Making the Milestones Work for You: Training your faculty to become better evaluators

Shalini Reddy MD
Shanu Gupta MD
Janet Riddle MD

Disclosures

- We will not be discussing off-label use of any medications in this presentation

Objectives

By the end of this one hour workshop, participants will be able to:

- Discuss how the milestones can be used for rater training

- Participate in a demonstration of two types of rater training

- Formulate a draft of a plan to implement rater training at their home institution

Format

- Introductory didactic
- Small groups training exercises
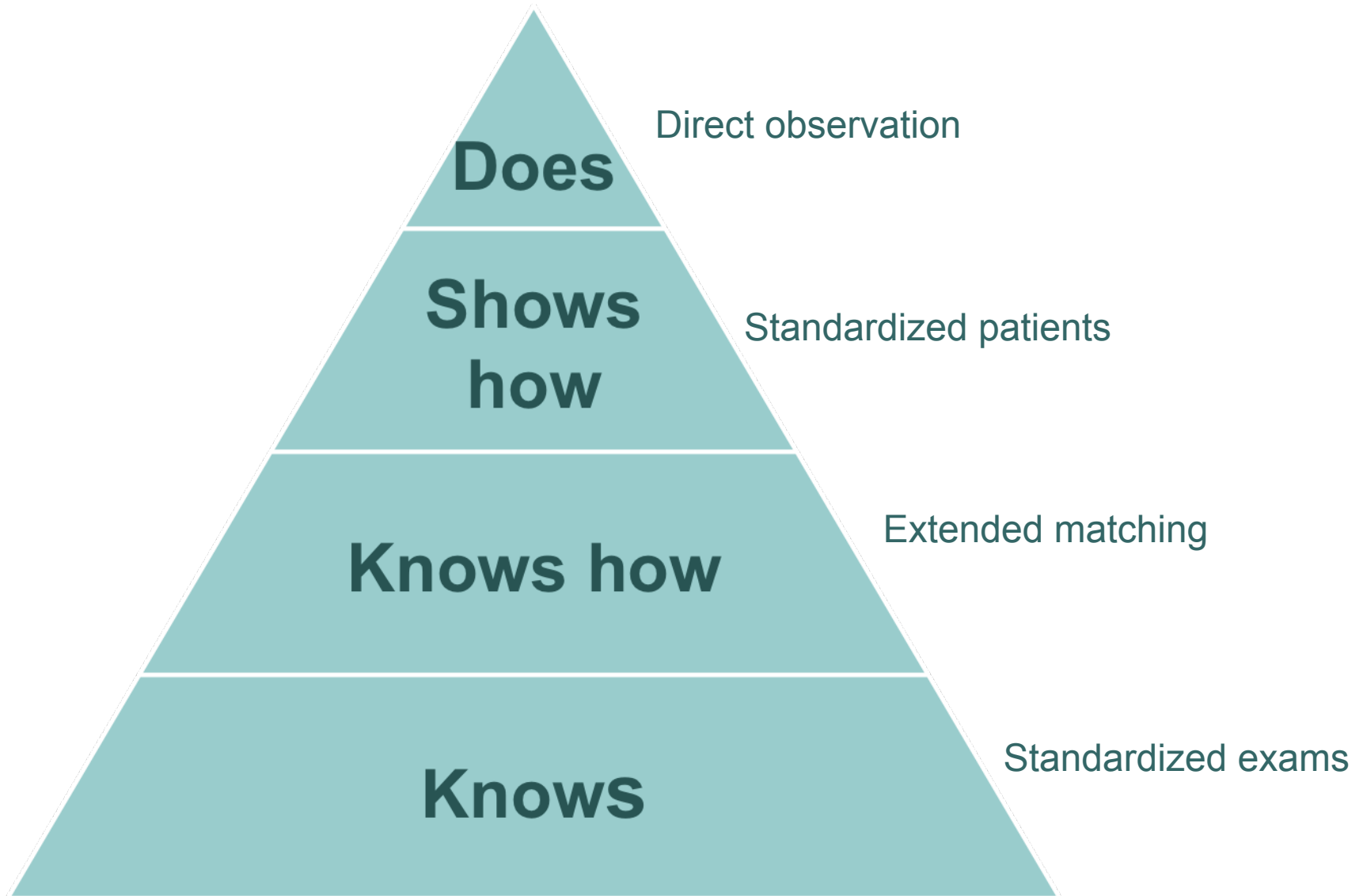- Large group discussion

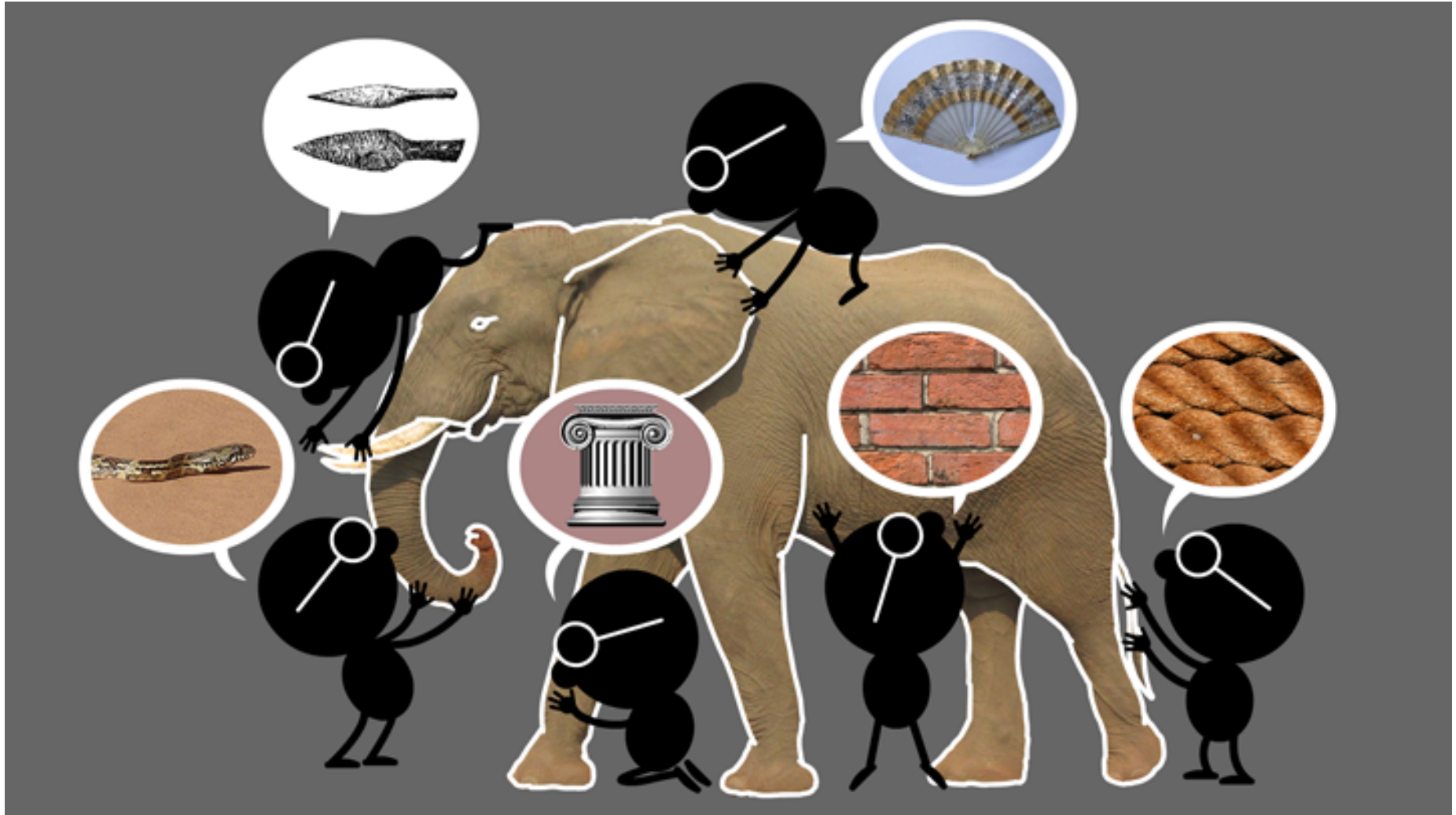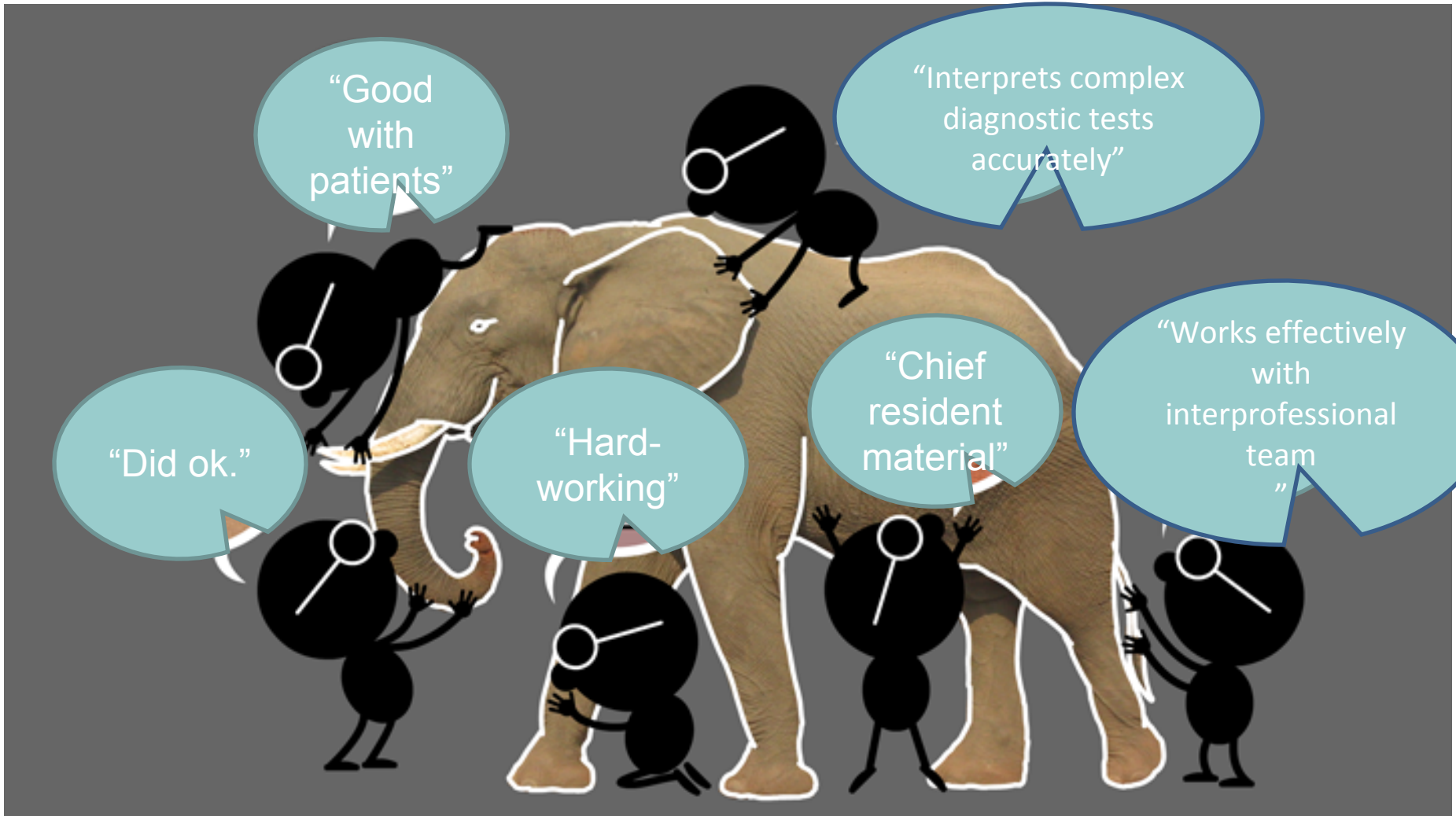Why is it important to train your raters?

Improving the Quality of your Evaluations

- Improve your evaluation forms
- Improve the raters
- Remove the raters

Recommendations

- Educate raters

- Establish the meaning of ratings

- Make promotion and grading decisions via a faculty group review

- Give raters feedback about stringency and leniency

McGaghie et al, 2008

Rater Training

- Performance dimension

- Frame of reference

- Direct observation of competence

- Training in behavioral observation

Systematic Rater Errors

- Examiner bias

- Halo effect
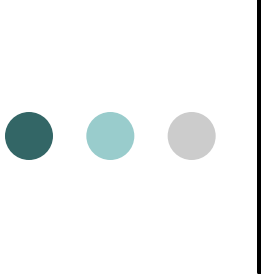
- Recency effect

- Stringency bias

- Central tendency error

Holmboe, 2008

Rater Error Training

- Familiarize raters with common errors
- Reduces rater errors
- But…mixed effects on rating accuracy

Feldman M, Lazzara EH, Vanderbilt AA, Diaz Granados D; "Rater Training to Support High-Stakes Simulation-Based Assessments. J. Cont Ed in Health Prof. 32(4):279-286, 2012
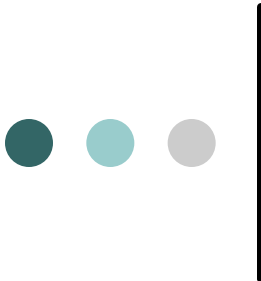
Common Rater Errors

- Halo/horn effect
- Central tendency
- Leniency
- Stringency

Feldman M, Lazzara EH, Vanderbilt AA, Diaz Granados D; "Rater Training to Support High-Stakes Simulation-Based Assessments. J. Cont Ed in Health Prof. 32(4):279-286, 2012

Halo Effect: "They're good at X so must be good at everything."

- Patient care   1 2 3 4 5 6 7 8 **9**
- Med knowledge   1 2 3 4 5 6 7 8 **9**
- SBP   1 2 3 4 5 6 7 **8** 9
- PBLI   1 2 3 4 5 6 7 **8** 9
- Prof   1 2 3 4 5 6 7 8 **9**
- ICS   1 2 3 4 5 6 7 8 **9**

Central Tendency: "Everyone is Average"

- Patient care   1 2 3 4 5 ⑥ 7 8 9
- Med knowledge   1 2 3 4 5 ⑥ 7 8 9
- SBP   1 2 3 4 5 ⑥ 7 8 9
- PBLI   1 2 3 4 5 ⑥ 7 8 9
- Prof   1 2 3 4 5 ⑥ 7 8 9
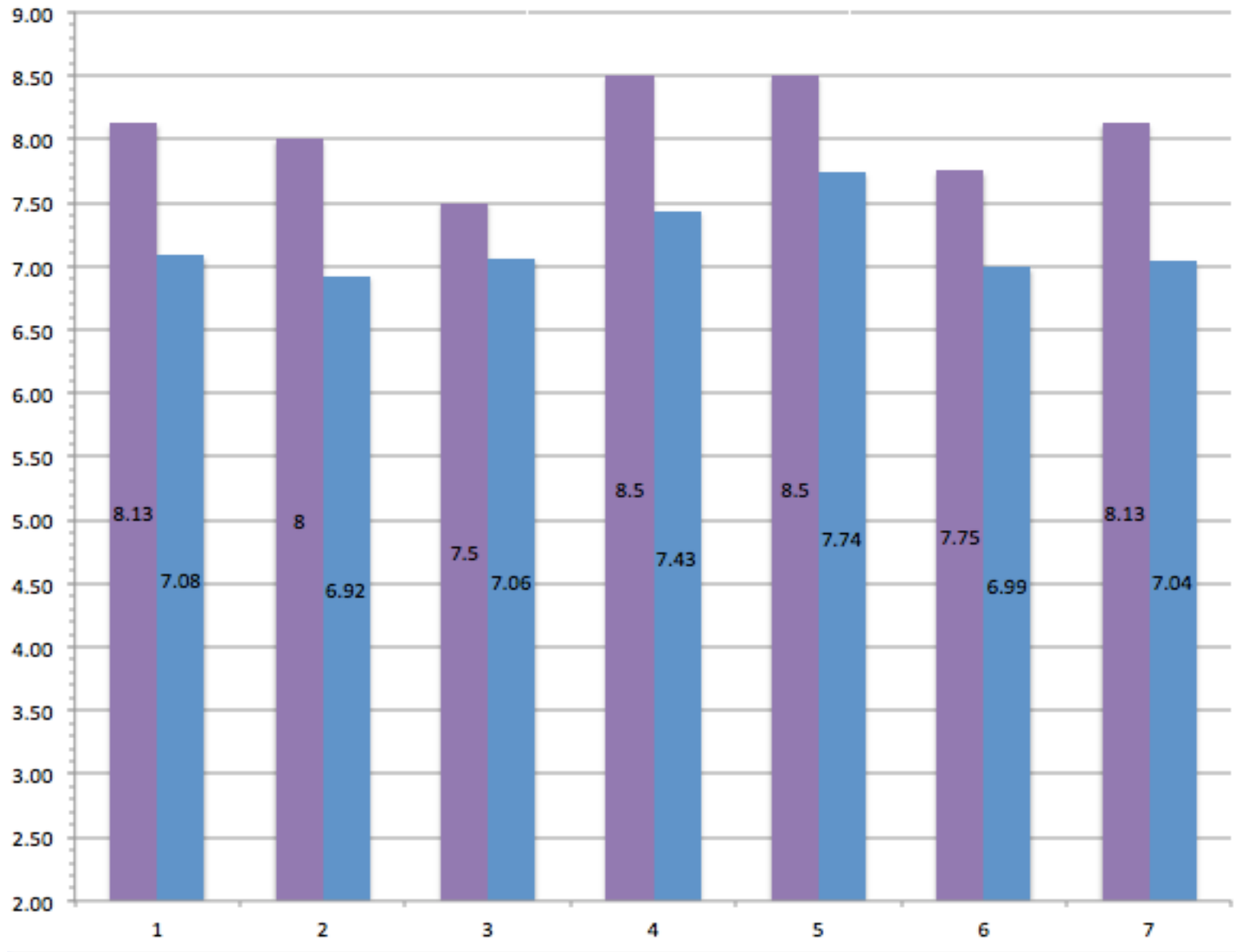- ICS   1 2 3 4 5 ⑥ 7 8 9

Leniency: "Everyone is Above Average"

- Patient care   1 2 3 4 5 6 7 **8** 9
- Med knowledge   1 2 3 4 5 6 7 **8** 9
- SBP   1 2 3 4 5 6 7 **8** 9
- PBLI   1 2 3 4 5 6 7 8 **9**
- Prof   1 2 3 4 5 6 7 8 **9**
- ICS   1 2 3 4 5 6 7 **8** 9

Stringency: "What's wrong with our selection committee?"

- Patient care   1 2 3 ④ 5 6 7 8 9
- Med knowledge   1 2 3 4 ⑤ 6 7 8 9
- SBP   1 2 3 4 ⑤ 6 7 8 9
- PBLI   1 2 3 ④ 5 6 7 8 9
- Prof   1 2 3 4 5 ⑥ 7 8 9
- ICS   1 2 3 4 ⑤ 6 7 8 9

DME Collaborative for Active Learning in
Medicine

Attending Ratings of Learners on GMF 2013-2014
N = 17

Legend:
- PC
- MK
- PBLI
- ICS
- PROF
- SBP
- OVERALL
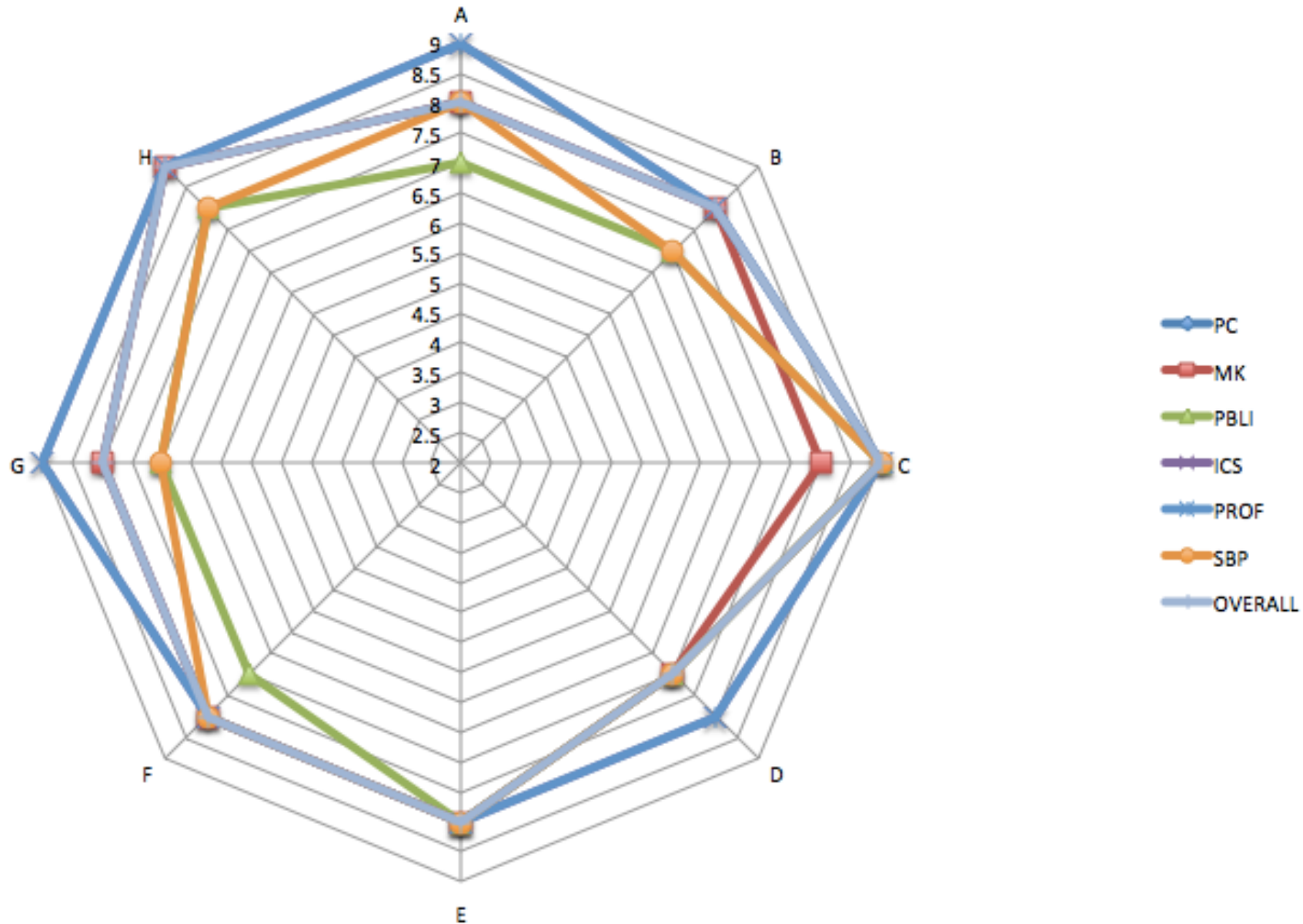
Performance Dimension Training

- Review definitions and criteria for each dimension of performance

Performance Dimension Training

- What constitutes competence in...?
- Raters define the dimensions of a competence
- Teach raters a common language

Milestones can be the Language

20. Communicates effectively with patients and caregivers. (ICS1)				
Critical Deficiencies			Ready for unsupervised practice	**Aspirational**
Ignores patient preferences for plan of care

Makes no attempt to engage patient in shared decision-making

Routinely engages in antagonistic or counter-therapeutic relationships with patients and caregivers | Engages patients in discussions of care plans and respects patient preferences when offered by the patient, but does not actively solicit preferences.

Attempts to develop therapeutic relationships with patients and caregivers but is often unsuccessful

Defers difficult or ambiguous conversations to others | Engages patients in shared decision making in uncomplicated conversations

Requires assistance facilitating discussions in difficult or ambiguous conversations

Requires guidance or assistance to engage in communication with persons of different socioeconomic and cultural backgrounds | Identifies and incorporates patient preference in shared decision making across a wide variety of patient care conversations

Quickly establishes a therapeutic relationship with patients and caregivers, including persons of different socioeconomic and cultural backgrounds

Incorporates patient-specific preferences into plan of care | Role models effective communication and development of therapeutic relationships in both routine and challenging situations

Models cross-cultural communication and establishes therapeutic relationships with persons of diverse socioeconomic backgrounds |
| ☐ ☐ | ☐ ☐ | ☐ ☐ | ☐ ☐ | ☐ ☐ |
| Comments: | | | | |

INTERPERSONAL AND COMMUNICATION SKILLS

Communicates effectively with patients and caregivers. (ICS1)					
Patient preferences	Ignores patient preferences for plan of care	Engages patients in discussions of care plans and respects patient preferences when offered by the patient, but does not actively solicit preferences	Engages patients in shared decision making in uncomplicated conversations	Identifies and incorporates patient preference in shared decision making across a wide variety of patient care conversations	Role models effective communication and development of therapeutic relationships in both routine and challenging situations
Therapeutic relationship	Routinely engages in antagonistic or counter-therapeutic relationships with patients and caregivers	Attempts to develop therapeutic relationships with patients and caregivers but is often unsuccessful	Requires guidance or assistance to engage in communication with persons of different socioeconomic and cultural backgrounds	Quickly establishes a therapeutic relationship with patients and caregivers, including persons of different socioeconomic and cultural backgrounds	Models cross-cultural communication and establishes therapeutic relationships with persons of diverse socioeconomic backgrounds
Difficult conversations	Makes no attempt to engage patient in shared decision-making	Defers difficult or ambiguous conversations to others	Requires assistance facilitating discussions in difficult or ambiguous conversations		
				Incorporates patient-specific preferences into plan of care	
Communicates effectively in interprofessional teams					
team communication	Utilizes communication strategies that hamper collaboration and teamwork	Uses unidirectional communication that fails to utilize the wisdom of the team	Inconsistently engages in collaborative communication with appropriate members of the team	Consistently and actively engages in collaborative communication with all members of the team	Role models and teaches collaborative communication with the team to enhance patient care, even in challenging settings and with conflicting team member opinions
Collaboration	Verbal and/or non-verbal behaviors disrupt effective collaboration with team members	Resists offers of collaborative input	Inconsistently employs verbal, non-verbal, and written communication strategies that facilitate collaborative care	Verbal, non-verbal and written communication consistently acts to facilitate collaboration with the team to enhance patient care	

Frame of Reference Training

- Achieve consistency among faculty in applying criteria – to distinguish levels of performance

*
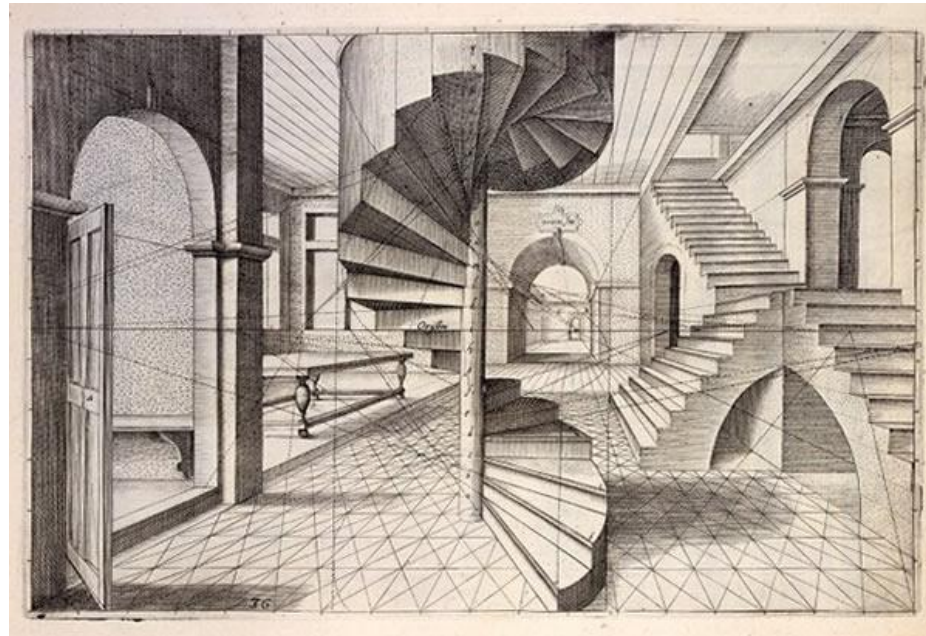
Frame of Reference Training

- Goal: discriminate between variations in performance

- What does a "level 1" resident look like? a "level 2"?

Holmboe, E. S., Hawkins, R. E., & Huot, S. J. (2004). Effects of Training in Direct Observation of Medical Residents' Clinical CompetenceA Randomized Trial. *Annals of Internal Medicine*, *140*(11), 874-881.

20. Communicates effectively with patients and caregivers. (ICS1)				
Critical Deficiencies			**Ready for unsupervised practice**	**Aspirational**
Ignores patient preferences for plan of care Makes no attempt to engage patient in shared decision-making Routinely engages in antagonistic or counter-therapeutic relationships with patients and caregivers	Engages patients in discussions of care plans and respects patient preferences when offered by the patient, but does not actively solicit preferences. Attempts to develop therapeutic relationships with patients and caregivers but is often unsuccessful Defers difficult or ambiguous conversations to others	Engages patients in shared decision making in uncomplicated conversations Requires assistance facilitating discussions in difficult or ambiguous conversations Requires guidance or assistance to engage in communication with persons of different socioeconomic and cultural backgrounds	Identifies and incorporates patient preference in shared decision making across a wide variety of patient care conversations Quickly establishes a therapeutic relationship with patients and caregivers, including persons of different socioeconomic and cultural backgrounds Incorporates patient-specific preferences into plan of care	Role models effective communication and development of therapeutic relationships in both routine and challenging situations Models cross-cultural communication and establishes therapeutic relationships with persons of diverse socioeconomic backgrounds
☐	☐ ☐	☐ ☐	☐ ☐	☐ ☐
Comments:				

20

Training in the Art of Observation

- Increase number of observations

- Provide a means for recording observations

- Prepare for observations
    - Purpose
    - Positioning
    - Minimize interruptions

Holmboe, 2008

Review: What is Rater Training?

- Rater error training
- Performance dimension training
- Frame of reference training
- Training in the art of observation

Creating a Curriculum for Rater Training

1. Problem identification and general needs assessment
2. Targeted needs assessment
3. Goals & objectives
4. Educational strategies
5. Implementation
6. Evaluation

Kern, D. E., Thomas, P. A., & Hughes, M. T. (Eds.). (2010). *Curriculum development for medical education: a six-step approach*. JHU Press.

Steps 1, 2 & 3

1. General needs assessment:
   - Ratings are inaccurate and unreliable
2. Targeted needs assessment
   - What do you wish *your* evaluators did better?
3. Goals and objectives
   - What do you hope people will do better as a result of your curriculum?

4. Educational Strategies

- Rater training is effective
- Four commonly used strategies for training raters
  - Rater error training
  - Performance dimension training
  - Frame of reference training
  - Behavioral observation training

Feldman M, Lazzara EH, Vanderbilt AA, DiazGranados D (2012). "Rater Training to Support High-Stakes Simulation-Based Assessments. J. Cont Ed in Health Prof. 32(4), 279-286.

Roch, S. G., Woehr, D. J., Mishra, V., & Kieszczynska, U. (2012). Rater training revisited: An updated meta-analytic review of frame-of-reference training. *Journal of Occupational and Organizational Psychology*, 85(2), 370-395.

Hedge, J. W., & Kavanagh, M. J. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. *Journal of Applied Psychology*, 73(1), 68.

5. Implementation

- Return to your group:
- How will you use this at your home institution?
- What are the resources that are available?
- What barriers do you anticipate encountering?

Take Home Points

- Rater training is *at least* as important as improving your forms
- Rater training works
- Follow Kern's 6 Steps for Curriculum Development
- Let your goals and objectives guide your educational strategy
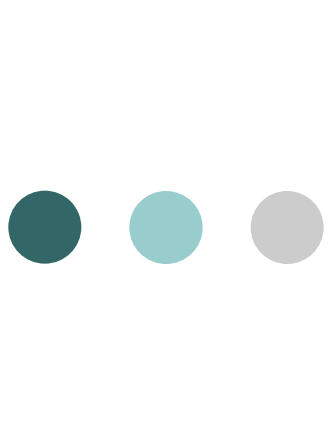
Your Turn…

Observational Assessment – References

Cook DA, Dupras DM, Beckman TJ, Thomas KG, Pankratz VS. Effect of rater training on reliability and accuracy of mini-CEX scores: a randomized, controlled trial. J Gen Intern Med. 2008;24:74-79.

Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. Med Educ. 2012;46:28-37.

Dudek N, Dojeiji S. Twelve tips for completing quality in-training evaluation reports. Med Teach. 2014 Jul 2:1-5 [Epub ahead of print].

Dudek NL, Marks MB, Woods TJ, Dojeiji S, Bandiera G, Hatala R, Cooke L, Sadownik L. Quality evaluation reports: can a faculty development program make a difference? Med Teach. 2012;34:e725-731.

Gingerich A, Regehr G, Eva KW. Rater-based assessments as social judgments: rethinking the etiology of rater errors. Acad Med. 2011;86:S1-S7.

Ginsburg S, Eva K, Regehr G. Do in-training evaluation reports deserve their bad reputations? A study of the reliability and predictive ability of ITER scores and narrative comments. Acad Med. 2013;88:1539-1544.

Hauer KE, Holmboe ES, Kogan JR. Twelve tips for implementing tools for direct observation of medical trainees' clinical skills during patient encounters. Med Teach. 2011;33:27-33.

Holmboe ES, Direct observation by faculty. In Holmboe ES, Hawkins RE (eds). *Practical Guide to the Evaluation of Clinical Competence.* Philadelphia: Mosby Elsevier, 2008.

Holmboe ES, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence. A randomized trial. Ann Intern Med. 2004;140:874-881.

Kogan JR, Conforti L, Bernabeo E, Iobst W, Holmboe E. Opening the black box of clinical skills assessment via observation: a conceptual model. Med Educ. 2011;45:1048-1060.

McGaghie WC, Butter J, Kaye M. Observational assessment. In Downing S, Yudkowsky R (eds). *Assessment in Health Professions Education.* New York & London: Routledge, 2009.

Park YS, Riddle J, Tekian A. Validity evidence of resident competency ratings and the identification of problem residents. Med Educ. 2014;48:614-622.

Regehr G, Ginsburg S, Herold J, Hatala R, Eva K, Oulanova O. Using "standardized narratives" to explore new ways to represent faculty opinions of resident performance. Acad Med. 2012;87:419-427.

Thomas MR, Beckman TJ, Mauck KF, Cha SS, Thomas KG. Group assessments of resident physicians improve reliability and decrease halo error. J Gen Intern Med. 2011;26:759-764.

Williams RG, Klamen DA, McGaghie WC. Cognitive, social, and environmental sources of bias in clinical performance ratings. Teach Learn Med. 2003;15:270-292.

Williams RG, Dunnington GL, Klamen DL. Forecasting residents' performance – partly cloudy. Acad Med. 2005;80:415-422.