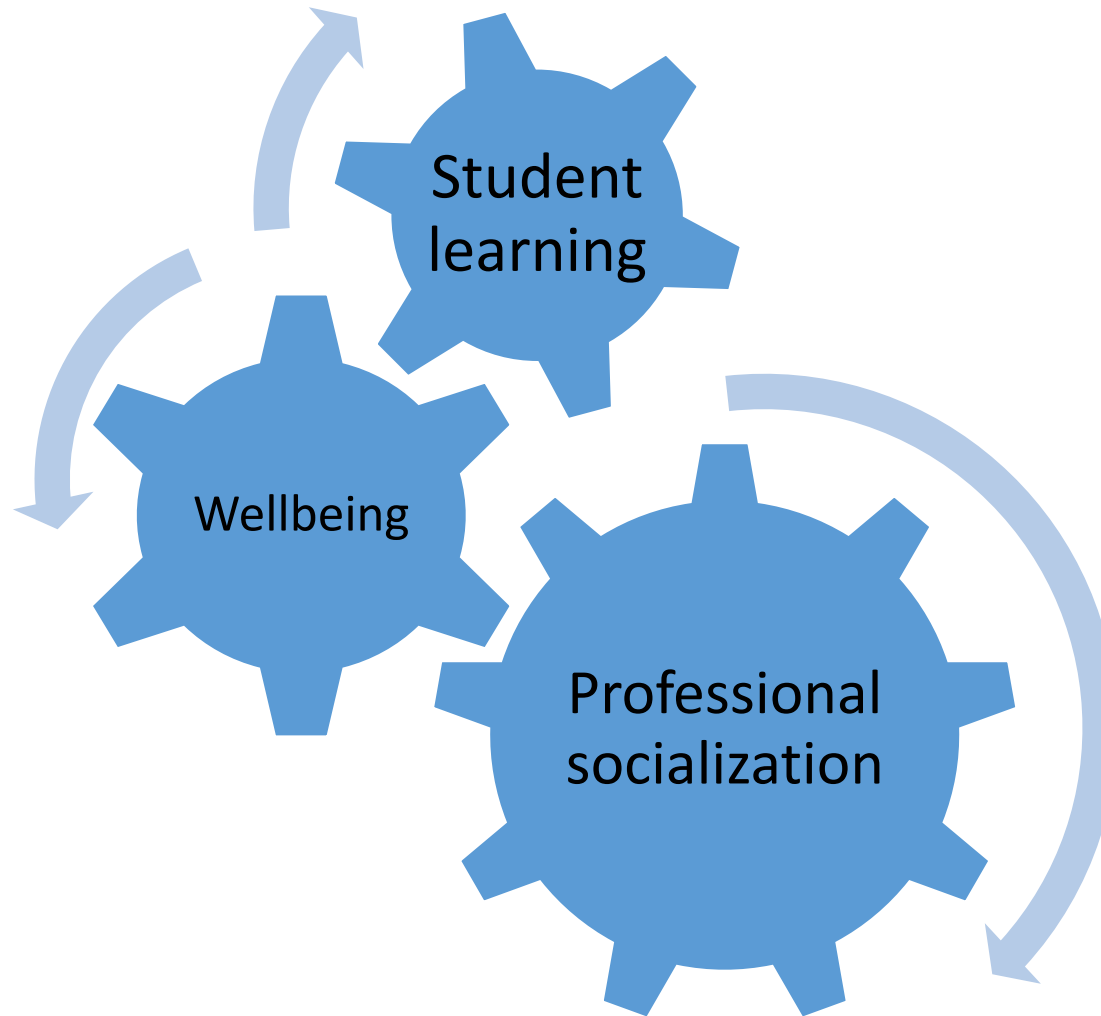
Bias and error in measurement



The focus of student assessment



The focus of the learning environment



Miller's pyramid



Figure 1 Miller's original pyramid. Reprinted with permission from Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9 suppl):s63–s67.

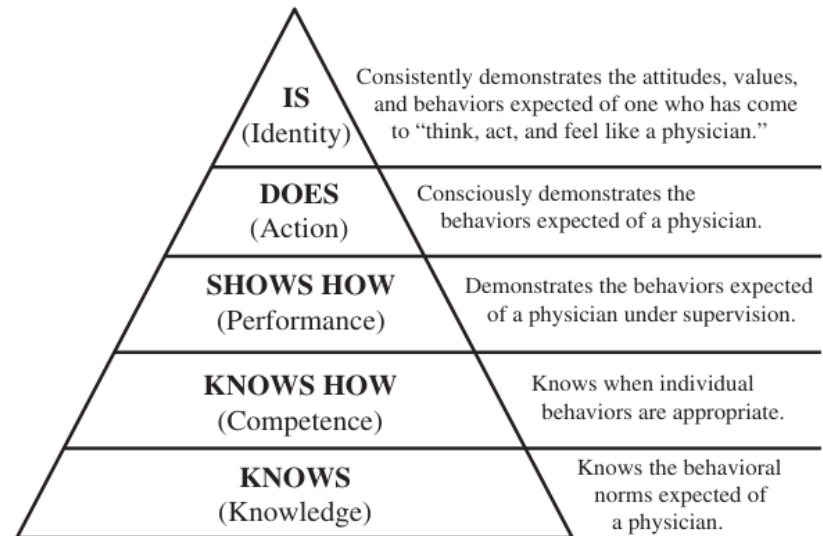
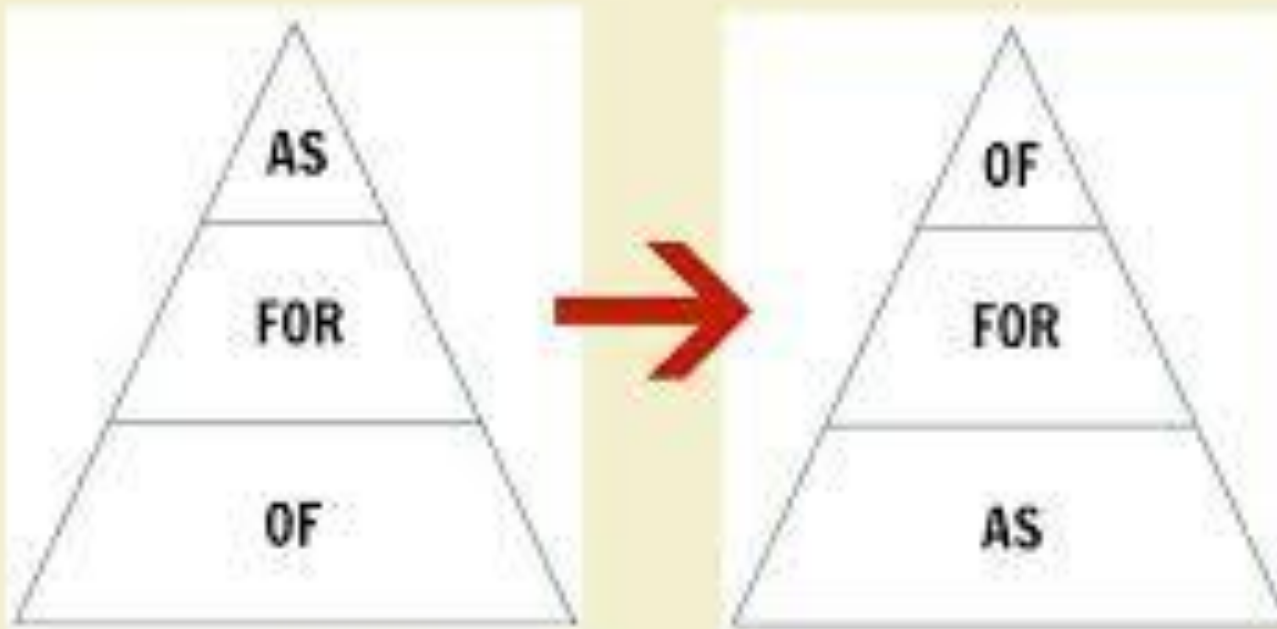


Figure 2 The amended version of Miller's pyramid with the addition of "Is" and an outline of what is to be assessed at each level. Sources: Adapted with permission from Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9 suppl):s63–s67. Quotation from Merton, 1957.²²

The ultimate goal of teaching and assessment



Goal of assessment



**Traditional VS Reconfigured
Assessment Pyramid**

Different approaches to student assessment

Assessment	Of Learning	For Learning	As Learning
Type	Summative	Formative	Formative
What	Teachers determine the progress or application of knowledge or skills against a standard.	Teachers and peers check progress and learning to help learners to determine how to improve.	Learner takes responsibility for their own learning and asks questions about their learning and the learning process and explores how to improve.
Who	Teacher	Teacher & Peers	Learner & Peers
How	Formal assessments used to collect evidence of student progress and may be used for achievement grading on grades.	Involves formal and informal assessment activities as part of learning and to inform the planning of future learning.	Learners use formal and informal feedback and self-assessment to help understand the next steps in learning.
When	Periodic report	Ongoing feedback	Continual reflection
Why	Ranking and reporting	Improve learning	Deeper learning and learning how to learn
Emphasis	Scoring, grades, and competition	Feedback, support, and collaboration	Collaboration, reflection, and self-evaluation

Van Der Vleuten's utility formula

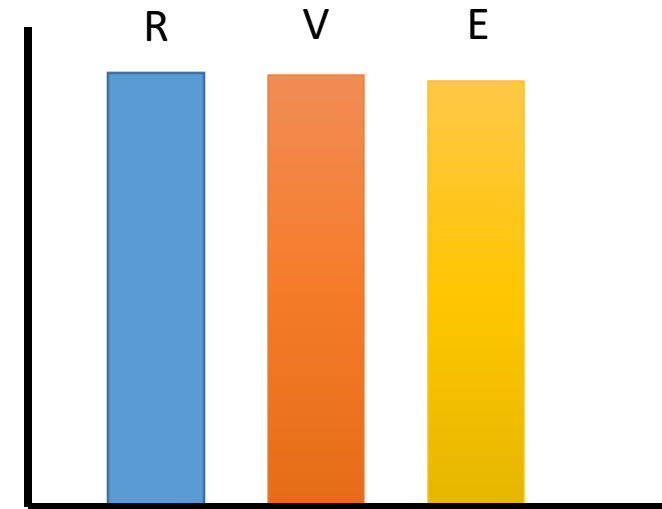
$$\square U = R_w * V_w * E_w * A_w * C_w$$

Five interdependent components of:

- 1) Reliability
- 2) Validity
- 3) Educational impact
- 4) Acceptability
- 5) Cost

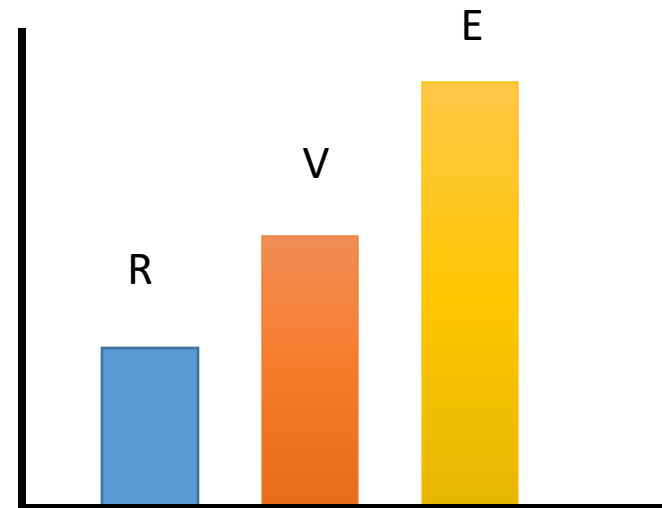
- Not a mathematics multiplication relative
- Not the same weights
- Weight of each element shifts depending on the assessment's purpose and stakes

Assessment attributes for different purposes

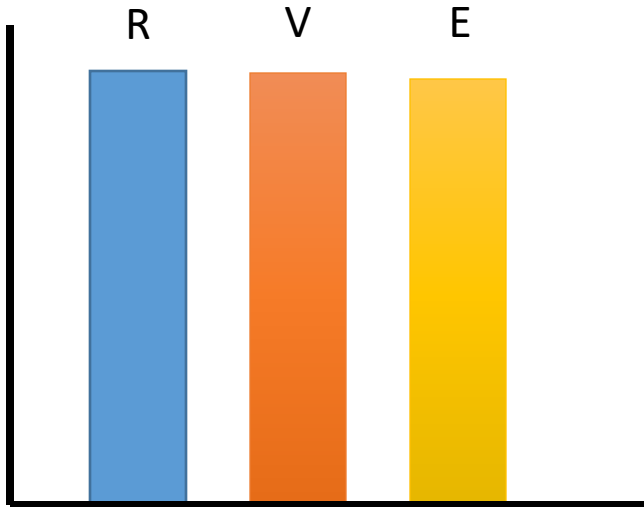


Ideal

Formative

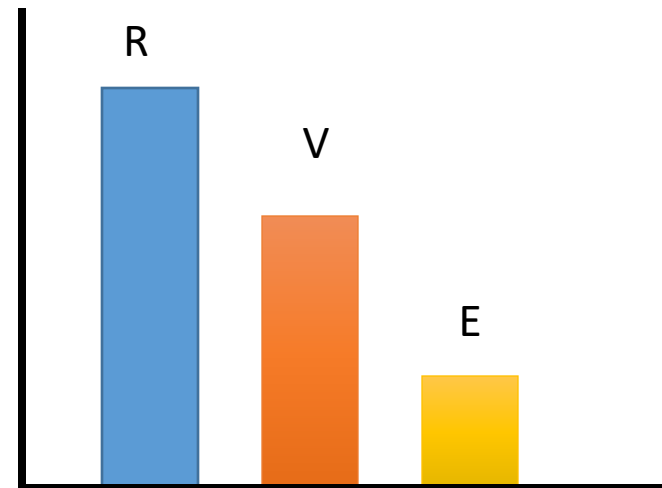


Assessment attributes for different purposes



Ideal

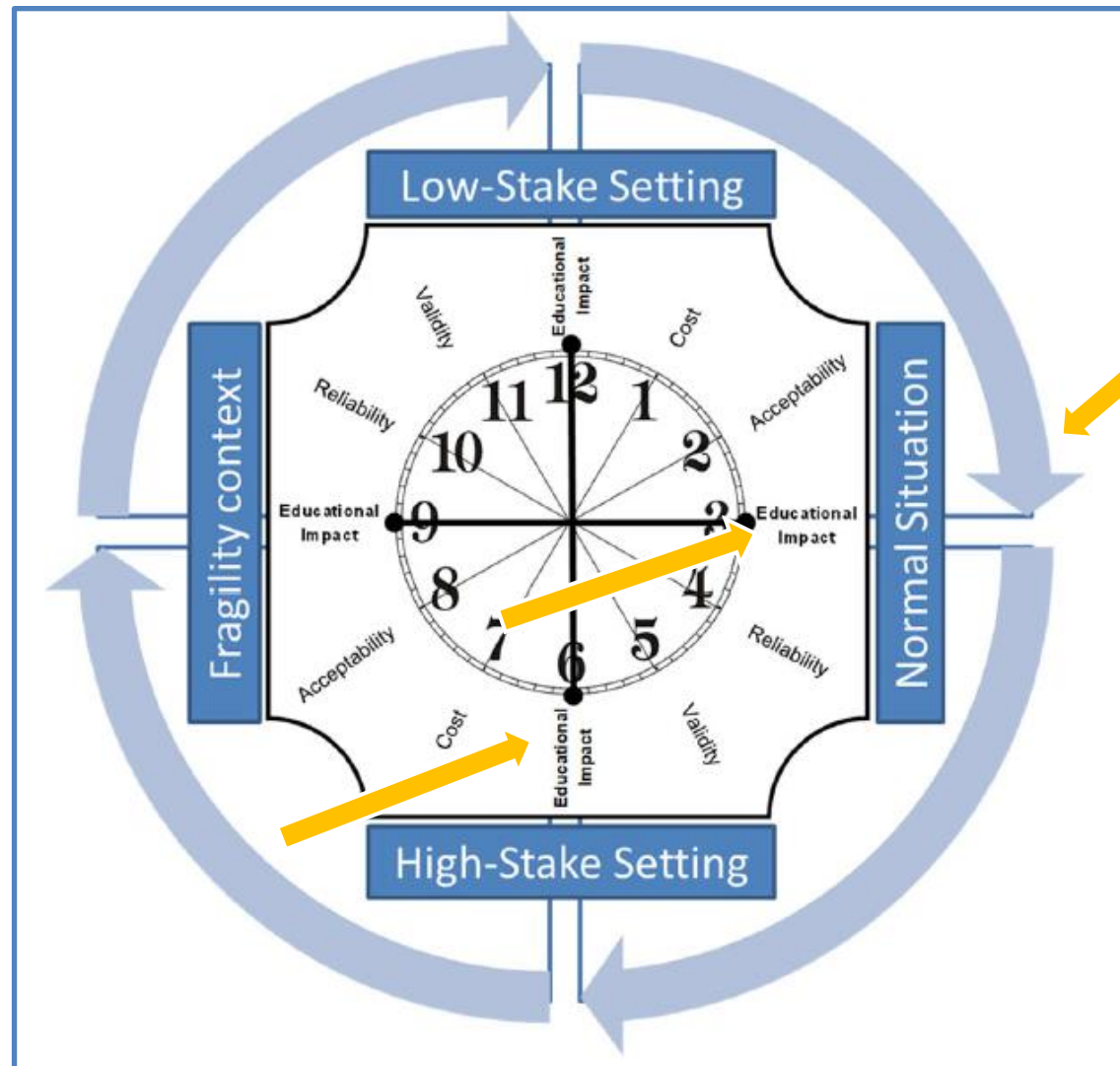
Selection



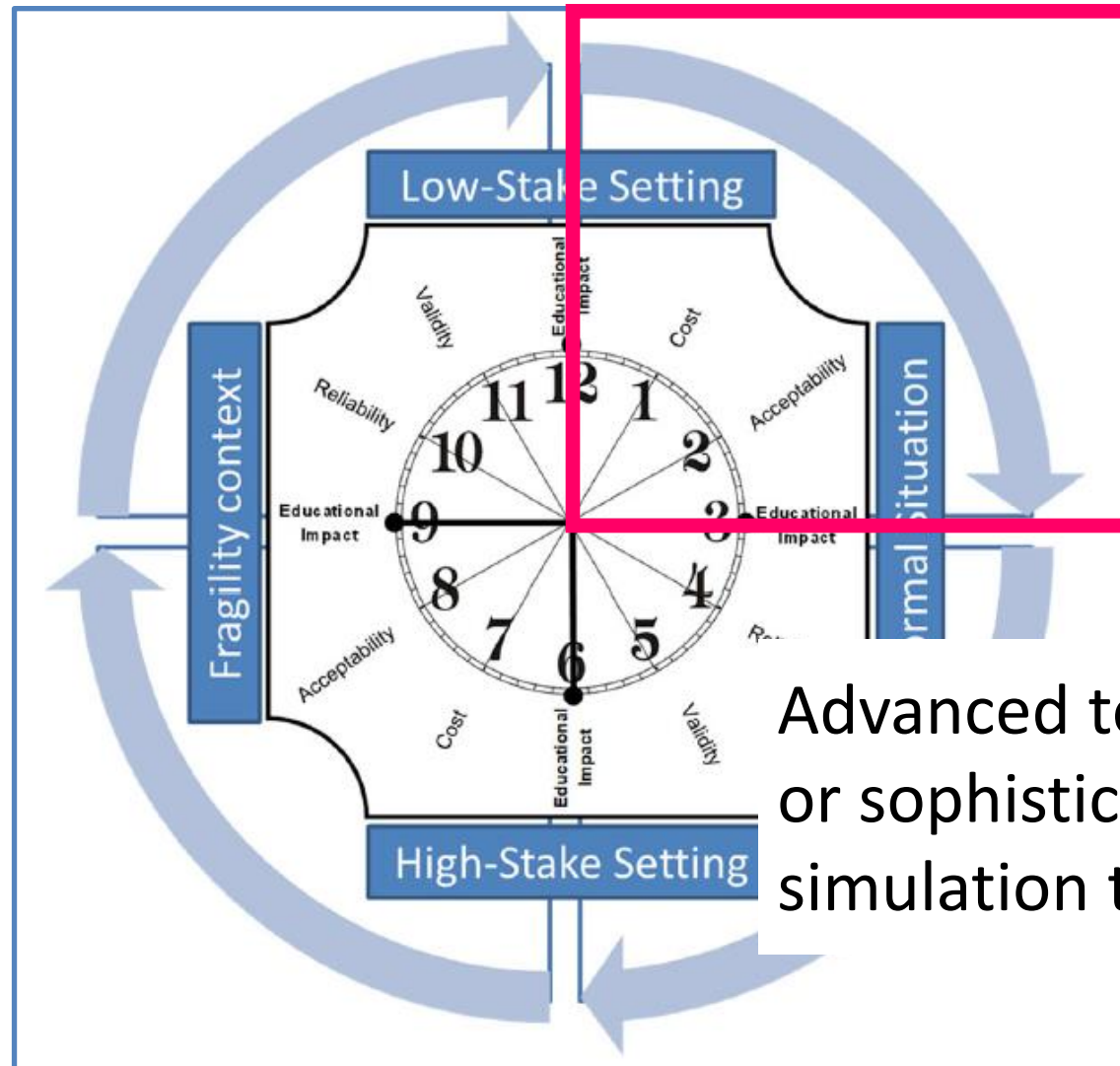
Your summary up to now?



Assessment clock as a model to prioritize students' assessment in fragile contexts



Assessment clock as a model to prioritize students' assessment in fragile contexts



Advanced technologies
or sophisticated
simulation tools

Computerized manikin(High-Fidelity Simulation)

- ❑ Able to breathing, talking, mimicking heart sounds and child-birth



An online game

EIRGENODPC

--	--	--	--	--	--	--	--	--	--

MAAGRIN

--	--	--	--	--	--	--

AVLRIETRE

--	--	--	--	--	--	--	--	--

LLRCEA

--	--	--	--	--	--

OIPECINSR

--	--	--	--	--	--	--	--	--

OCCP

--	--	--	--

CPSA

--	--	--	--

REIC

--	--	--	--

ERQYU

--	--	--	--	--

YGRE

--	--	--	--

HIACLN

--	--	--	--	--	--

AXNSTY

--	--	--	--	--	--

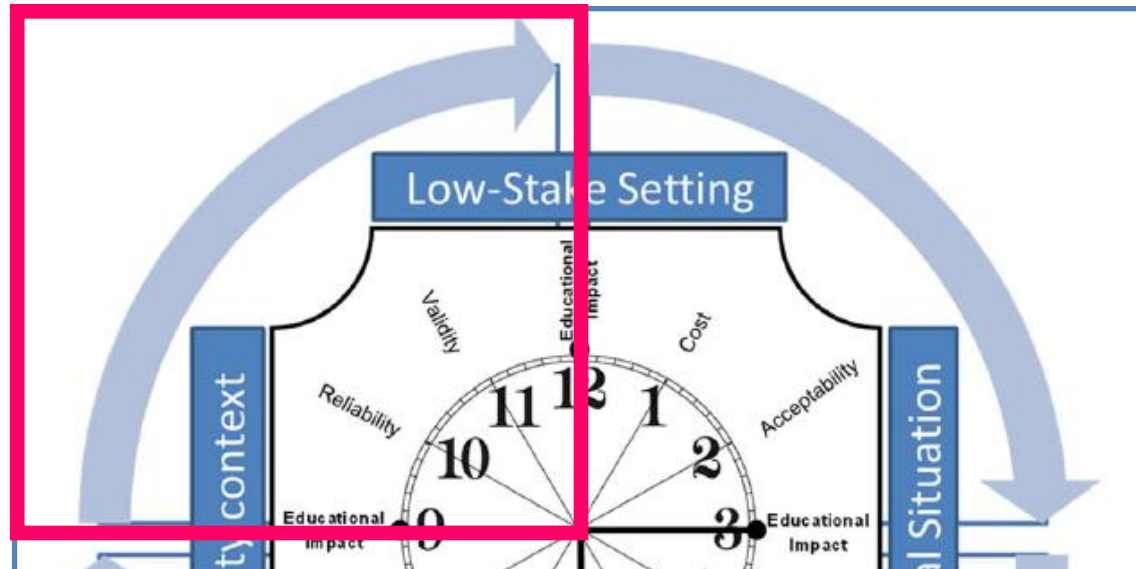
OAONLBE

--	--	--	--	--	--	--

AEATABSD

--	--	--	--	--	--	--	--

Assessment clock as a model to prioritize students' assessment in fragile contexts



Open-book or take-home exams that emphasize knowledge application and reasoning over recall, written assignments and projects – portfolios, reflective essays, research tasks – that allow asynchronous progression

Portfolio



Portfolio

“Compiled by learners themselves, containing evidence of how (learning) goals were met and competence progressed. They often include students’ self-assessments and plans for improving competence.”

(Driessen et al., 2008)

Inclusion of reflective texts and self-analyses as the minimal requirement for portfolio



The content of a Portfolio

Evidence of self-directed learning (summaries, designed questions, emails, certificates, ...)

Products, such as reports, papers, patient management plans, letter of discharge, critical appraisals of a topic

Products, such as reports, papers, patient management plans, letter of discharge, critical appraisals of a topic



Evaluations, such as test scores, feedback forms (e.g., mini-clinical evaluation exercise (mini-CEX)), feedback from peers and tutors, letters from patients expressing appreciation, certificates.

Mini-clinical evaluation exercise (Mini-CEX)

Direct observation of the practitioner's performance in real clinical encounters with patients in the workplace

Then discussing the diagnosis and management with the practitioner and giving them feedback on the encounter

Judging the practitioner in a number of clinical domains and on his or her overall clinical care



Mini-Clinical Evaluation Exercise (Mini-CEX)

1

Average Mini-CEX encounter: around 15 to 25 minutes

2

Giving immediate feedback

3

Identifying strengths and weaknesses

4

Should be performed on multiple occasions with different patients and different assessors

5

Good reliability with between 10 and 14 encounters over time

A Mini-CEX form

Mini-Clinical Evaluation Exercise (CEX)

Evaluator: _____ Date: _____

Resident: _____ ☐ R-1 ☐ R-2 ☐ R-3

Patient Problem/Dx: _____

Setting: ☐ Ambulatory ☐ In-patient ☐ ED ☐ Other _____

Patient: Age: _____ Sex: _____ ☐ New ☐ Follow-up

Complexity: ☐ Low ☐ Moderate ☐ High

Focus: ☐ Data Gathering ☐ Diagnosis ☐ Therapy ☐ Counseling

1. Medical Interviewing Skills (☐ Not observed)

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

2. Physical Examination Skills (☐ Not observed)

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

3. Humanistic Qualities/Professionalism

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

4. Clinical Judgment (☐ Not observed)

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

5. Counseling Skills (☐ Not observed)

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

6. Organization/Efficiency (☐ Not observed)

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

7. Overall Clinical Competence (☐ Not observed)

1	2	3	4	5	6	7	8	9
UNSATISFACTORY			SATISFACTORY			SUPERIOR		

Mini-CEX Time: Observing _____ Mins Providing Feedback: _____ Mins

Evaluator Satisfaction with Mini-CEX

LOW	1	2	3	4	5	6	7	8	9	HIGH
-----	---	---	---	---	---	---	---	---	---	------

Resident Satisfaction with Mini-CEX

LOW	1	2	3	4	5	6	7	8	9	HIGH
-----	---	---	---	---	---	---	---	---	---	------

Comments: _____

Direct observation of procedural skills (DOPS)

Specifically for assessing practical skills

Direct observation of a practitioner by an assessor while undertaking a procedure **on a real patient**

Judging specific components of the procedure and providing feedback after the event

- ❑ Mean observation time for DOPS: varies according to the procedure assessed
- ❑ On average, feedback time takes an additional third of the procedure observation time (Wilkinson et al., 1998).



Applicant name:

Assessor name:

Date of procedure:

Procedure observed:

Domain	Satisfactory	Needs improvement	Unsatisfactory
Professional approach (includes communication, consent, and patient)	☐	☐	☐
Knowledge (anatomy, technique)	☐	☐	☐
Appropriate Pre-Procedure Preparation	☐	☐	☐
Appropriate Analgesia	☐	☐	☐
Aseptic Technique	☐	☐	☐
Technical Ability	☐	☐	☐
Appropriate Post-Procedure Recommendations	☐	☐	☐

Domain	Satisfactory	Needs improvement	Unsatisfactory
Overall Ability to Perform Procedure	☐	☐	☐

Comments:

A DOPS
form

CBD

- Case Based Discussion
- In UK and Australasia

CSR

- Chart Stimulated Recall
- In US

CBD/CSR

“A structured interview in which practitioners discuss aspects of a case in which they have been involved to explore their underlying reasoning, decision making and ethical understanding”



CBD/CSR

Can be used in a variety of settings, such as clinics, wards or assessment units, and on different clinical problems

Selecting suitable cases by practitioner and giving his or her written records of these to the assessor in advance

Presenting the case, which has been selected by the assessor, to her/him

Asking questions by assessor to probe the practitioner's clinical reasoning and professional judgement

Chart Stimulated Recall (CSR)

Exam questions based on the patients' files (3-6 files)

To assess knowledge, skills, decision skills, problem solving skills

To assess communication skills & professionalism



Each exposure 20 minutes (5 minutes for feedback)

Case-based discussion assessment form



Australian Medical Council Limited

Logo placement area

Candidate and assessor information

AMC candidate name		AMC candidate number	
Assessor name		Assessor position	

This Case-based discussion assesses the following domains (multiple options can be selected):

History	☐	Physical Examination	☐	Management/Counselling	☐
Clinical Judgement	☐	Communication Skills	☐	Working in a team	☐
Cultural Competence	☐	Professionalism	☐	Patient Safety and Quality of Care	☐

Patient information

Age		Gender		Setting (E.g. ED/GP/Ward)		
Real Patient	Direct observation of an encounter with a real patient is mandatory.					
Candidate involved in patient's care	Y ☐ N ☐					
Problem(s)						
Please circle Clinical area	Adult medicine	Adult surgery	Women's health	Child health	Mental health	Emergency medicine

Please record a rating for each criterion on the scale 1 (extremely poor) to 5 (extremely good). A score of 1-2 is considered below expected level, 3 at expected level and 4-5 above expected level, at the standard of an Australian graduate at the end of PGY1. The criteria where there are no N/O (not observable in this encounter) boxes are mandatory and must be rated for each assessment. Assessors should note that over all the encounters observed it is expected that all attributes are observed and scored at least once. Support all ratings with an explanation / example in the comments box.

Candidate assessment criteria

	Below expected level	At expected level	Above expected level
1. Clinical record keeping	1 ☐	2 ☐	3 ☐
2. Differential diagnosis and summary list	1 ☐	2 ☐	3 ☐
3. Management plan – Investigations, treatment and follow-up	1 ☐	2 ☐	3 ☐
4. Clinical judgement / clinical reasoning	1 ☐	2 ☐	3 ☐

Global rating

An overall rating of this doctor's performance and professionalism in all areas. The global rating is not an algorithmic calculation of the candidate assessment criteria ratings but a judgement about the overall performance of the candidate.

Not competent

☐

Competent

☐

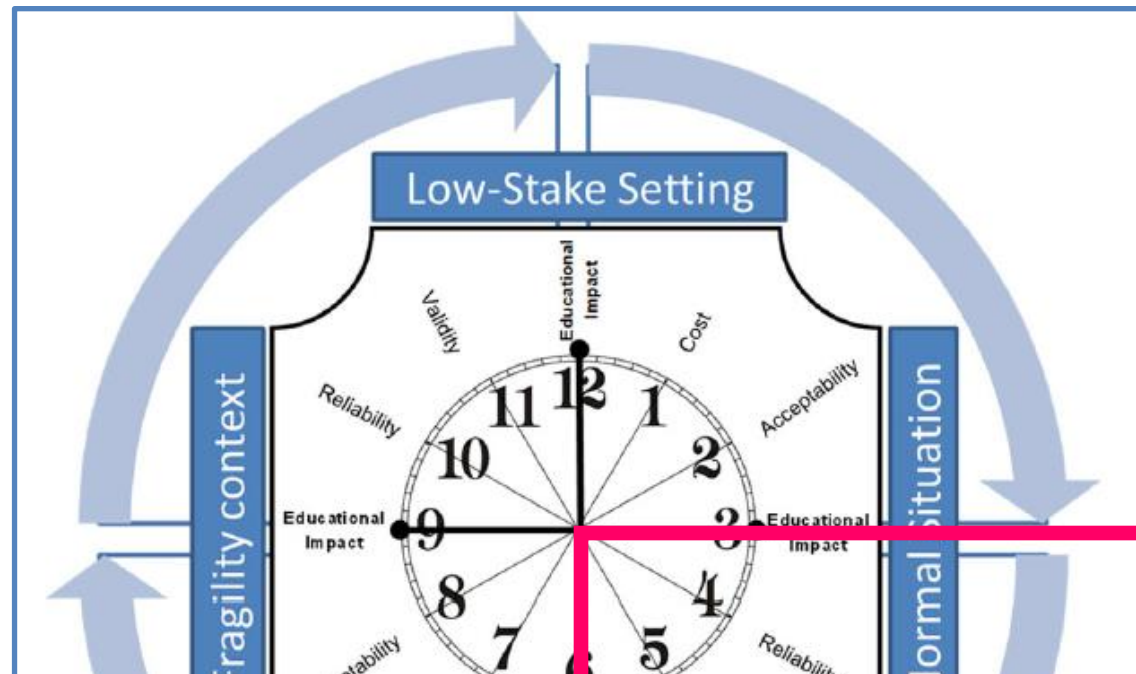
Assessors comments (compulsory) Please describe what was effective, what could be improved and your overall impression. If required, please specify suggested actions for improvement and a timeline.

A CBD form

Capstone projects



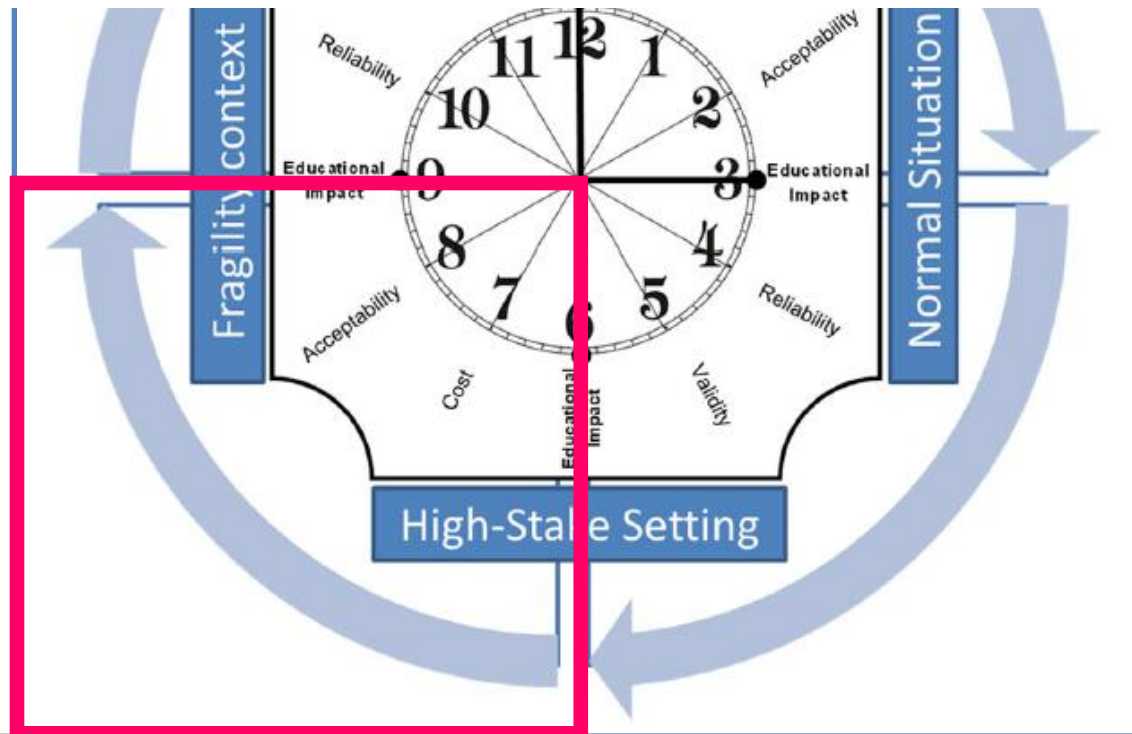
Assessment clock as a model to prioritize students' assessment in fragile contexts



Rigorous assessment process involving detailed **blueprinting**, careful **item development**, **examiner training**, and **robust standard-setting** procedures to ensure fair and defensible decisions.

Assessment clock as a model to prioritize students' assessment in fragile contexts

Remote OSCEs or online structured viva voce examinations



Case Study: Curriculum delivery in medical education during an emergency (COVID-19) – University of Sharjah, United Arab Emirates

“During the COVID-19 pandemic, a structured, institution-wide response to sustain curriculum delivery and assessment was implemented. Faculty were trained in using Blackboard, Zoom, and Microsoft Teams. Teaching activities were redesigned to include e-lectures, e-PBL, virtual patient encounters, and simulation-based sessions. Weekly coordination meetings ensured curriculum consistency and quality assurance. To enhance engagement in theoretical components, the college introduced online Team- Based Learning (TBL) using Blackboard Lockdown Browser, breakout rooms, and Microsoft Forms. Final-year dental students’ summative assessment included MEQs, MCQs, OSCEs, and oral exams, conducted using Blackboard and MS Teams. E-OSCEs were launched using Microsoft Teams. Students remained in assigned virtual rooms while faculty rotated across cases. Targeted manuals were disseminated to guide faculty. Each manual was accompanied by training sessions, video tutorials, and ongoing support to ensure consistent implementation.”

Embedding resilience within assessment

“Heightened importance of well-being in fragile settings – especially in relation to assessment – Wadi et al. introduced the Systematic Assessment for Resilience (SAR) framework. Although designed to embed resilience within assessment during normal conditions, SAR offers practical and actionable guidance that remains highly relevant during periods of disruption”.



Lessons learned from the assessment clock?



Which of the following phrases/statements about student assessment during times of educational fragility are correct?

- 1) Relying on meticulous standards we usually during normal situations
- 2) **Compromise**
- 3) Arbitrary compromise
- 4) **Clear understanding of which assessment aspects can be adjusted without undermining the overall integrity of the system**
- 5) Cost and acceptability at the heart of implementation
- 6) **Blueprinting, careful item development, examiner training, and robust standard-setting procedures in summative exams during fragility**

Summary

- ✓ Fundamentals of student assessment
- ✓ Challenges of virtual formative examinations
- ✓ Actions to buffer challenges of virtual formative student assessment
- ✓ Assessment clock as a model to prioritize students' formative assessment in different contexts

Your summary of the entire session?



Take home message

“Even the most advanced assessment tools may remain unused if assessors or learners feel unprepared or unconvinced – for instance, incorporating highly technical simulation equipment into a final OSCE may be theoretically ideal but practically unfeasible without adequate training and buy-in.”



References



- 1) Mohamed Hassan Taha, Majed Wadi, Abdelrahim Mutwakei Gattar, Esra Abdallah Abdalwahed Mahgoub, Ghaith Alfakhry, David Taylor & Mohamed Elhassan Abdalla (10 Jan 2026): **Building resilient health professions education in fragile contexts: AMEE guide No. 182**, Medical Teacher, DOI: 10.1080/0142159X.2025.2606070
- 2) Dent J, Harden RM, Hunt D. **A practical guide for medical teachers**, e-book. Elsevier health sciences; 2021.
- 3) Jalili Mohammad, Khabaz Mafinejad Mahboobeh, Gandomkar Roghayeh, Mortaz Hejri Sara. **Principals and methods of student assessment in health professions**. Tehran. The Academy of Medical Sciences Islamic Republic of Iran. 1st edition, 1396.
- 4) Al-Eraky M, Marei H. **A fresh look at Miller's pyramid: assessment at the 'Is' and 'Do' levels**. Medical education. 2016;50(12):1253-7.
- 5) Cruess RL, Cruess SR, Steinert Y. **Amending Miller's pyramid to include professional identity formation**. Academic Medicine. 2016 ;91(2):180-5.
- 6) Google images

Appreciate your attention.
Any questions?

