The Courage to Change:
Fostering a culture of continuous educational program quality improvement

KAISER PERMANENTE
BERNARD J. TYSON
SCHOOL OF MEDICINE

John L. Dalrymple, MD
Dean and CEO

I have nothing to disclose.

Objectives

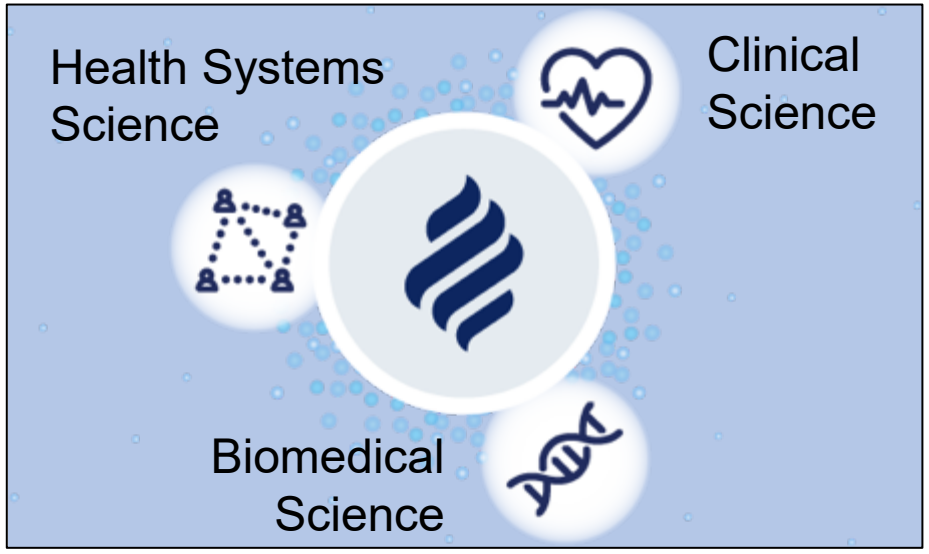
1. Describe the value of continuous program evaluation for strengthening educational quality and learner outcomes

2. Identify common personal, cultural, and program-level challenges that limit meaningful educational program evaluation and quality improvement

3. Outline practical, data-informed strategies for ongoing assessment and improvement of educational programs

My first 90 days.

An Integrated Curriculum



Did it hit the mark?

Why should we Evaluate?

Program Evaluation

What it is

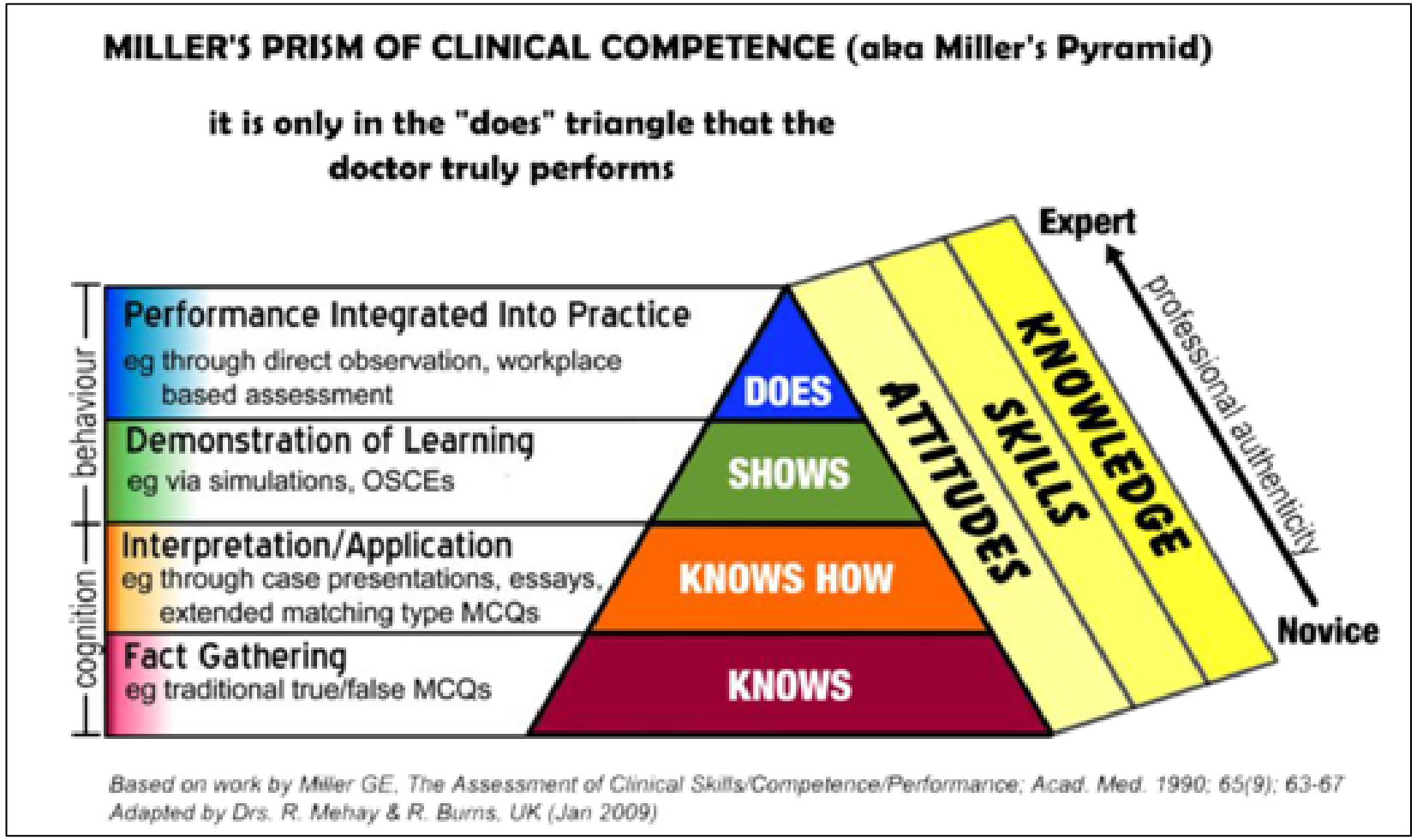
In health professions education, **program evaluation** is the systematic collection and analysis of data to make defensible value judgments about an educational program's merit, worth, or effectiveness. While it shares methodologies with research, its primary purpose is to **improve** a specific program rather than to prove generalizable truths.

Why we need it

- Achieving goals and objectives
- Continuous program quality improvement
- Appropriate use of resources

Program evaluation models and related theories: AMEE Guide No. 67
Frye AW & Hemmer P. *Medical Teacher 2012*

How does this relate to individual learners?



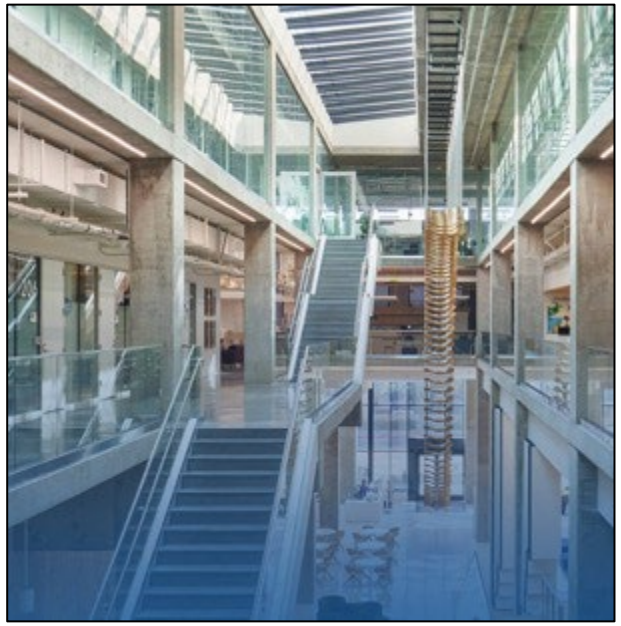
Opportunities to teach ***and*** assess:

- ✓ *Foundational knowledge and skills*
- ✓ *Core competencies*
- ✓ *Advanced abilities and readiness for independence*

Why do we avoid it?

Challenges and Barriers

- Prefer to do what's comfortable
- Easier to assess individuals
- Requires resources
- Lack frameworks or skills
- Requires data/inputs and analysis
- Takes time; Requires many cycles
- Too many constituents involved
- Change is hard; overwhelming
- Not part of the culture
- It's boring!



Time, Resources, Logistics

- Overlapping Functions & Workload
- Resource Constraints
- Short Clerkships/Rotations

Methodological & Technical Difficulties

- Limited Evaluation Skills
- Measurement Challenges
- Lack of Validated Tools
- Unclear Evaluation Aims

Institutional & Cultural Barriers

- Resistance to Change & Fear
- Low Priority/Recognition
- Lack of Institutional Support

Pedagogical & Assessment Challenges

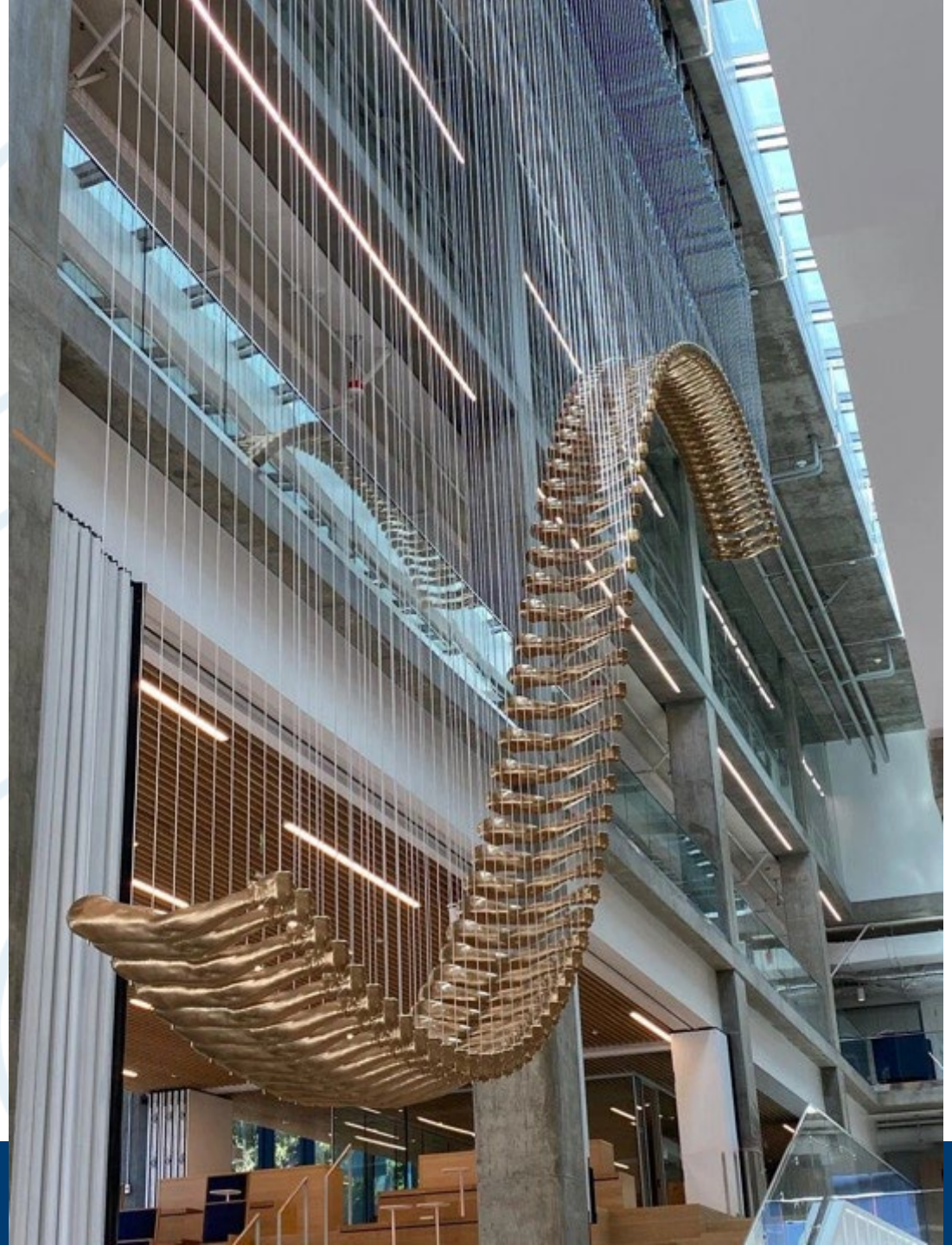
- Dual Roles
- Focus on Grades over Feedback
- Interprofessional Education (IPE) Complexities

Data & Analysis Issues

- Bias and Subjectivity
- Lack of Follow-up
- Data Overload and Poor Quality

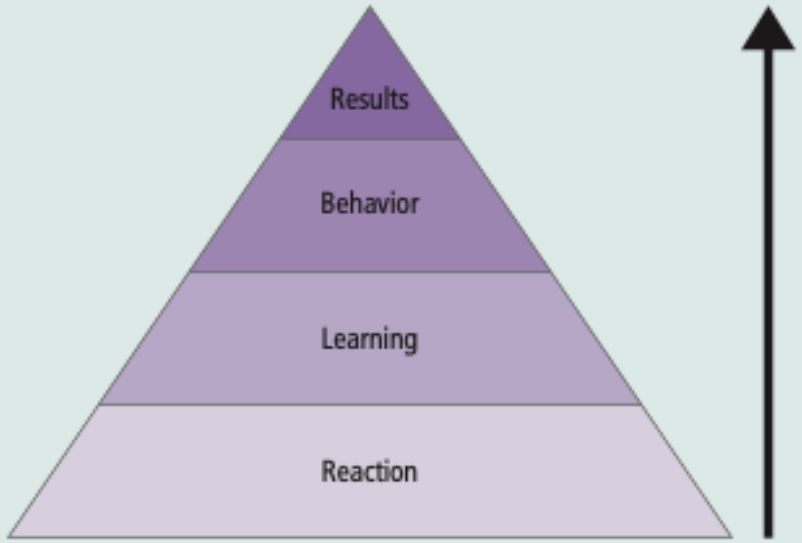
Aren’t you eager (curious) to see results?

- *How did the learners do?*
- *What did the course survey show?*
- *How did our graduates fare?*
- *Who got recognitions and teaching awards?*
  (and why – what are they doing that is working?)



Strategies, data sources, dashboards, methods

KIRKPATRICK'S TRAINING EVALUATION MODEL[4]



What will be ASSESSED? **Programmatic effects or effectiveness**

What do the LEVELS OF THE PYRAMID indicate?

Moving up the pyramid, what can the learner demonstrate?

Results—Influence lasting change in the workplace?
Behavior—Lead to behavior change in the program?
Learning—Acquire the intended knowledge, skills, or attitudes?
Reaction—Enjoy the training and feel it is relevant to their work?

WHEN to choose Kirkpatrick's Model?

When evaluating the effects of a program on trainees, patients, or populations.

W.K. Kellogg Foundation
Logic Model Development Guide
Updated, January 2004

The *WHAT*: Logic Model Definition

Basically, a logic model is a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your program, the activities you plan, and the changes or results you hope to achieve.

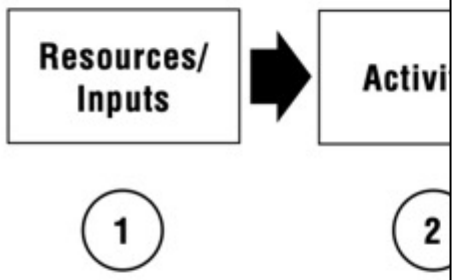


Figure 1. The Basic Logic Mode

The most basic logic mode
words and/or pictures to d
and how these activities ar

Logic Model Development
Program Implementation Template – Exercise 1 & 2

RESOURCES	ACTIVITIES	OUTPUTS	SHORT- & LONG-TERM OUTCOMES	IMPACT
In order to accomplish our set of activities we will need the following:	*In order to address our problem or asset we will accomplish the following activities:*	*We expect that once accomplished these activities will produce the following evidence or service delivery:*	*We expect that if accomplished these activities will lead to the following changes in 1–3 then 4–6 years:*	*We expect that if accomplished these activities will lead to the following changes in 7–10 years:*

With a little help from CoPilot

Are learners meeting Program Objectives?

Four Key Areas of Inquiry:

How do learners perform in this phase of the curriculum?

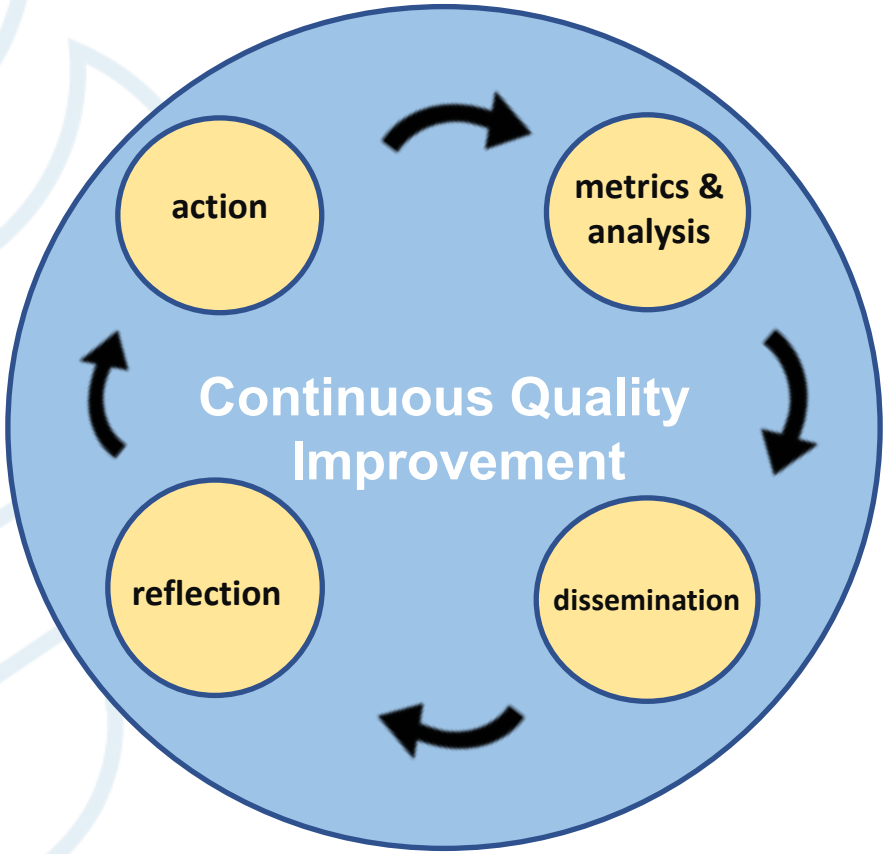
How do learners experience courses, rotations, supervisors and faculty?

Are learners adequately prepared to succeed in next phase of the curriculum/training? their careers?

Are we fostering a supportive learning environment and growth mindset?

Continuous Quality Improvement (CQI) in Health Professions Education

Possible Interventions:
- Curricular change
- Course development
- Pedagogical refinement
- Professional development



Reporting:
- Both scheduled and ad hoc
- Proactive and timely
- Easy to grasp graphical format

Tactics:
- Identify, amplify best practices
- Fill gaps
- Address areas for improvement

Key Constituents:
- Governance Committees
- Course Directors
- Educators/Supervisors
- Learners

Possible Data Sources

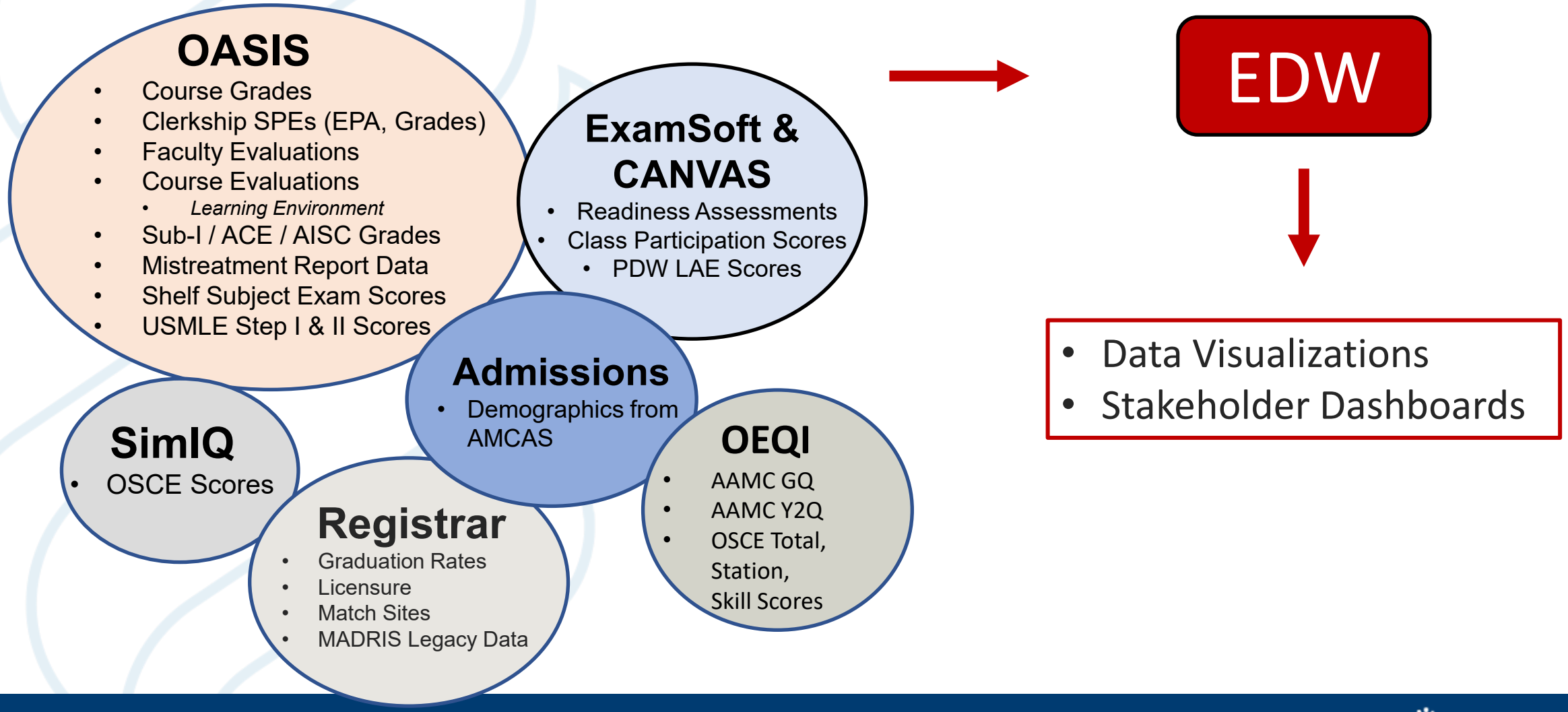
National Exams & Surveys

- ❑ Graduation questionnaires
- ❑ In-service and certification exams
- ❑ Nationally-administered learner surveys
- ❑ Nationally-derived exams
- ❑ Licensing exams
- ❑ Board examinations

Institutional assessments

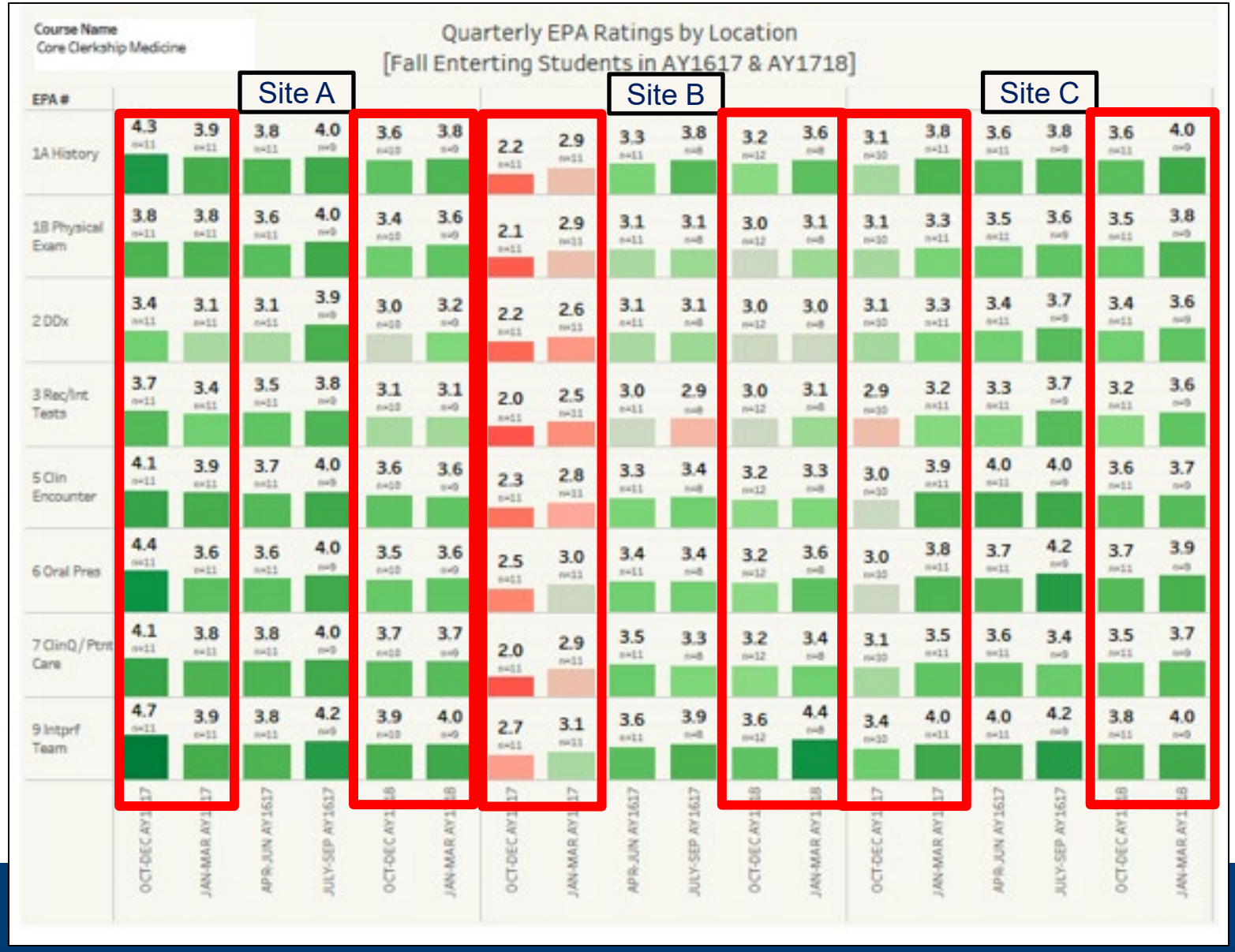
- ❑ Course, written narratives, grades
- ❑ Exams, quizzes, problem-sets
- ❑ Requirements checklist
- ❑ OSCEs
- ❑ Workplace based assessments
- ❑ Competency-based ratings
- ❑ Learner surveys (faculty, courses, LE)
- ❑ Peer assessments
- ❑ Post-graduation training surveys
- ❑ Graduation rates
- ❑ Alumni survey

Education Data Warehouse – Data Sources by Platform
(with thanks to Harvard Medical School Office of Medical Education)



Clerkship Director Dashboard – Site Comparisons

Workplace Based Assessments



Rating scale

1 (pre-entrustable)

3 (emerging)

5 (entrustable)

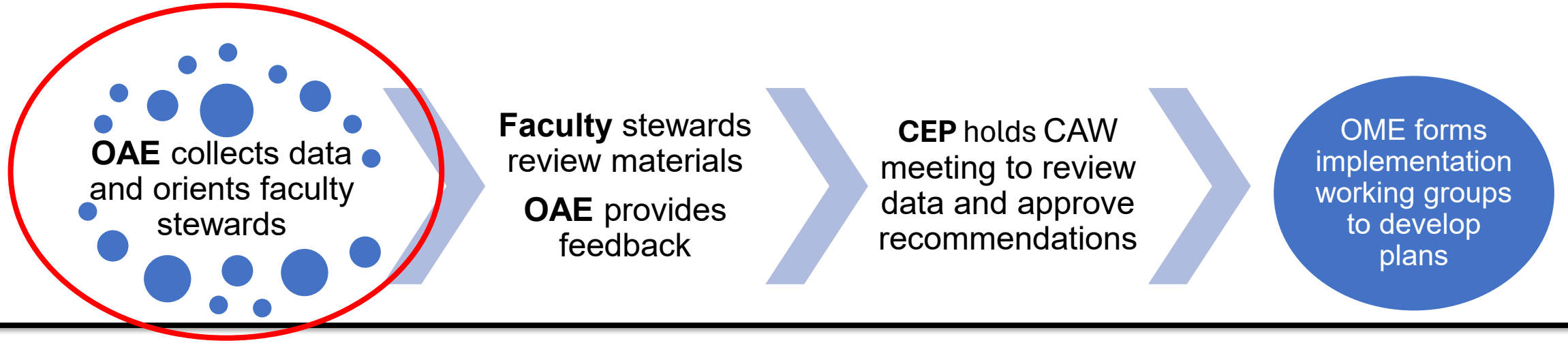
Data-informed Opportunities for Educational QI

Findings from Analysis	Possible Intervention
WBA* not aligned among supervisors	Professional Development
WBAs not being covered	Curriculum Development
Standardized exam performance	Gaps in content
OSCE performance	Vertical integration of clinical skills teaching
Grading disparities	Assessment equity; bias
Narrative quality	Professional Development
Professionalism lapses	Policy / Assessment Development
Educator performance	Professional Development
Learner readiness at key points	Program development

*Workplace Based Assessment

My first 90 days.

KPSOM Curriculum as a Whole (CAW) Review – Process*



Data Informing Recommendations

Data Steam #1: Relevant Curricular Data by EPO	Data Stream #2: CAW Written Survey	Data Stream #3: CAW Focus Groups	Data Stream #4: AAMC GQ
• Learning data (instructional coverage, assessment data, Competency Committee judgments)	• Graduating student experience ratings and comments	• Graduating student experience	• Student experience with national benchmarks
Adequacy of curricular coverage and assessment + student learning data	**What students found valuable (and not)** **How KPSOM measured up to its mission and national average**		

* With thanks to Dr. Carla Lupi, Senior Associate Dean for Medical Education, Sophia Harvey, and the Office of Medical Education

Changes stemming from CAW

Educational Program Objective (EPO) Domain (# of EPOs)	Changes to EPOs	Changes to Milestones	Changes to Instruction	Changes to Assessment
Patient Care (7)		Revised milestone language for Technology and Digital Care (CSDC), Urgent Care (PCUC)	Incorporate health information technology in HSS in Phase 3; enhanced instruction of PCUC; additional Transitions (PCTR) activities	EHR-specific instruction/ assessment in Phase 3; increased PCUC and PCTR simulation assessment
Lifelong Learning (3)		Revised milestone language for Wellbeing (LLWB), Uncertainty (LLUN)		Add LLWB assessment
System-based Practice (5)	Revised Quality Improvement (SBQI) graduation requirement	Revised milestone language for Patient Safety (SBPS)		Introduce Leadership Change (SBLC) portfolio assessment; enhance Systems Thinking and Design (SBST) assessment prompts
Interpersonal and Communication Skills (2)	Revised Roles and Teamwork (IPRT) and Collegial Communication (IPCC) graduation requirement		Pilot IPCC and IPRT instruction in Phase 2	Pilot IPCC and IPRT CASP assessment in Phase 2
Professionalism (3)			Increase instruction in Trustworthiness (PRTR) in REACH	Increase PRTR assessment in REACH

EPO Domains not affected by CAW: Population and Community Health, Interprofessional Collaboration and Teamwork, Medical Knowledge

* With thanks to Dr. Carla Lupi, Senior Associate Dean for Medical Education, Sophia Harvey, and the Office of Medical Education

Using an Outcomes-Logic-Model Approach to Evaluate a Faculty Development Program for Medical Educators

Elizabeth G. Armstrong, PhD, and Sylvia J. Barsion, PhD

Abstract

Purpose

This study used an outcomes-logic-model approach to examine the impact of participat professio Building provides examine learner o methods, approach

Metho

Structure conducte Medical Harvard Educator

the program in 1998, five in 1999, and six in 2000. Interviews were also conducted with four Faculty Scholars

knowledge about and confidence us learner-center teaching methods 10 16 (63%) said they gave fewer lectu ional %) repo e field e third rning p

ovided m miss and attaine

Logic Model Development
Program Implementation Template – Exercise 1 & 2

RESOURCES	ACTIVITIES	OUTPUTS	SHORT- & LONG-TERM OUTCOMES	IMPACT
In order to accomplish our set of activities we will need the following:	*In order to address our problem or asset we will accomplish the following activities:*	*We expect that once accomplished these activities will produce the following evidence or service delivery:*	*We expect that if accomplished these activities will lead to the following changes in 1–3 then 4–6 years:*	*We expect that if accomplished these activities will lead to the following changes in 7–10 years:*

Table 1

The Outcomes Logic Model for the Harvard Macy Institute Program for Physician Educators, Harvard Medical School*

Component	Description
Inputs What resources are dedicated to or consumed by the program?	• Funding from foundation (initially); tuition plus operational budget (currently) • Faculty and staff time (within HMS, Harvard University, and beyond) • Facilities at HMS and associated hospitals
Activities What does the program do with inputs to fulfill its mission?	• Systems to publicize program, screen potential participants, manage course logistics, etc. • Curriculum design that incorporates: (1) assessment of learning needs, (2) interactive learning and opportunities to practice, (3) sequenced and multifaceted activities, and (4) outcome evaluation. • Ongoing curriculum updating • Winter and spring sessions with evaluation during and after both sessions to judge whether program is meeting needs and is implemented as planned. • Systems to support medical educator networking and communities of practice before, during and after participation, e.g., HM-PE Web site, participant reunions, recruitment of program alumni as Faculty Scholars.
Outputs What are the direct outputs of program activities?	• Number of applicants and participants since the program began • Average number of hours a participant spends on the program and related activities • Number of participants from HMS • Number of participants from across the United States and internationally • Number of "hits" on HM-PE Web site
Outcomes What are the immediate and intermediate benefits for participants during and after program activities?	*Immediate* • Increased knowledge about active learning methods and greater capacity to be learner-centered educators. • New understanding of and appreciation for ways medical education is implemented in institutions nationally and globally. *Intermediate* • Belief that the program was "transformational," leading to an increased commitment to medical education as a primary career direction and stronger identity as a medical educator. • Expanded network of colleagues in medical education and communications with like-minded physician educators, e.g., via virtual communities, ongoing emails, collaborative activities.

* The outcomes logic model seeks to document to what extent the immediate and intermediate outcomes shown above are attained by program participants.

MEDICAL TEACHER
2022, VOL. 44, NO. 11, 1268–1276
https://doi.org/10.1080/0142159X.2022.2081142

MEDICAL TEACHER

The Harvard Medical School Pathways curriculum: A comprehensive curricular evaluation

Amy M. Sullivan[a], Edward Krupat[a], Jules L. Dienstag[b], Jakob I. McSparron[c], Stephen R. Pelletier[d], Sara B. Fazio[a], Thomas J. Fleenor[d], John L. Dalrymple[e], Edward M. Hundert[f] and Richard M. Schwartzstein[a]

[a]Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; [b]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [c]Department of Medicine, University of Michigan School of Medicine, Ann Arbor, MI, USA; [d]Office of Educational Quality Improvement, Harvard Medical School, Boston, MA, USA; [e]Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA; [f]Global Health and Social Medicine and Medical Education at Harvard Medical School, Boston, MA, USA

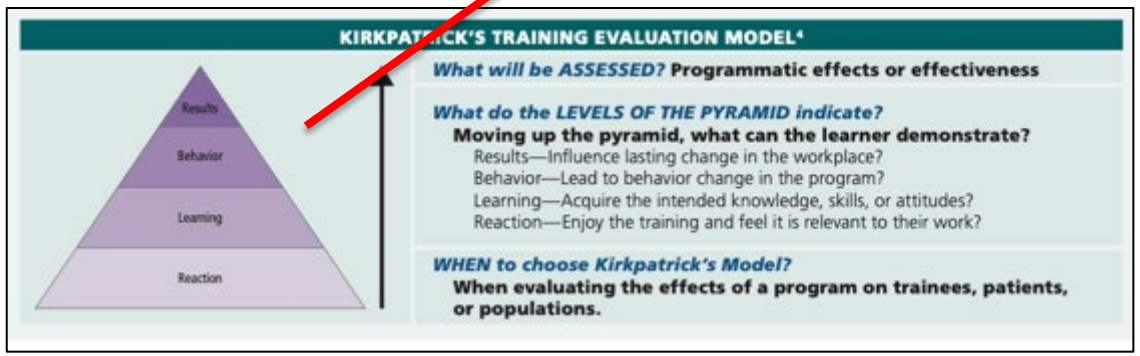


Table 1. Summary of measures used to compare student cohorts and faculty perceptions of student knowledge, skills, and attitudes in the prior and new curricula.

Category and New World Kirkpatrick (NWK) Level[a]	Constructs	Instruments	Timing of administration
Attitudes and perceptions of learning environment	Empathy	Jefferson Empathy Scale (JSE), a 20-item self-report instrument (pre- and post).	A,B
	Patient-centeredness	Patient-Practitioner Orientation Scale (PPOS), an 18-item survey instrument measuring patient-centered beliefs.	
NWK 1, 2, 4	Need for cognition	Need for Cognition Scale (NFCS), 18 items measuring predisposition toward effortful thinking	
	Curiosity, defined as an intrinsically-motivated drive for learning and information-seeking	I-Type and D-Type Scales, 5 items each	
	Learning environment	Educational Climate Inventory, 20-item instrument with three subscores: Centrality of learning and mutual respect; competitiveness and stress; and Passive learning and rote memorization.	B
		AAMC Year Two Questionnaire, annual survey for 2nd year students, includes learning climate items from the Medical School Learning Environment Survey (MSLES) and 8-item Oldenburg Burnout Inventory	D
	Student satisfaction with medical education	Focus groups	B
		Exit Surveys	F
Knowledge NWK 2	Foundational medical knowledge	USMLE Step 1 Exams.[b] Medical content knowledge in anatomy, behavioral sciences, biochemistry, biostatistics and epidemiology, microbiology, pathology, pharmacology, and physiology.	B
Skills NWK 3	Core clinical skills-summative	9-Station summative Objective Structured Clinical Examination (OSCE) (at beginning of final year). Summative assessment of observed student skills in conducting history, physical exam skills, communication and clinical reasoning in standardized test setting	E
		Clinical Capstone OSCE (within last four months)	E
	Student clinical readiness	Clerkship director faculty perceptions of each student cohort's preparation to learn and participate in the clinical environment. 11-item survey with ratings for students from each of the two cohorts. Items covered observed knowledge, physical examination and presentation skills, and openness to feedback.	C
	Clinical skills—formative, longitudinal	Entrustable professional activities focused on 13 specific skills. Longitudinal, formative assessment of student clinical skills across different phases of the curriculum, including pre-clerkship, clerkship and post-clerkship courses.	G

*Key: A = Administered at start of preclerkship period (Year 1); B = End of preclerkship period; C = First quarter of first year of clerkship; D = Mid-year 2nd year; E = Final year before graduation; F = Graduation; G = Administered across entire curriculum.

[a]New World Kirkpatrick (NWK) model levels: 1. Reaction (satisfaction, engagement), 2. Learning (attitudes, knowledge, skills), 3. Behaviors (application of learning), and 4. Leading indicators (here, perceived learning environment, curiosity, empathy, and burnout).

[b]USMLE: U.S. Medical Licensing Exam (3 years of data were used for Step 1, 2 years of data available for Step 2 CK)

Teaching and Learning in Nursing 13 (2018) 9–13

Contents lists available at ScienceDirect

Teaching and Learning in Nursing

journal homepage: www.jtln.org

Using the CIPP Model to Assess Nursing Education Program Qu and Merit

Megan Lippe *,[1], Patricia Carter [2]

University of Texas at Austin School of Nursing, Austin, TX 78701, USA

ARTICLE INFO

Article history:
Accepted 14 September 2017

Keywords:
Nursing education
Program evaluation
End-of-life care

ABSTRACT

Nursing programs must maintain a high-quality curriculum that g
evaluation of key components of nursing education is required.
and Product model allows for evaluation of the quality and merit
nursing program. Data analysis identified missing content, program
within the program. When used appropriately, Stufflebeam's mode
curriculum evaluation.

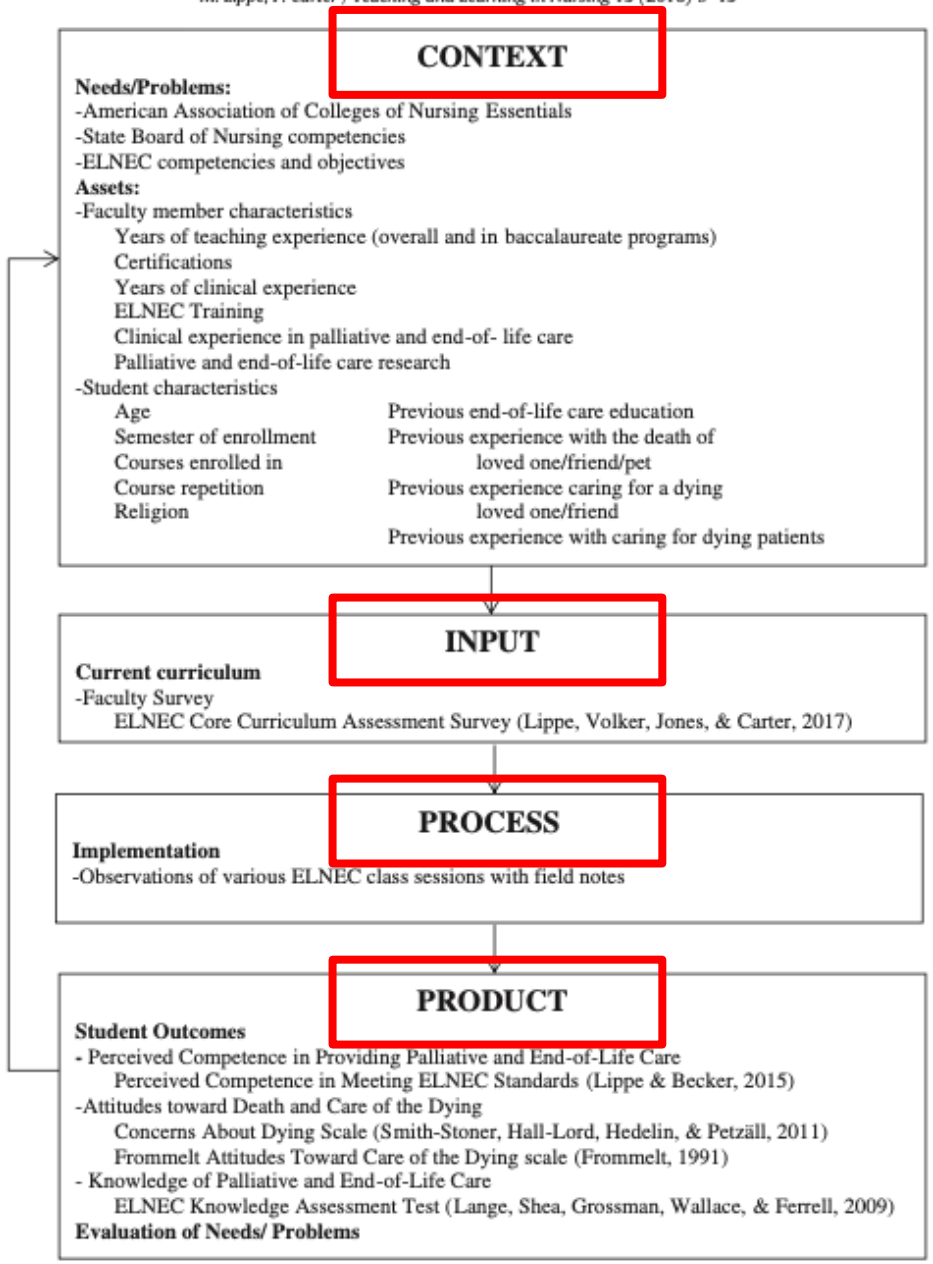


Fig. 1. Palliative care curriculum evaluation model (adapted from Stufflebeam's CIPP model). ELNEC = End-of-Life Nursing Education Consortium.

Call to Action

Lessons learned and tips to move forward

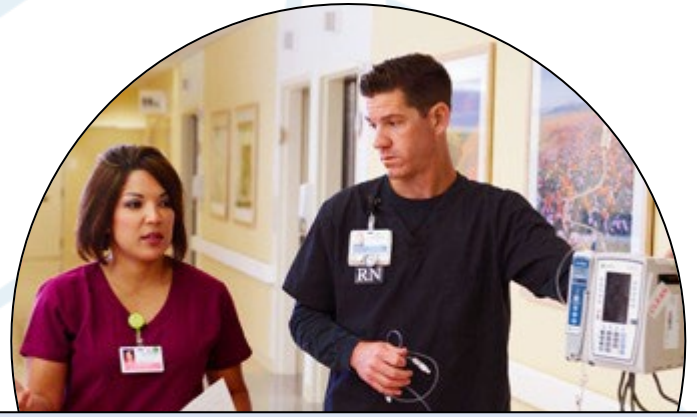
- ✓ Begin with the end in mind
- ✓ Upfront planning helps
- ✓ Assemble your team – find an ally, buddy, champion
- ✓ Use a framework, establish a process, develop skills
- ✓ **Change the culture**

  Frame in positive way – “we want to be better”

  Be curious – “how are we doing?”

  Espouse humility – “it’s OK to do this”

  Find the COURAGE! Be BRAVE – just do it!



What's holding you back?
How will you know you are achieving your programmatic goals?
What's your **NEXT STEP**?

Trust is at the core

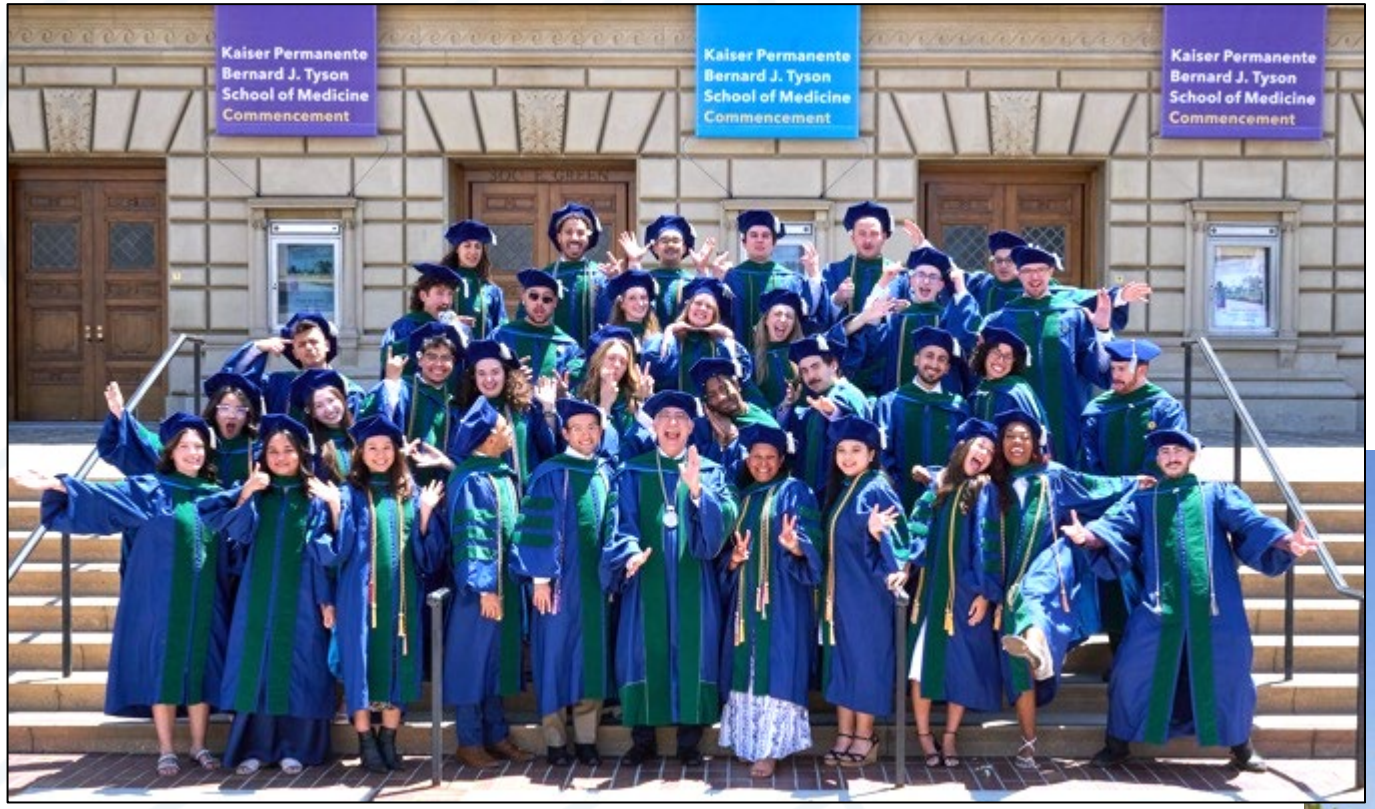
*"**Teachers must trust their learners**. The trust of teachers is won when learners attend, engage, respond, a the teachers and their vo takes time and repeated ex to be optimally developed both availab engagement tha contrast, trust i not attend or a resistant, dismissive, de unresponsive*

*"**Teachers make trusting judgments about their learners all the time.** Teachers must progressively relinquish control and supervision to their learners, attending to th indirect supervision with the tea within close proximity to supe distance; ultimately, the learne to supervise other*

We owe it to our learners and our faculty and society.

*"**Students, residents, and peer-learners must rs.** The level of trust in th the engagement in chers create trust by nstrating their own competence and professionalism, showing interest in the learners, and providing meaningful feedback that helps learners grow. Reciprocally, educators who harass or otherwise lose their learners' trust may continue to teach but will be unlikely to create effective learning. Trust that is eroded or lost can be impossible to recover."*

Thank You!



KAISER PERMANENTE
BERNARD J. TYSON
SCHOOL OF MEDICINE