



# **SAU WINTER MEETING** **FEBRUARY 6-7, 2026**

Houston, Texas • Marriot Marquis Houston

## ***Crossfire Debate***

**Are Milestones Measuring Progress or Missing the Mark?  
Building a Better Assessment Toolbox**

**Moderator: Christopher Jaeger, MD, MMSc**

**Speakers: Sydney McLean, MHA, & Tasha Posid, MA, PhD**

# FINANCIAL DISCLOSURE

We do not have any relationships to report with ACCME defined ineligible companies.

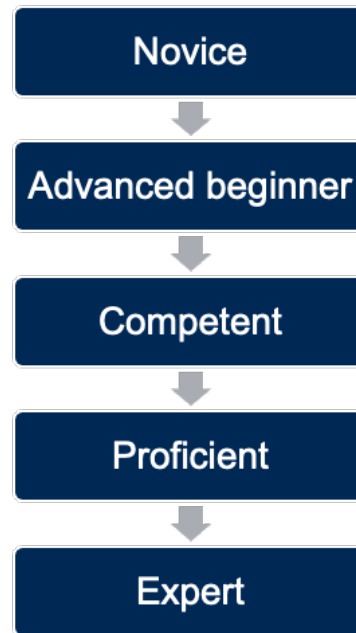
We will not be discussing unlabeled/investigational uses of medical devices or pharmaceuticals during this presentation.

# Introduction:

## *Are Milestones Measuring Progress or Missing the Mark?*

- Milestones designed to support learner development and guide curriculum design
- Milestones in Urology have been far from perfect (Sebesta et al., *Urology*, 2019)
- Great in theory... the question is whether they are *reliable, valid, and useful?*
- DEBATE!

*Disclaimer: this debate is not abolish vs. defend milestones*



5 Stage Model for Skill Acquisition  
Dreyfus & Dreyfus, 1986.

# Introduction: *Let's Debate!*

## Topics for Debate!

- (1) Good for Early Signal Detection and Standardization or not?
- (1) Developmental Tracking or Illusion of Precision?
- (1) Usefulness for Feedback, Remediation, and Excellence or not?



# Crossfire Debate



**Tasha  
Posid,  
MA, PhD**

Education  
Specialist



**Sydney  
McLean,  
MHA**

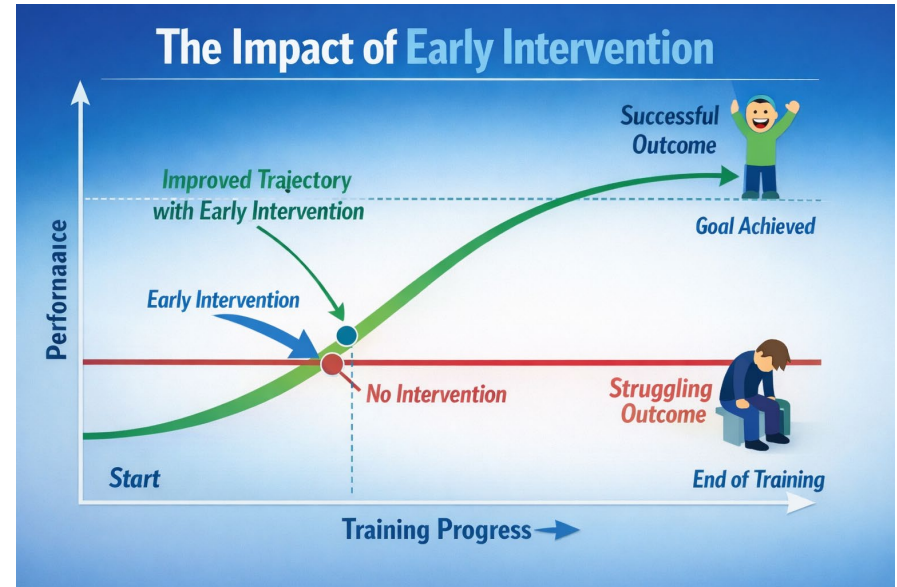
Director of  
Milestones  
Development



# Debate Topic 1:

## *Early Signal Detection and Standardization*

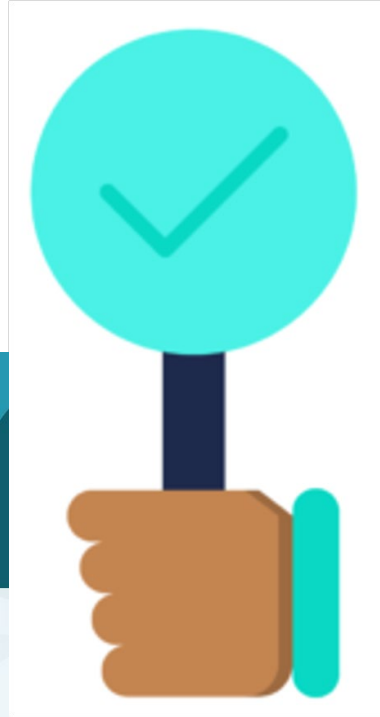
- Milestones track development over time
- Identifying struggling learners is always key
- Early remediation can change trajectory of struggling learners
- Can the Milestones really aid in signal detection?



# Debate Topic 1:

## *Early Signal Detection and Standardization*

- Standardized language → consistent identifications of development concerns
- Built by content experts using theoretical frameworks
  - Validity and shared understanding
- Longitudinal trends for early performance signals



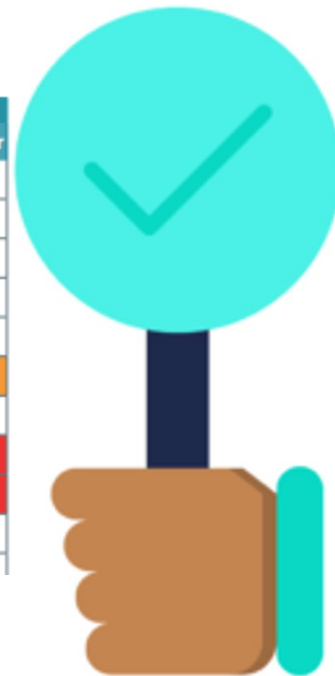
# Debate Topic 1: *Early Signal Detection and Standardization*

## PPV Tables By Specialty/Subspecialty

SPECIALTY: Urology

Subcompetency	Threshold	Yr1, Mid-Year	Yr1, Year-End	Yr2, Mid-Year	Yr2, Year-End	Yr3, Mid-Year	Yr3, Year-End	Yr4, Mid-Year	Yr4, Year-End	Yr5, Mid-Year
<b>SBP01</b>										
	≤ Lev 5.0		10.0%					9.6%	9.7%	9.9%
	≤ Lev 4.5		10.0%		9.8%	9.9%	9.9%	9.7%	9.8%	11.1%
	≤ Lev 4.0		10.0%		9.8%	9.9%	9.9%	10.1%	11.1%	15.8%
	≤ Lev 3.5	9.8%	10.0%	9.8%	10.0%	10.2%	11.2%	16.0%	19.4%	39.6%
	≤ Lev 3.0	9.6%	10.2%	10.1%	9.7%	12.0%	14.0%	30.4%	30.2%	69.0%
	≤ Lev 2.5	9.9%	8.8%	10.0%	9.6%	17.0%	22.6%	36.8%	84.0%	
	≤ Lev 2.0	9.6%	9.3%	12.6%	11.2%	30.9%		76.9%	40.6%	99.0%
	≤ Lev 1.5	9.6%	8.8%	17.9%	19.6%					99.0%
	≤ Lev 1.0	7.7%	9.2%	18.5%						

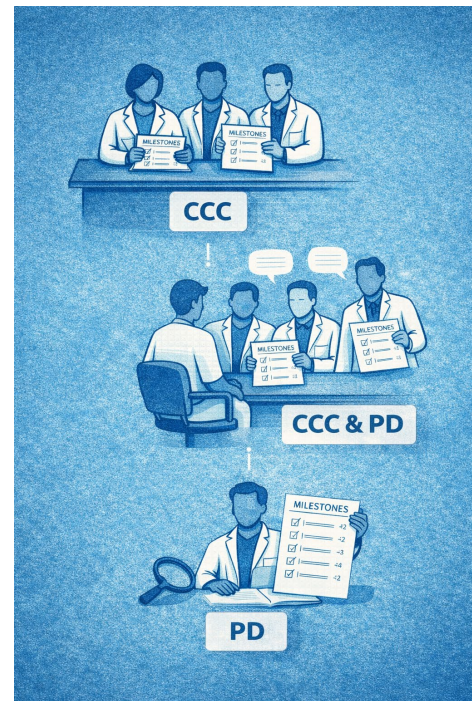
***See patterns. Track progress. Intervene early.***



# Debate Topic 1: *Early Signal Detection and Standardization*

***“Standardized” framework ≠ standardized use***

- Signal detection is only as good as how the milestones are used, and use varies widely
  - *Inconsistent / no faculty development*
  - *Implementation*
  - *Review / feedback*
- If direct observation and evidence aren't feeding the system, the signal is inherently weak



# Debate Topic 1:

## *Early Signal Detection and Standardization*

### *When do the Milestones Fail?*

- Milestones often treated as an output, not a process
  - *Faculty aren't trained to use them*
  - *Data aren't systematically collected*
  - *Ratings substitute for observation*
- By the time a concern appears in milestones, it's often late, not early

***Milestones don't fail because the framework is flawed***



## Debate Topic 2: *Developmental Tracking vs. Illusion of Precision*

- Levels 1-5 imply objectivity, but are they really that objective?
- Signal vs. noise?
- Meaningfully track development or just an oversimplification?



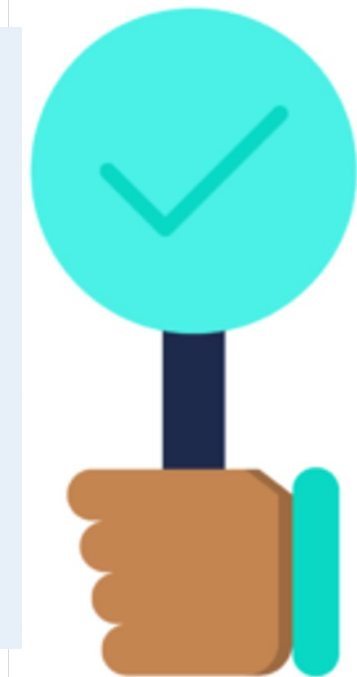
## Debate Topic 2: *Developmental Tracking vs. Illusion of Precision*

### **Milestones for Program-Level Insight**

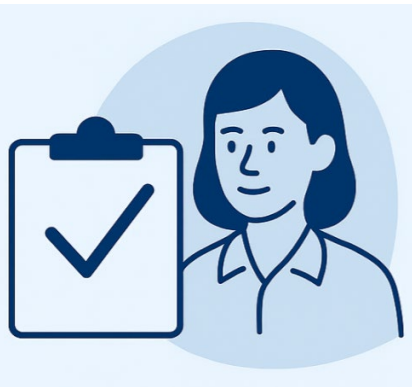
*Well-suited for global, longitudinal tracking across the program.*

### **Value for PDs & CCCs**

*Helps visualize learner growth trajectories over time.*



## Debate Topic 2: *Developmental Tracking vs. Illusion of Precision*



### **Importance of High-Quality Inputs**

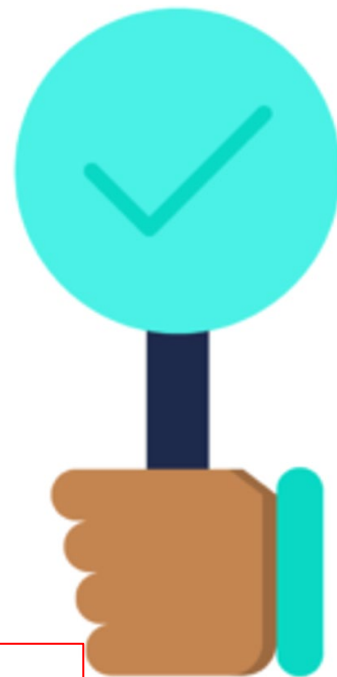
*Meaningful, consistent assessments make Milestones more actionable.*

### **Understanding Learner Development**

*Provides a broad view of where a resident sits on their developmental continuum (e.g., overall progression).*



***Track development, not decimal places.***



## Debate Topic 2: *Developmental Tracking vs. Illusion of Precision*

*It takes more than Milestones to be a doctor*

- Milestones imply precision that doesn't exist

**Milestones provide:**

- ✗ Broad anchors
- ✗ Wide interpretation
- ✗ Clustered ratings

**Residents actually receive:**

- ✓ Faculty comments
- ✓ Coaching conversations
- ✓ Direct observation

- High-quality input is required, but inconsistently available



## Debate Topic 2: *Developmental Tracking vs. Illusion of Precision*

### *Learners don't interact with Milestones day-to-day*

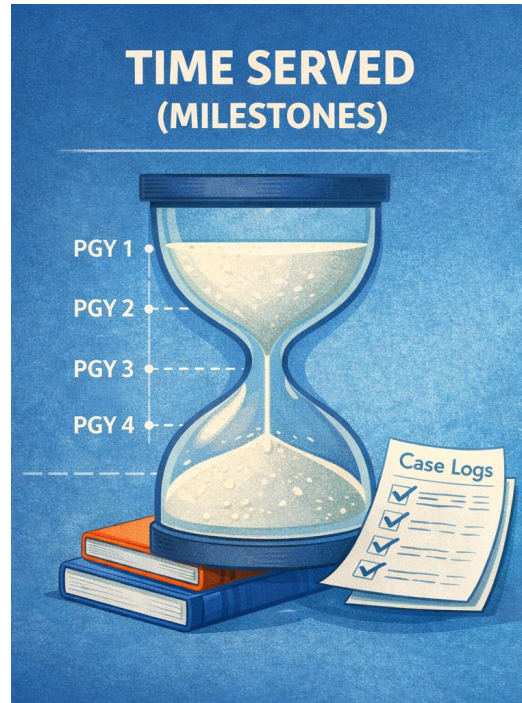
- Limited value for real-time, actionable feedback
- ACGME fails when competence becomes a **checkbox** instead of a **trajectory**
- Milestones are useful for global developmental tracking, but they were never designed to:
  - *Drive daily feedback*
  - *Guide moment-to-moment improvement*

***Milestones summarize performance — they don't generate it.***



# Debate Topic 2: *Developmental Tracking vs. Illusion of Precision*

*If Milestones don't teach, what does?*



# Debate Topic 3:

## *Usefulness for Feedback, Remediation, and Excellence*

- Feedback is the **engine** behind learning in surgical education
- Are Milestones a practical tool to use for feedback?
- Can Milestones be used in remediation?
- Can Milestones be used to promote/identify excellence or is it just a race to the minimum?



---

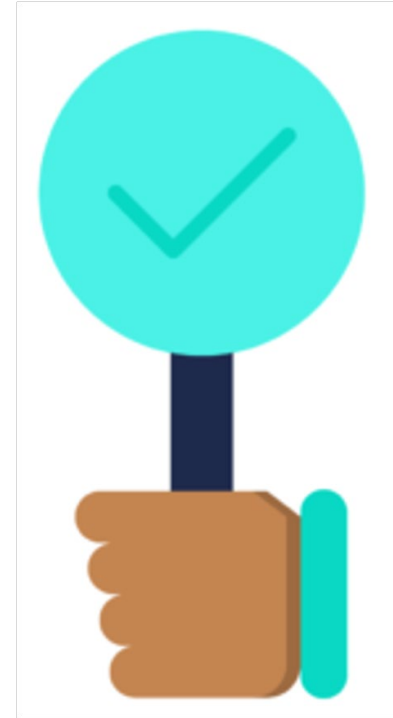
Thank you Dr. Kraft for this image that I keep using in my talks!

# Debate Topic 3:

## *Usefulness for Feedback, Remediation, and Excellence*

### How Milestones Strengthen Programs

- *Identify gaps in curriculum or rotation structure.*
- *Help programs audit whether training opportunities align with intended expectations.*
- *Reveal **system-level issues**, not just learner-level concerns.*



# Debate Topic 3:

## *Usefulness for Feedback, Remediation, and Excellence*

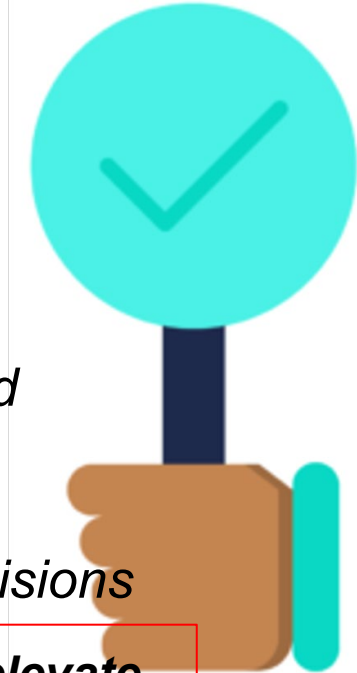
### Recognizing Excellence

- *Level 5 offers a clear framework for identifying and describing exceptional performance.*

### Supporting Remediation

- *Provide a structured, competency-based foundation for developing targeted remediation plans*
- *Defensible structure for high stakes decisions*

***Spot issues early, address them thoughtfully, elevate outstanding performance***



## Debate Topic 3:

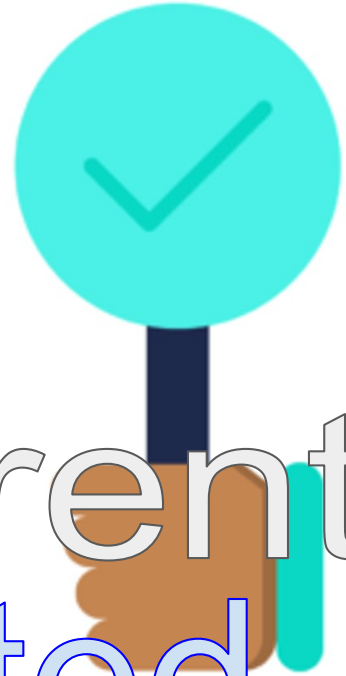
### *Usefulness for Feedback, Remediation, and Excellence*

Practical

Scalable

Transparent

Real-World Tested



# Debate Topic 3: *Usefulness for Feedback, Remediation, and Excellence*

*If Milestones don't teach, what does?*

*Challenge: Learning happens everywhere!*



# Debate Topic 3: *Usefulness for Feedback, Remediation, and Excellence*

*If Milestones don't teach, what does?*

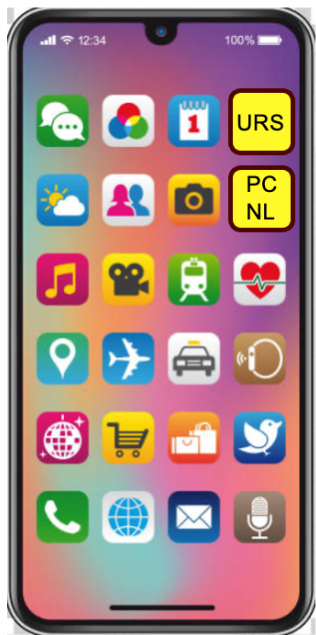
*Challenge: Learning happens everywhere!*



- Direct observation
- EPAs / skill assessments
- Narrative assessment
- Coaching
- Simulation / OSCEs

# Debate Topic 3:

## *Usefulness for Feedback, Remediation, and Excellence*



### Creation:

- Internally created
- Underwent iterative review / feedback
- Critical steps X procedure
- Housed in Qualtrics (free via institution)
- Easy use:
  - Save link in Notes app
  - Put shortcut on phone home screen

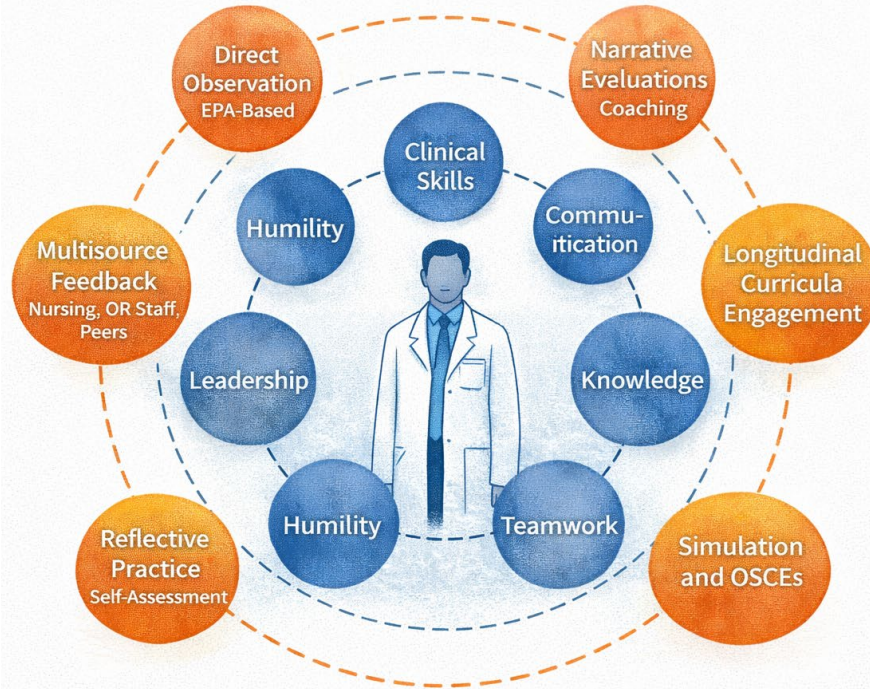


Average time to complete assessment  
= **119.0 seconds**  
(SD=42.6)



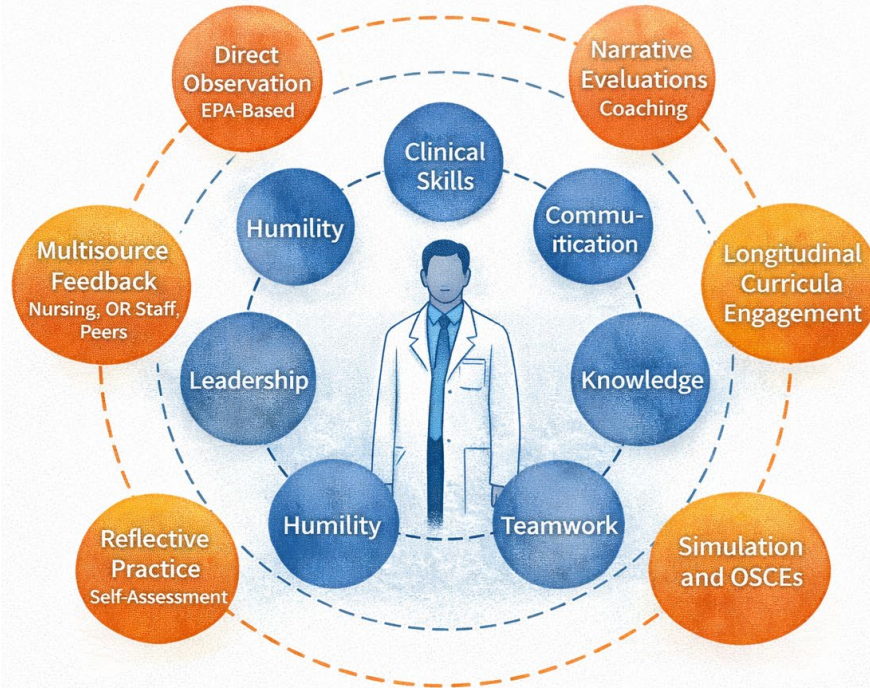
# Debate Topic 3:

## *Usefulness for Feedback, Remediation, and Excellence*



# Debate Topic 3:

## *Usefulness for Feedback, Remediation, and Excellence*

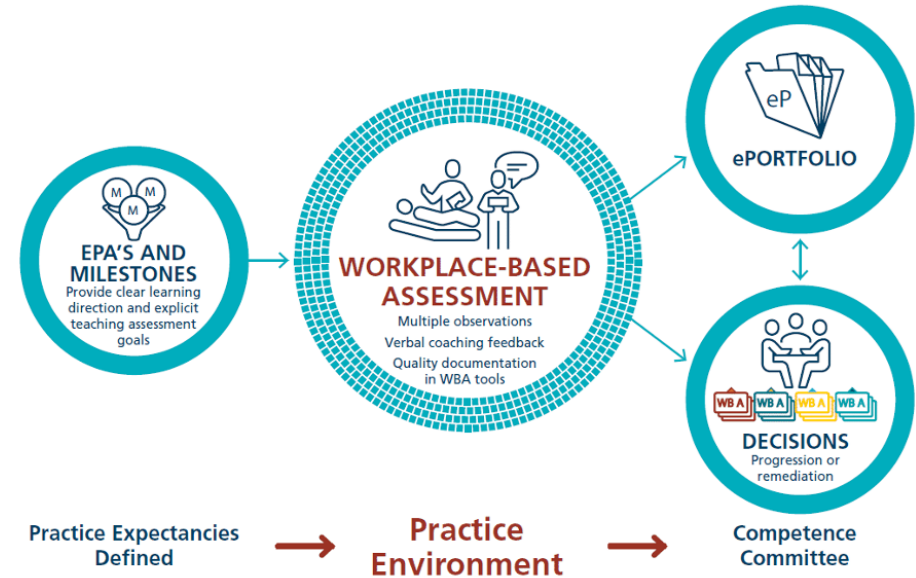


***The assessment toolbox isn't about adding work, it's about recognizing and organizing the work we already do***

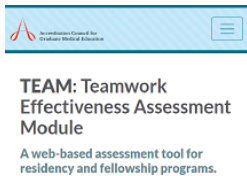
# Conclusions

## *Are Milestones Measuring Progress or Missing the Mark?*

- Milestones are not going away!
- The real risk is over-reliance on the Milestones alone
- Milestones are the toolbox and there are plenty of other tools that should fill out the toolbox
- High stakes decisions should be informed by compilation of tools



Dudek et al., *Royal College of Physicians and Surgeons*, 2017.



# Urology Milestones Maze

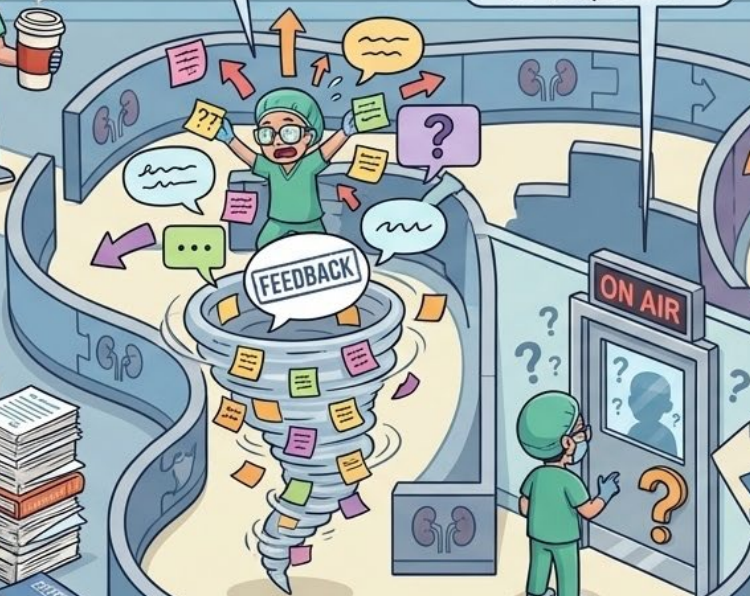
START

FINISH

**Level 1 – Orientation**  
"Too many forms, not enough coffee"



**Level 2 – Feedback Frenzy**  
"Notes everywhere, actionable nowhere"



**Level 3 – Entrustment Enigma**  
"Can I really do this unsupervised?"



**Level 4 – Remediation Roundabout**  
"Oops... back to basics"



**Level 5 – Competence Conundrum**  
"I think I'm ready... maybe?"





# SAU WINTER MEETING FEBRUARY 6-7, 2026

Houston, Texas • Marriot Marquis Houston

## *Thank you!*



**Christopher Jaeger, MD, MMSc**

[christopher.jaeger@corewellhealth.org](mailto:christopher.jaeger@corewellhealth.org)  
X: @Jaeger\_Chris



**Sydney McLean, MHA**

[smclean@acgme.org](mailto:smclean@acgme.org)  
[www.linkedin.com/in/sydney-mclean-acgme](https://www.linkedin.com/in/sydney-mclean-acgme)



**Tasha Posid, MA, PhD**

[Tasha.Posid@osumc.edu](mailto:Tasha.Posid@osumc.edu)  
X: @TashaPosidPhD