



Constructing Rubrics for Performance-based Assessments

Jordan Holmes, MA, MRT(N), CHSE
Manager, Centre for Learning,
Innovation and Simulation
jholmes@michener.ca @jordanholmes

Objectives

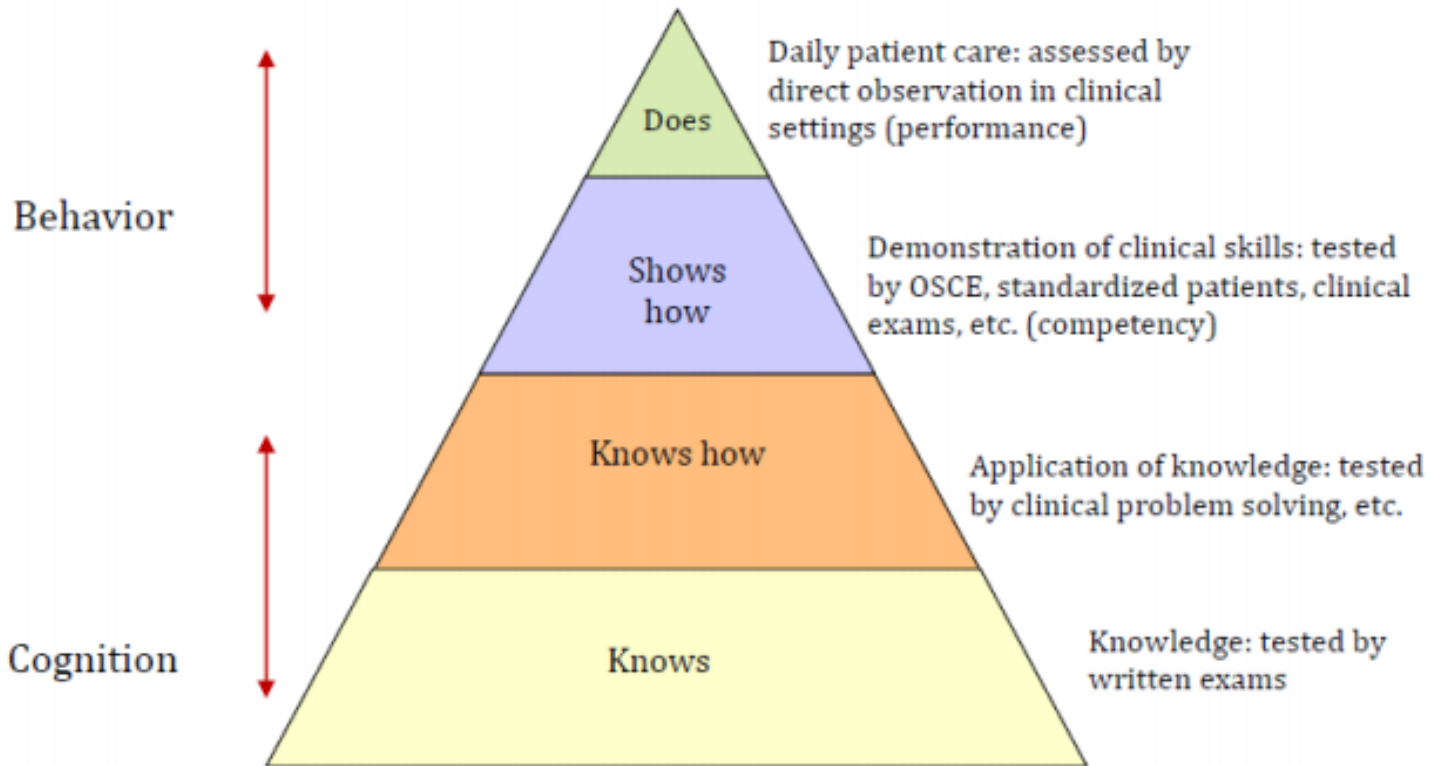
At the end of this presentation, you will be able to:

1. Describe the fundamental principles of performance-based assessment
2. Describe how to construct rubrics that accurately and fairly assess student performance
3. Consider when to implement procedural checklists versus global rating scales
4. Describe how to establish a pass score for a performance-based assessment

What is Performance-based Assessment?

- The appraisal of a learner's proficiency in demonstrating a particular competency or set of competencies
 - May include knowledge, skills, judgement:
Observable criteria
 - Time-bound
- Two general purposes:
 - Formative: assessment *for* learning
 - Summative: assessment *of* learning

Miller's Pyramid of Assessment



Adapted from: Ramani S, Leinster S, AMEE Guide no 34: Teaching in the clinical environment. *Medical Teacher*, 2008;30(4):347-364.

Example – Venipuncture Checklist

	Criterion	Rating
1	Washed hands	Complete/Not Complete
2	Checked that all necessary equipment is available and ready to use	Complete/Not Complete
3	Put on gloves	Complete/Not Complete
4	...	

Checklists – Strengths and Limitations

Strengths

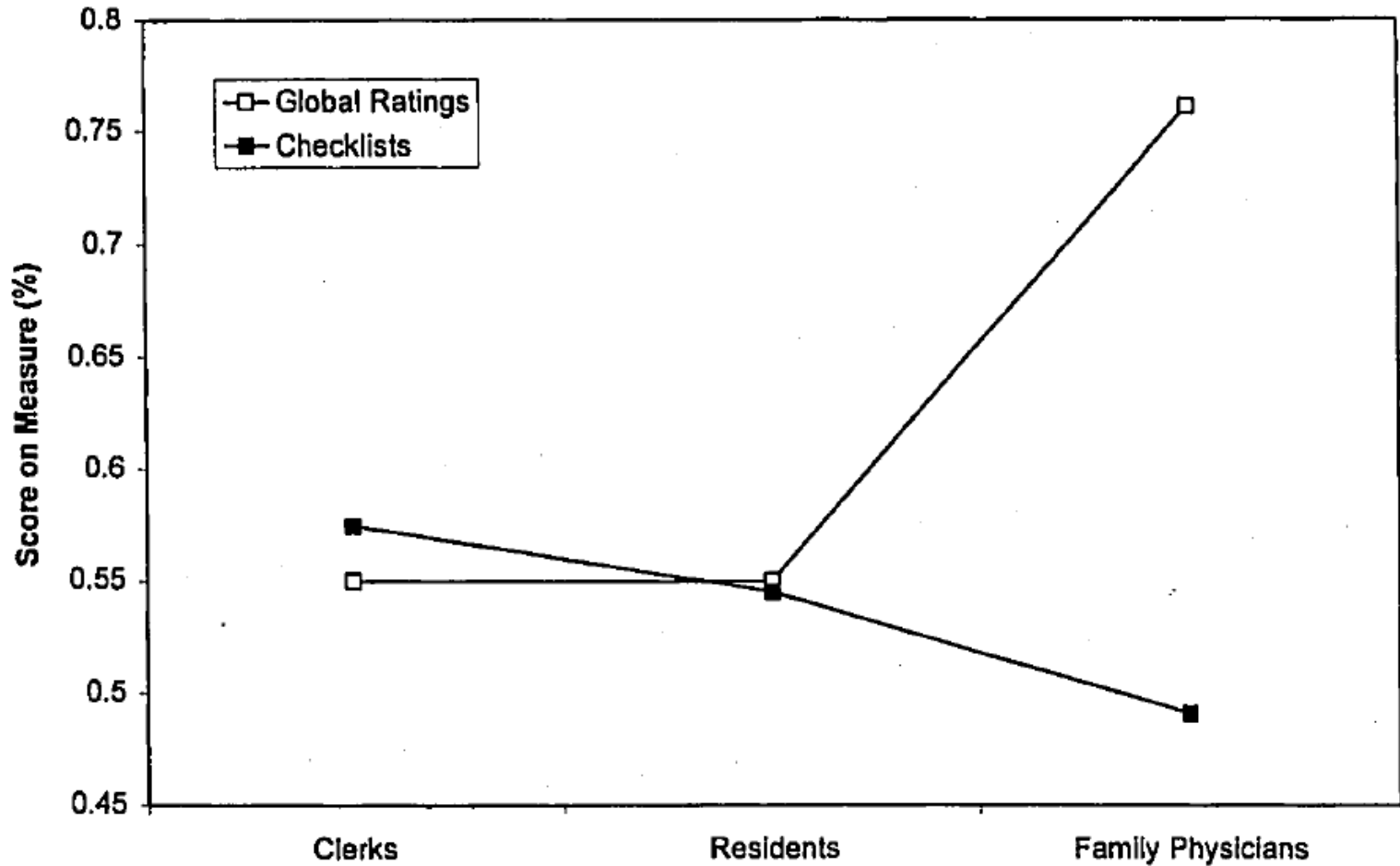
- “Recipe” – easy for rater to follow, intuitive
- Rater can be less familiar with clinical task
- Provides template for formative feedback

Limitations

- Binary (e.g., complete/not complete) ratings may result in a loss of information
- **May reward thoroughness at the expense of expertise**

Ilgen, J. S., Ma, I. W. Y., Hatala, R., & Cook, D. A. (2015). A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Medical Education*, 49(2), 161–173. <https://doi.org/10.1111/medu.12621>

Do Checklists Capture Expertise?



Hodges, B., Regehr, G., McNaughton, N., Tiberius, R., & Hanson, M. (1999). OSCE checklists do not capture increasing levels of expertise. *Academic Medicine*, 74(10), 1129–1134.

Global Rating Scales (GRSs)

GLOBAL RATING SCALE OF OPERATIVE PERFORMANCE

Please circle the number corresponding to the candidate's performance in each category, irrespective of training level

Respect for Tissue :

1	2	3	4	5
Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissue appropriately with minimal damage

Time and Motion :

1	2	3	4	5
Many unnecessary moves		Efficient time/motion but some unnecessary moves		Clear economy of movement and maximum efficiency

Instrument Handling :

1	2	3	4	5
Repeatedly makes tentative or awkward moves with instruments by inappropriate use of instruments		Competent use of instruments but occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness

Knowledge of Instruments :

1	2	3	4	5
Frequently asked for wrong instrument or used inappropriate instrument		Knew names of most instruments and used appropriate instrument		Obviously familiar with the instruments and their names

Martin, J. A., Regehr, G., Reznick, R., Macrae, H., Murnaghan, J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery*, 84(2), 273–278.

<https://doi.org/10.1002/bjs.1800840237>

Checklists vs. GRSs

- A recent systematic review (Ilgen et al., 2015) examined the validity evidence of checklists vs. GRSs in simulation-based assessment of health professionals:
 - Inter-rater reliability similar between checklists and GRSs
 - Inter-item and inter-station reliability higher for GRSs
 - Separate tasks require separate checklists – single GRS can be used across tasks
 - GRSs may better capture nuanced elements of expertise
 - Checklists appropriate for well-defined procedures with clear best-practice (e.g., venipuncture)

GRS Construction

- Where possible, start with published scales in the literature that have demonstrated to pass basic psychometric properties (ANTS, BARS, OSATS, OCAT, mini-CEX, etc.)
- Scale should be a collaboration – developed and validated by multiple content experts
- Well defined anchors on rating scale (e.g., what does “exceeds expectations” vs. “meets expectations” look like for each criterion?)

Setting a Pass Score

- Most frequent method for high-stakes:
 - Angoff: expert judges indicate probability that a “borderline” student would accomplish each item – total test pass score set to average of item probabilities
- Examinees are allowed to miss a fixed percentage of items, **regardless of item importance...**
- Is this a defensible practice in a competency-based program?

Safety-anchored Pass Score

- Yudkowsky et al. (2014): identify what procedural items are **essential** to patient/clinician safety or the outcome of the procedure
- Set a **conjunctive** pass score, i.e., all essential items plus a percentage of non-essential items
- Students challenge exam as many times as it takes until they are able to demonstrate mastery – deliberate practice and feedback in-between

Questions?

jholmes@michener.ca

@jordanolmes on Twitter

References

Hodges, B., Regehr, G., McNaughton, N., Tiberius, R., & Hanson, M. (1999). OSCE checklists do not capture increasing levels of expertise. *Academic Medicine*, 74(10), 1129–1134.

Ilgen, J. S., Ma, I. W. Y., Hatala, R., & Cook, D. A. (2015). A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Medical Education*, 49(2), 161–173. <https://doi.org/10.1111/medu.12621>

Martin, J. A., Regehr, G., Reznick, R., Macrae, H., Murnaghan, J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery*, 84(2), 273–278. <https://doi.org/10.1002/bjs.1800840237>

Yudkowsky, R., Tumuluru, S., Casey, P., Herlich, N., & Ledonne, C. (2014). A patient safety approach to setting pass/fail standards for basic procedural skills checklists. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 9(5), 277–282. <https://doi.org/10.1097/SIH.0000000000000044>