
THE UNIVERSITY OF ILLINOIS
COLLEGE OF MEDICINE
CHICAGO PEORIA ROCKFORD URBANA

# Assessing Learners Using Mastery Learning: Issues in Validity, Standard Setting, and Applications

University of Tokyo

Medical Education Lecture

Yoon Soo Park, PhD

November 2, 2017

# Overview

1. Overview of Mastery Learning
   - Time-Based Model
   - Mastery-Based Learning

2. Implications for Standard Setting
   - Borderline / Marginal Performance
   - Standard Setting for Mastery Learning

3. Competency-Based and Time-Variable Programs
   - Education in Pediatrics Across the Continuum (EPAC)
   - Orthopedic Residency Program in Canada

# "Tea-Steeping" Model (1)

Hodges BD. *Academic Medicine.* 2010
Snell LS, Frank JR. *Medical Teacher*. 2010

Tea Bag

"Steep" in Hot Water

Fixed Time

"Good" Tea!

# "Tea-Steeping" Model (2)

Hodges BD. *Academic Medicine.* 2010
Snell LS, Frank JR. *Medical Teacher*. 2010

Tea Bag

Medical Student

"Steep" in Hot Water

Medical School

"Good" Tea!

Competent Physicians

Fixed Time

Four or Six Years

# Why Mastery Learning?

## Making the Case for Mastery Learning Assessments: Key Issues in Validation and Justification

Matthew Lineberry, PhD, Yoon Soo Park, PhD, David A. Cook, MD, MHPE, and Rachel Yudkowsky, MD, MHPE

Abstract

Theoretical and empirical support is increasing for mastery learning, in which learners must demonstrate a minimum level of proficiency before completing a given educational unit. Mastery learning approaches aim for uniform achievement of key objectives by allowing learning time to vary and as such are a course-level analogue to broader competency-based curricular strategies. Sound assessment is the cornerstone of mastery learning systems, yet the nature of assessment validity and justification for mastery learning differs in important ways from standard assessment models. Specific validity issues include (1) the need for careful definition of what is meant by "mastery" in terms of learners' achievement or readiness to proceed, the expected retention of mastery over time, and the completeness of content mastery required in a particular unit; (2) validity threats associated with increased retesting; (3) the need for reliability estimates that account for the specific measurement error at the mastery versus nonmastery cut score; and (4) changes in item- and test-level score variance over retesting, which complicate the analysis of evidence related to reliability, internal structure, and relationships to other variables. The positive and negative consequences for learners, educational systems, and patients resulting from the use of mastery learning assessments must be explored to determine whether a given mastery assessment and pass/fail cut score are valid and justified. In this article, the authors outline key considerations for the validation and justification of mastery learning assessments, with the goal of supporting insightful research and sound practice as the mastery model becomes more widespread.

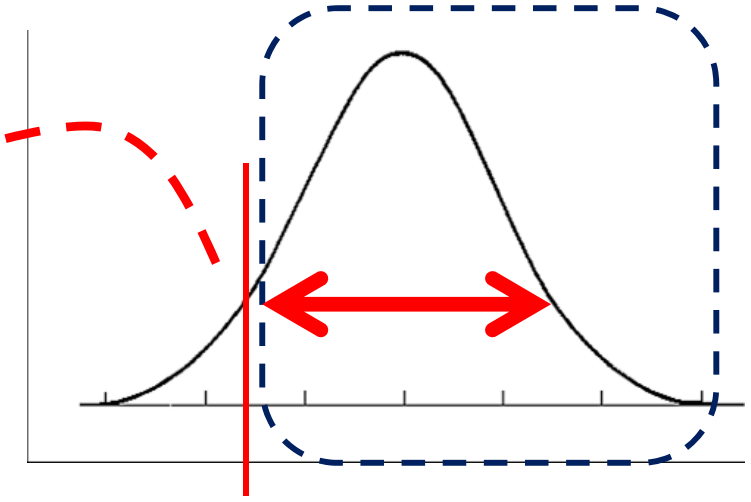
Academic Medicine, Vol. 90, No. 11 / November 2015

# Traditional Learning:
# Time-Based Model (1)

Learning Time → Fixed


Learning Outcome → Variable

Passing Score:
Minimum
Competence



# Mastery Learning (1)

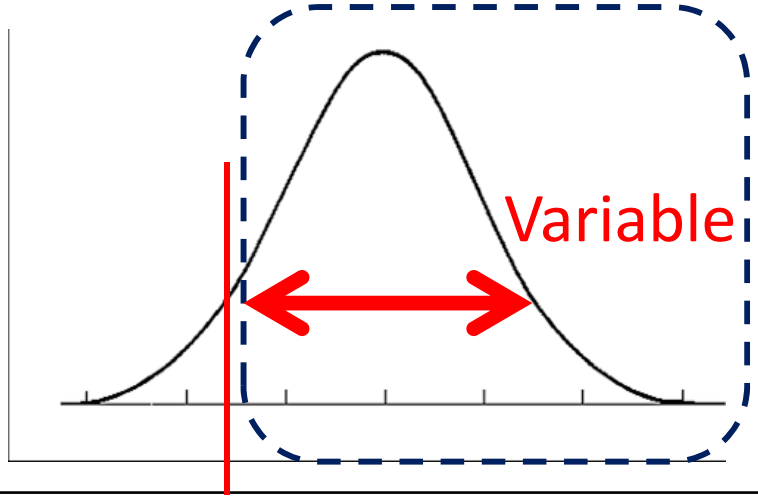
- Developed in 1968 by Benjamin Bloom

## Learning Time

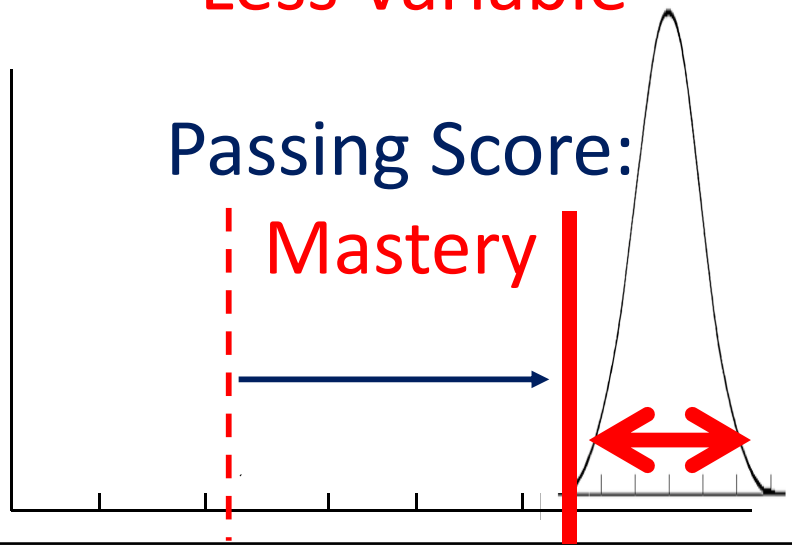

Fixed


Variable

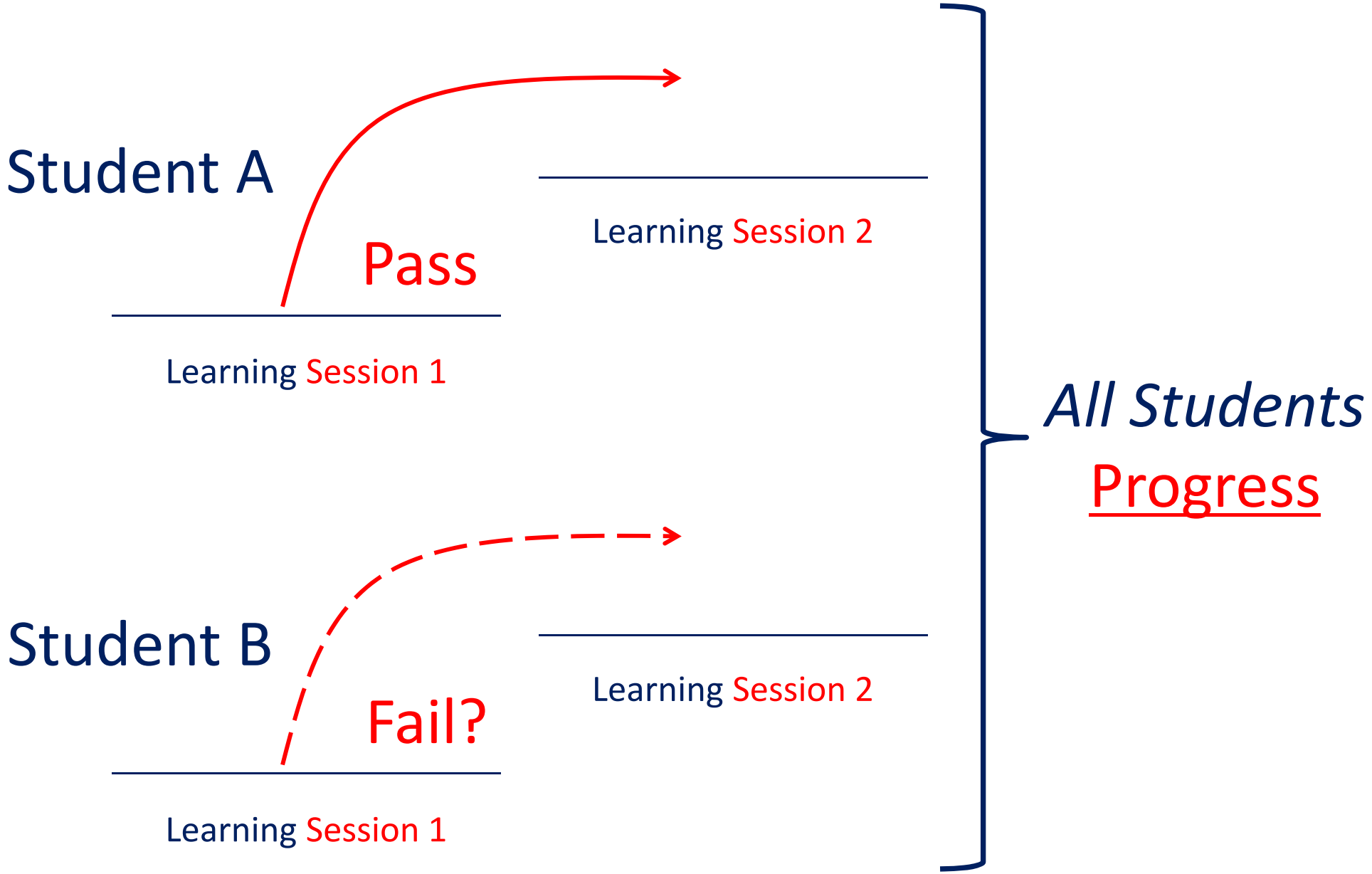
## Learning Outcome



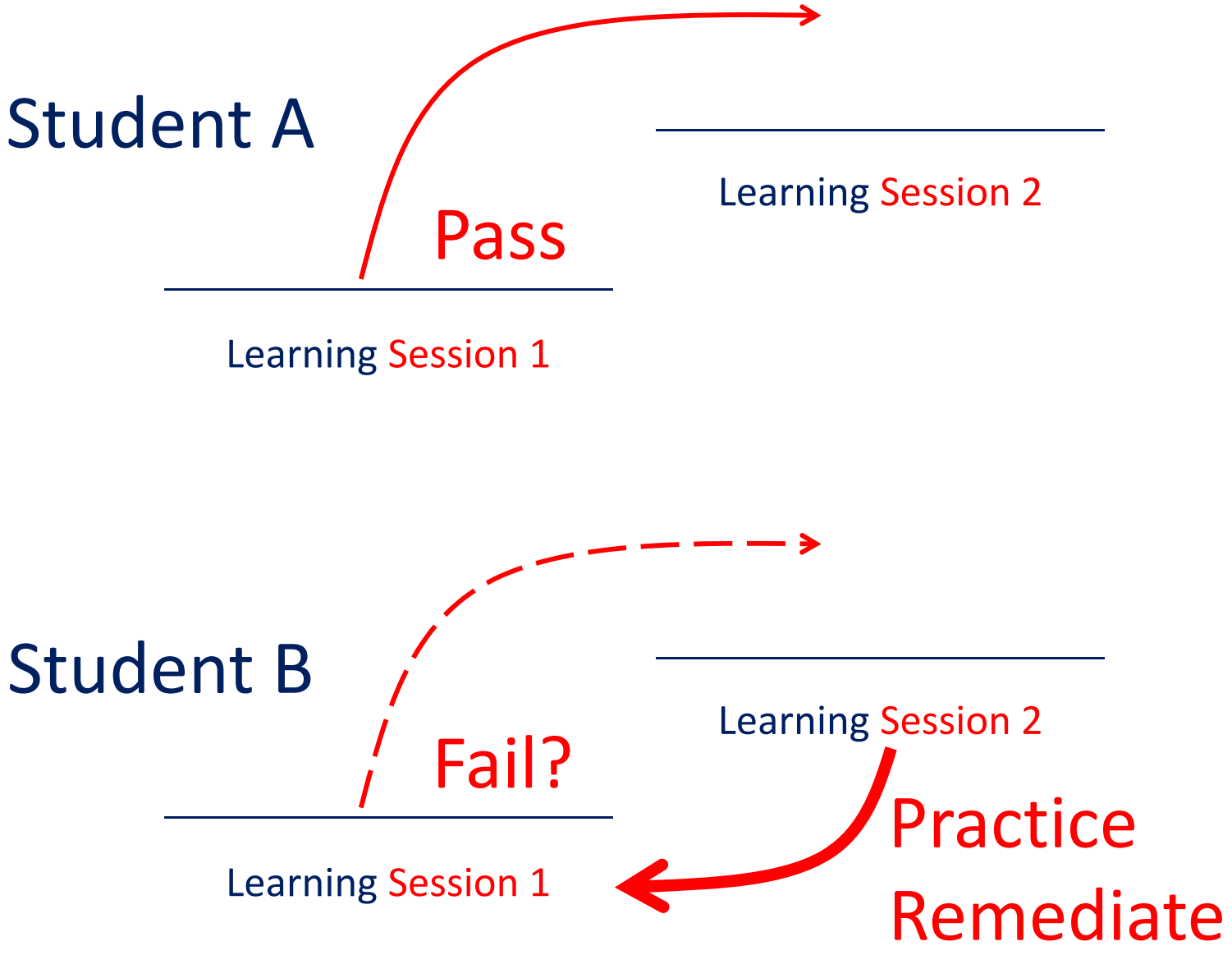

Less Variable



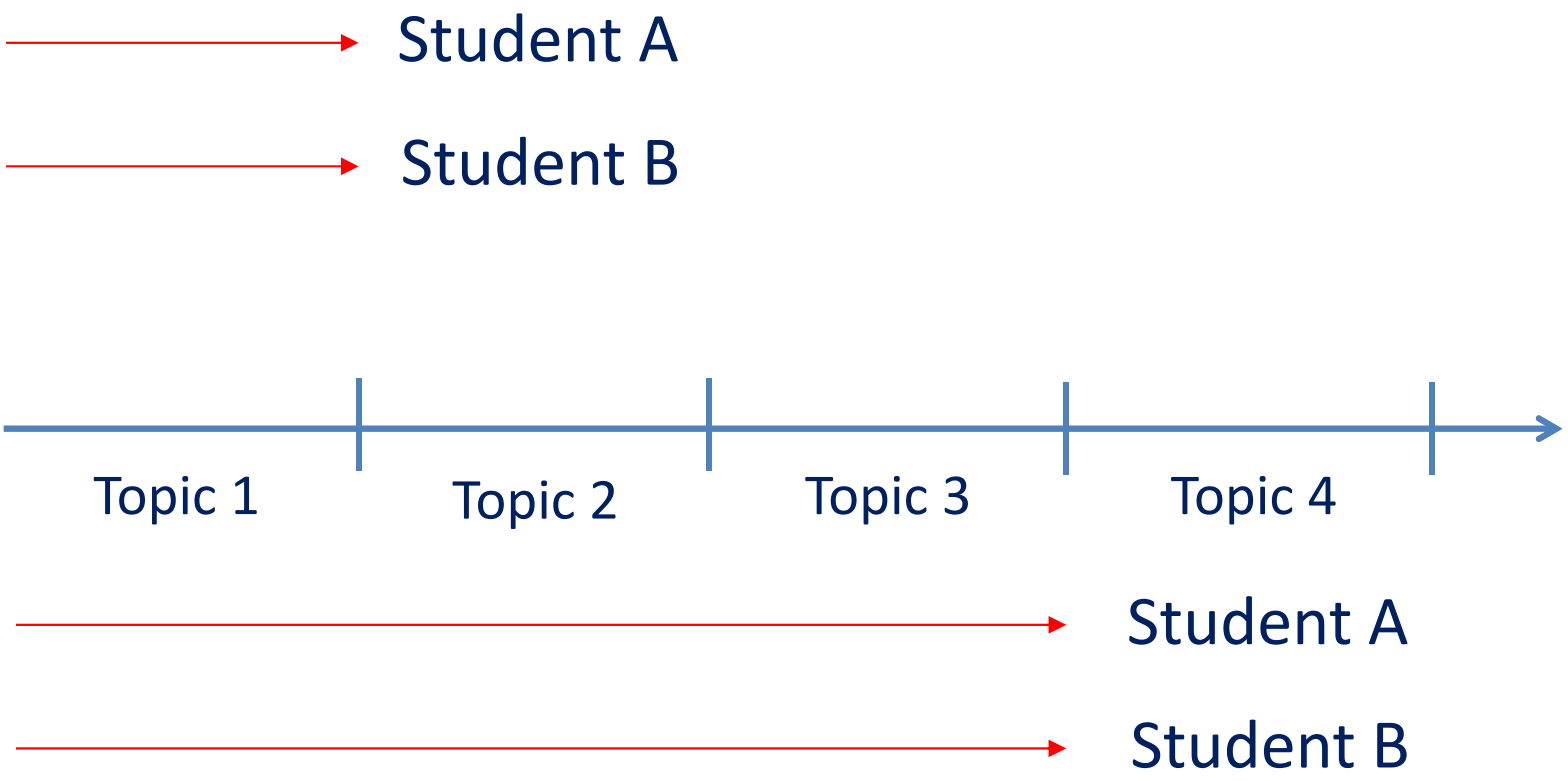
# Time-Based Learning (2)



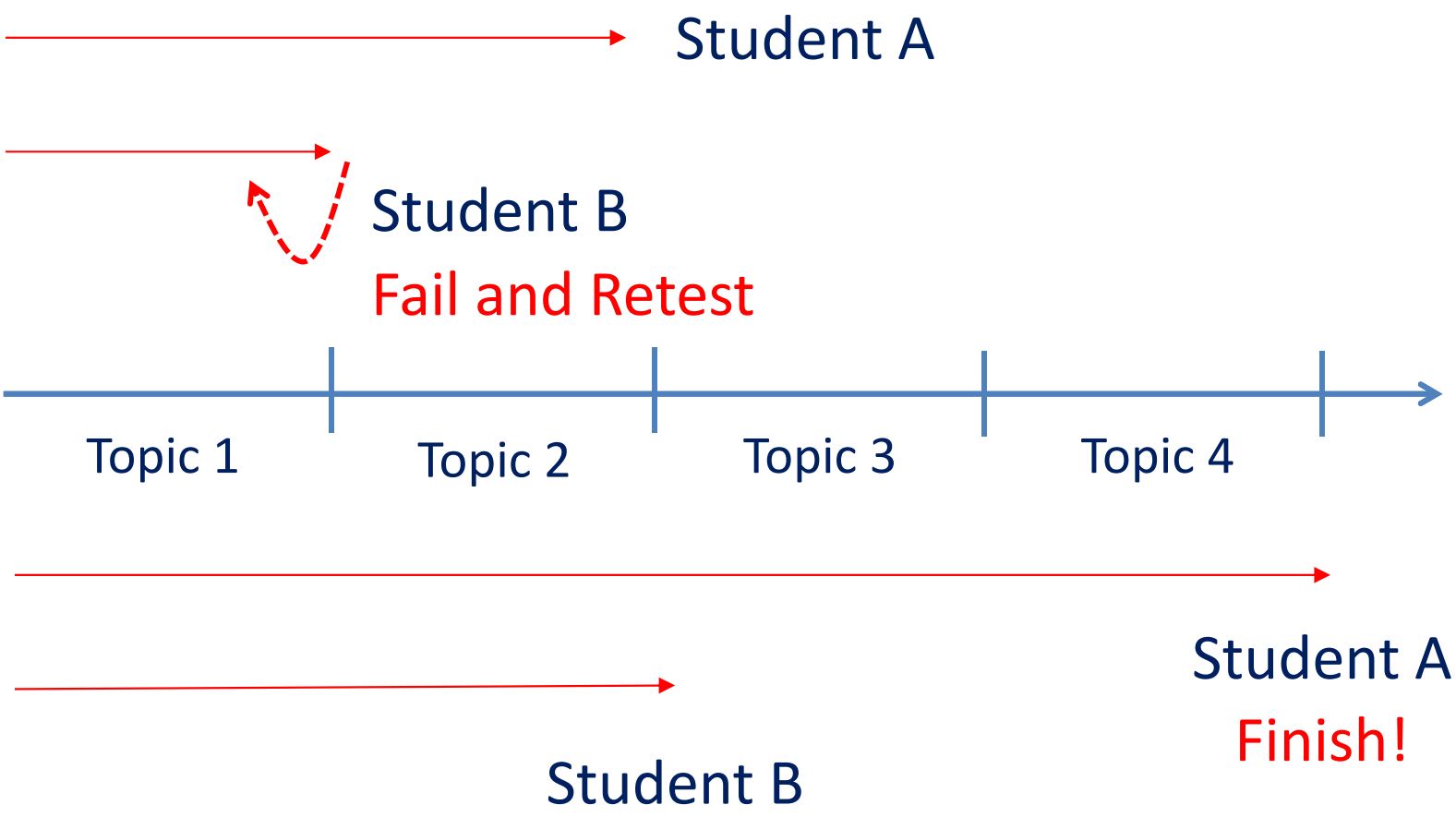
# Mastery Learning (2)



# Time-Based Traditional Learning (3)

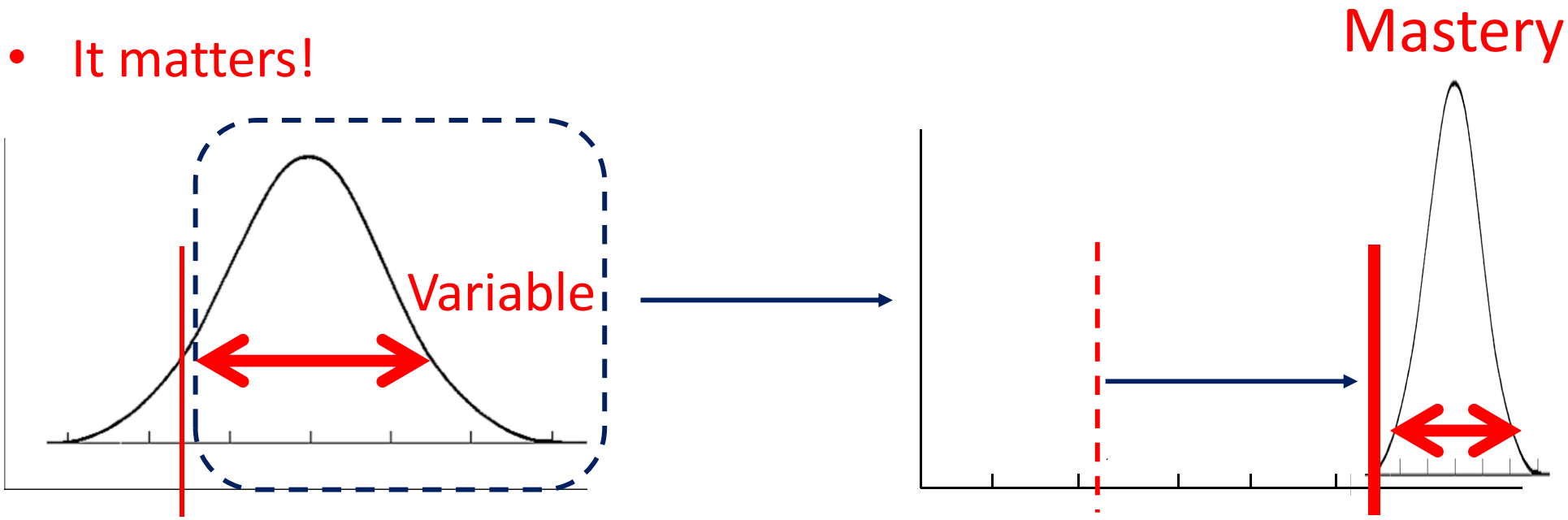


# Mastery Learning (3)



# Mastery Learning (4)

- Everyone is well prepared!
- It matters!




- But – what is the evidence?

# Evidence from Laparoscopic Skill

**Laparoscopic simulation training with proficiency targets improves practice and performance of novice surgeons**

Paul G. Gauger, M.D.[a,b,*], Linnea S. Hauge, Ph.D.[a,b],
Pamela B. Andreatta, Ed.D., M.F.A.[b], Stanley J. Hamstra, Ph.D.[b],
Miranda L. Hillard, B.S.[b], Eamonn P. Arble, B.A.[a], Steven J. Kasten, M.D.[a],
Patricia B. Mullan, Ph.D.[b], Paul S. Cederna, M.D.[a], Rebecca M. Minter, M.D.[a,b]

The American Journal of Surgery (2010) 199, 72–80

- Intervention group ⟶ task-specific proficiency criteria
- Performance Outcome
  - Meet target
  - Depth perception
  - Bimanual dexterity
  - Efficiency
  - Tissue handling
  - Autonomy
  - Overall competence

Improvement relative to Control Group

# Randomized Controlled Trial

## Simulation-Based Mastery Learning Improves Patient Outcomes in Laparoscopic Inguinal Hernia Repair

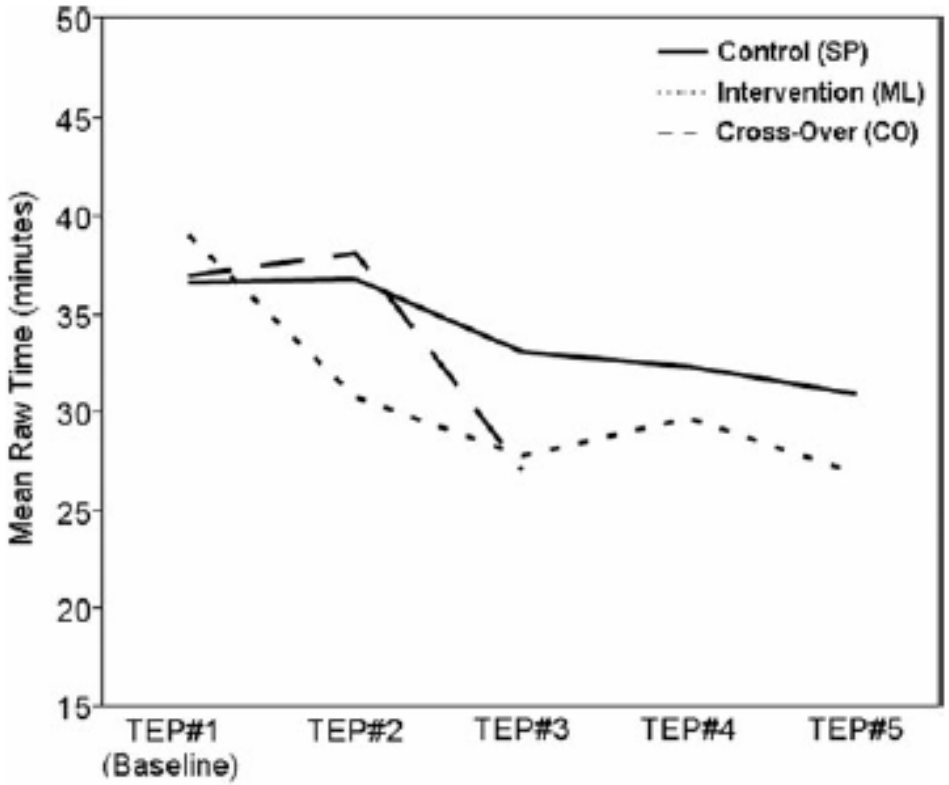
*A Randomized Controlled Trial*

*Benjamin Zendejas, MD, MSc,*† David A. Cook, MD, MHPE,†‡§ Juliane Bingener, MD,* Marianne Huebner, PhD,¶ William F. Dunn, MD,†§ Michael G. Sarr, MD,* and David R. Farley, MD*†*

*Annals of Surgery* • Volume 254, Number 3, September 2011

- Totally Extraperitoneal (TEP) inguinal hernia repair
  - 219 TEP repairs
  - 146 patients

- ↑ speed
- ↑ operative performance
- ↓ complications
- ↓ need for overnight stay

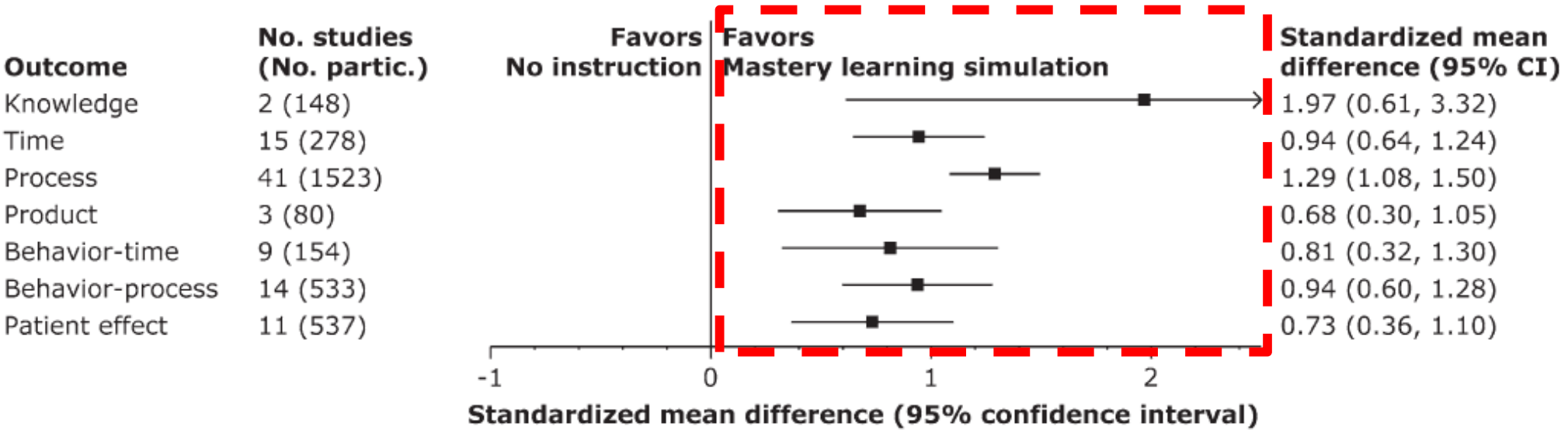


# Systematic Review and Meta-Analysis

## Mastery Learning for Health Professionals Using Technology-Enhanced Simulation: A Systematic Review and Meta-Analysis

David A. Cook, MD, MHPE, Ryan Brydges, PhD, Benjamin Zendejas, MD, MSc, Stanley J. Hamstra, PhD, and Rose Hatala, MD, MSc

Academic Medicine, Vol. 88, No. 8 / August 2013



# Standard Setting

**Setting Mastery Learning Standards**

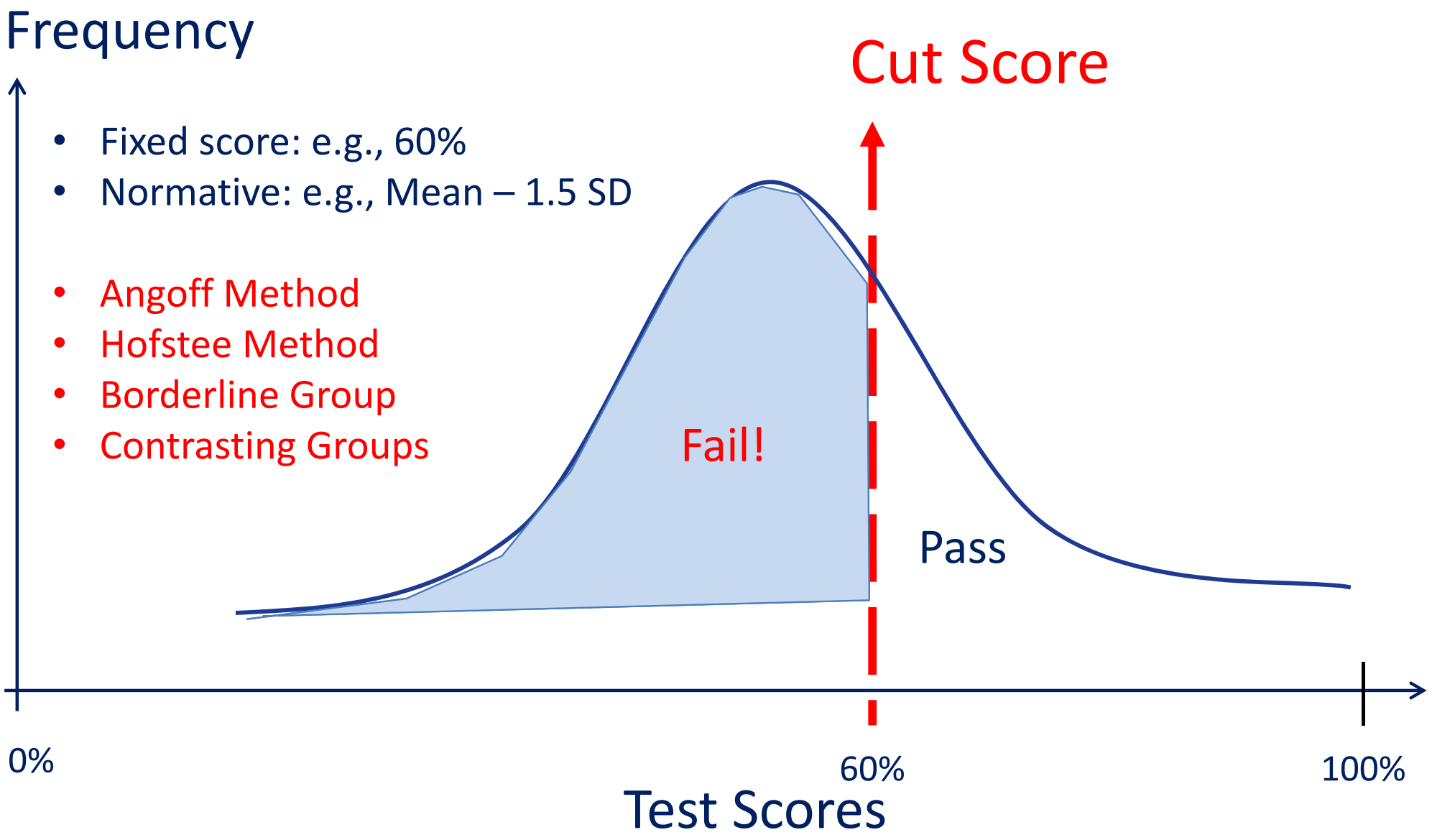
Rachel Yudkowsky, MD, MHPE, Yoon Soo Park, PhD, Matthew Lineberry, PhD, Aaron Knox, MD, and E. Matthew Ritter, MD

Academic Medicine, Vol. 90, No. 11 / November 2015

*How are mastery standards determined?*

THE UNIVERSITY OF ILLINOIS
COLLEGE OF MEDICINE
CHICAGO PEORIA ROCKFORD URBANA

# *How* do you currently set cut scores for your exams?




- Fixed score: e.g., 60%
- Normative: e.g., Mean – 1.5 SD

- Angoff Method
- Hofstee Method
- Borderline Group
- Contrasting Groups

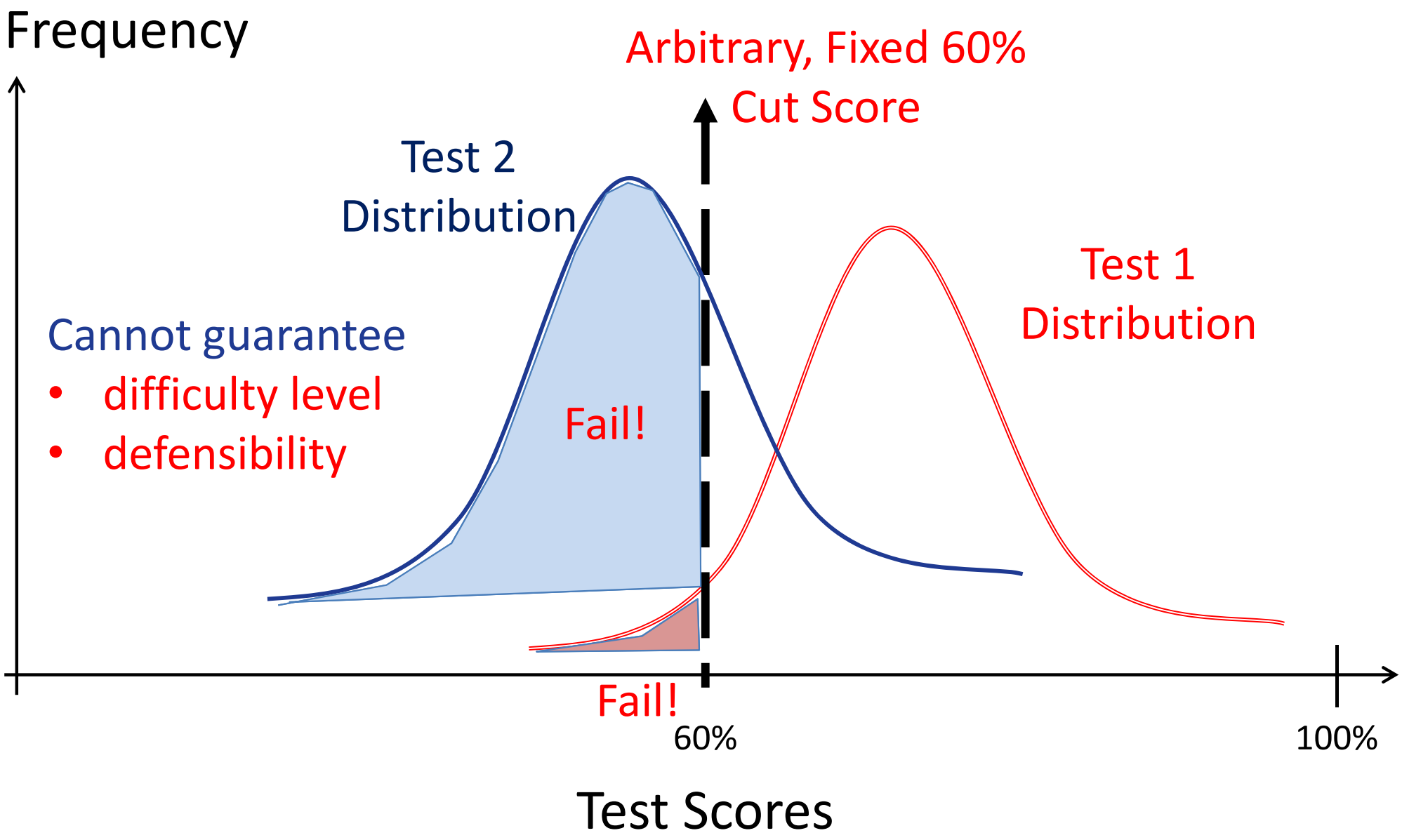
# What is standard setting?

- "Absolute" passing standard (e.g., 60%)
- "Norm-referenced" passing standard (e.g., Mean – 1.5 SD)

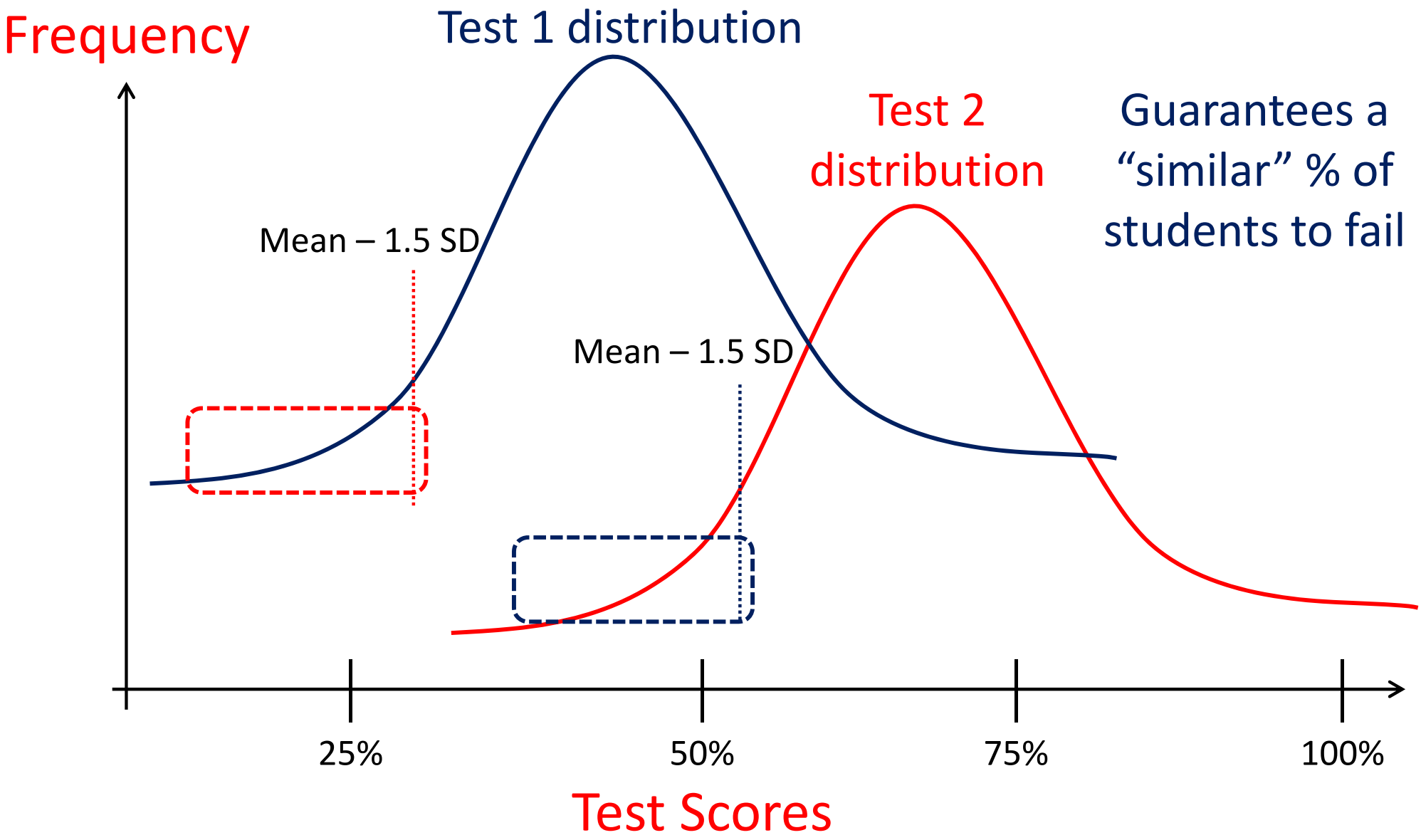
- "Criterion-based" passing standard
- Passing standard agreed by subject matter experts

THE UNIVERSITY OF ILLINOIS
COLLEGE OF MEDICINE
CHICAGO PEORIA ROCKFORD URBANA

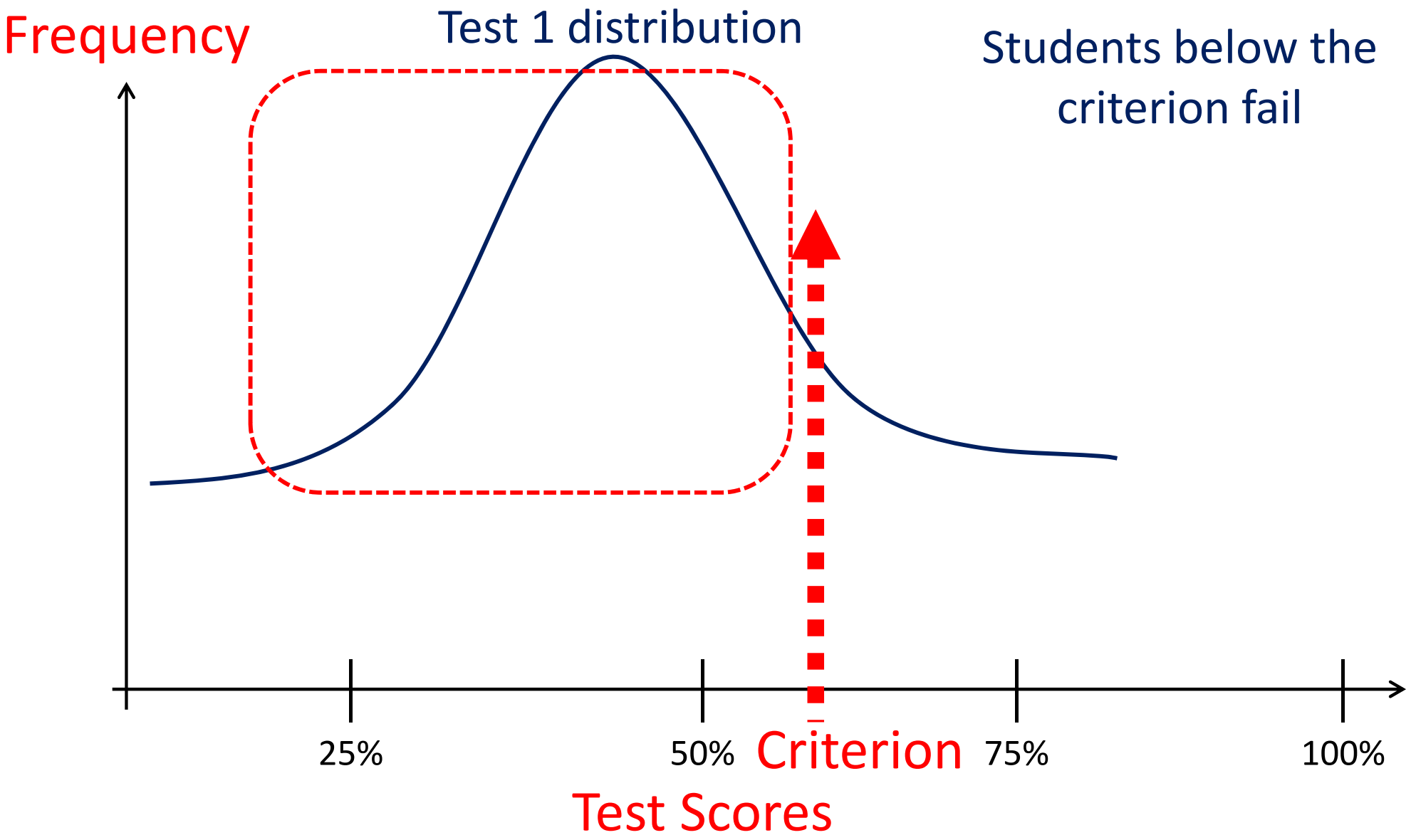
# Problems with "Fixed" Cut Score



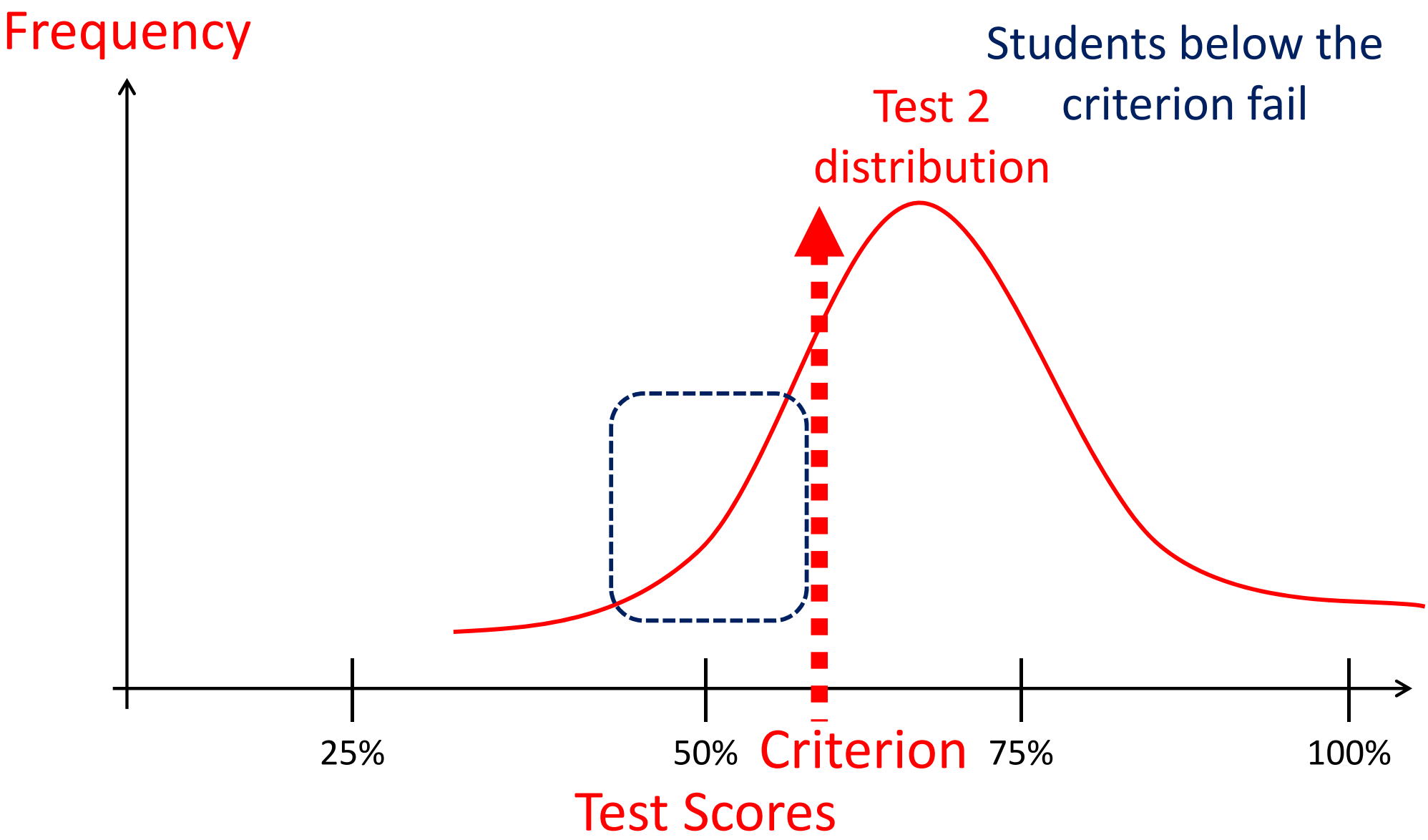
# Normative standard



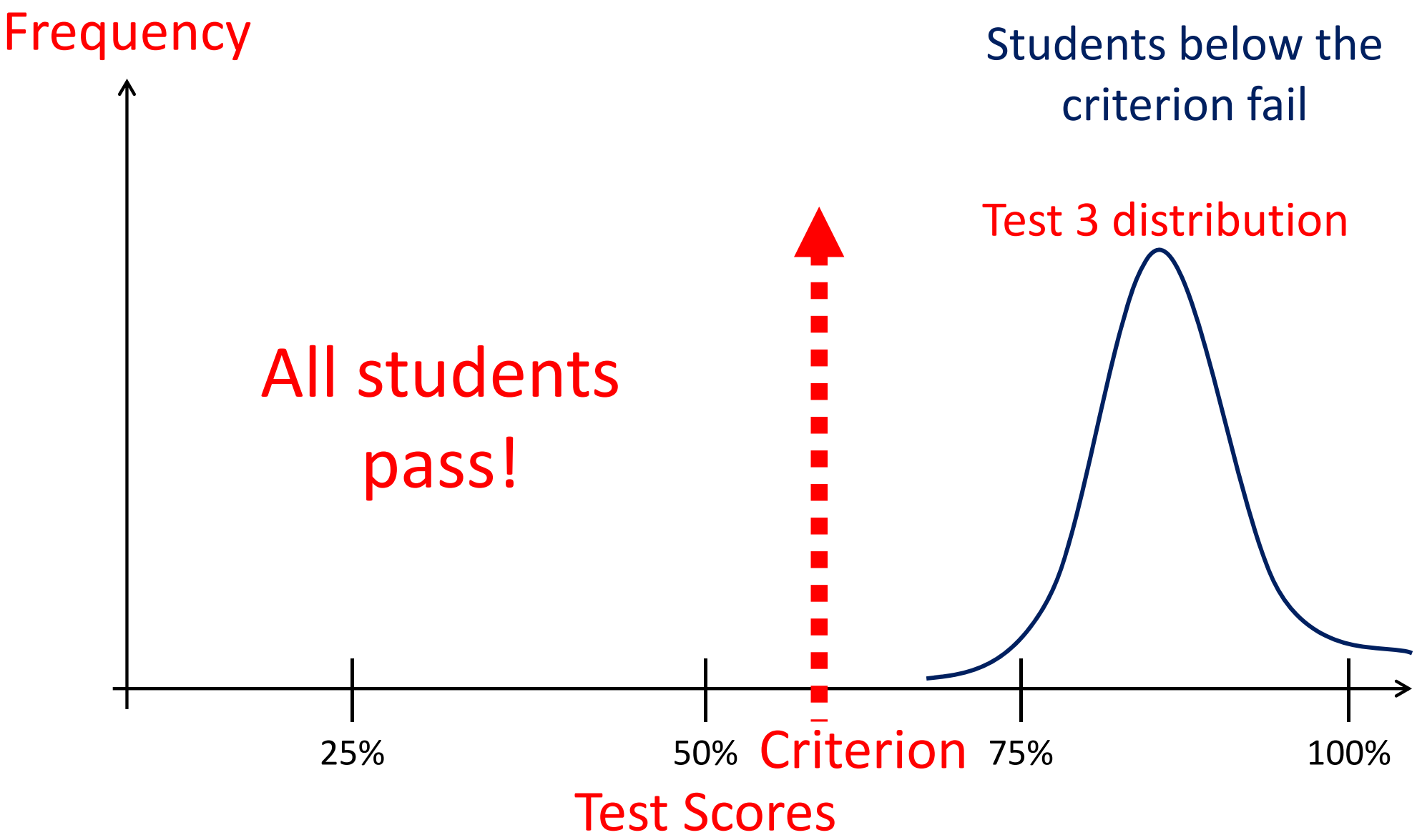
# Criterion-based standard (1)



# Criterion-based standard (2)

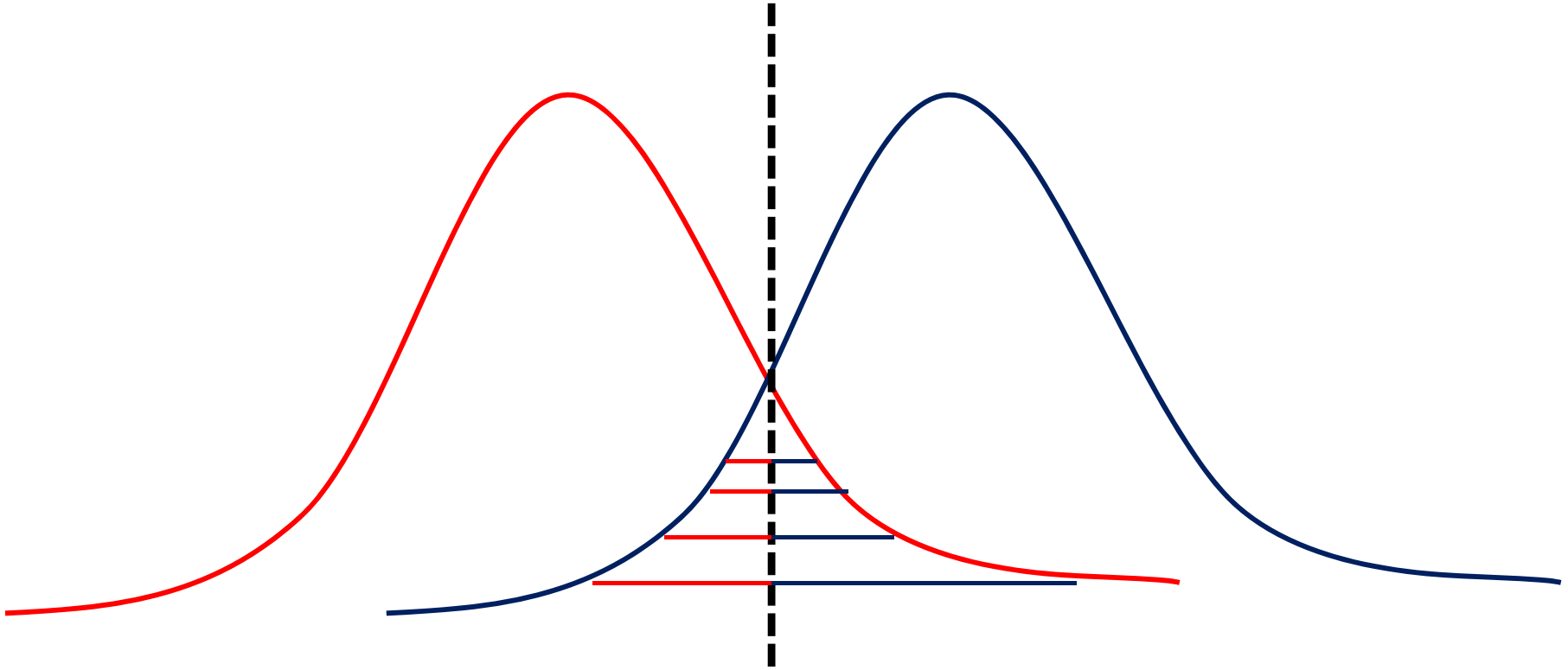


# Criterion-based standard (3)



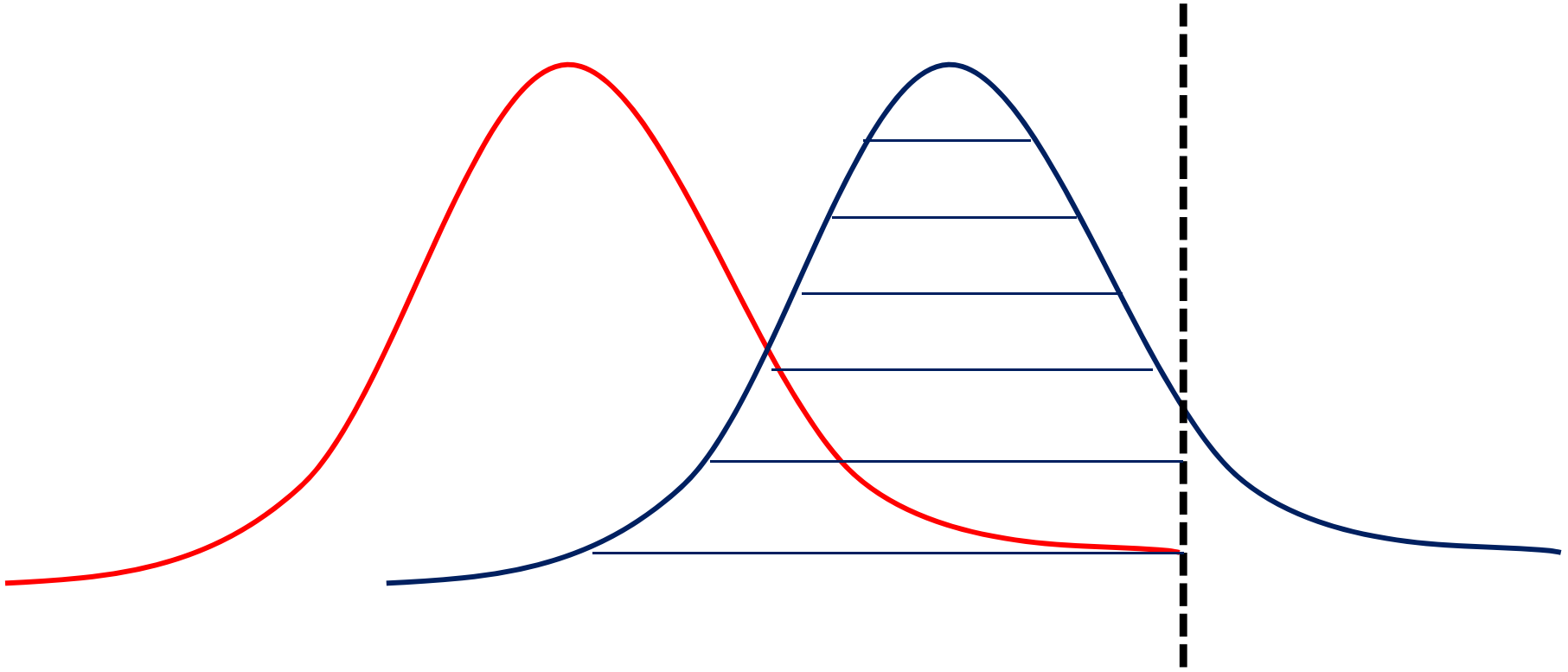
# Classification Error (1)

Red = True Fail, Blue = True Pass



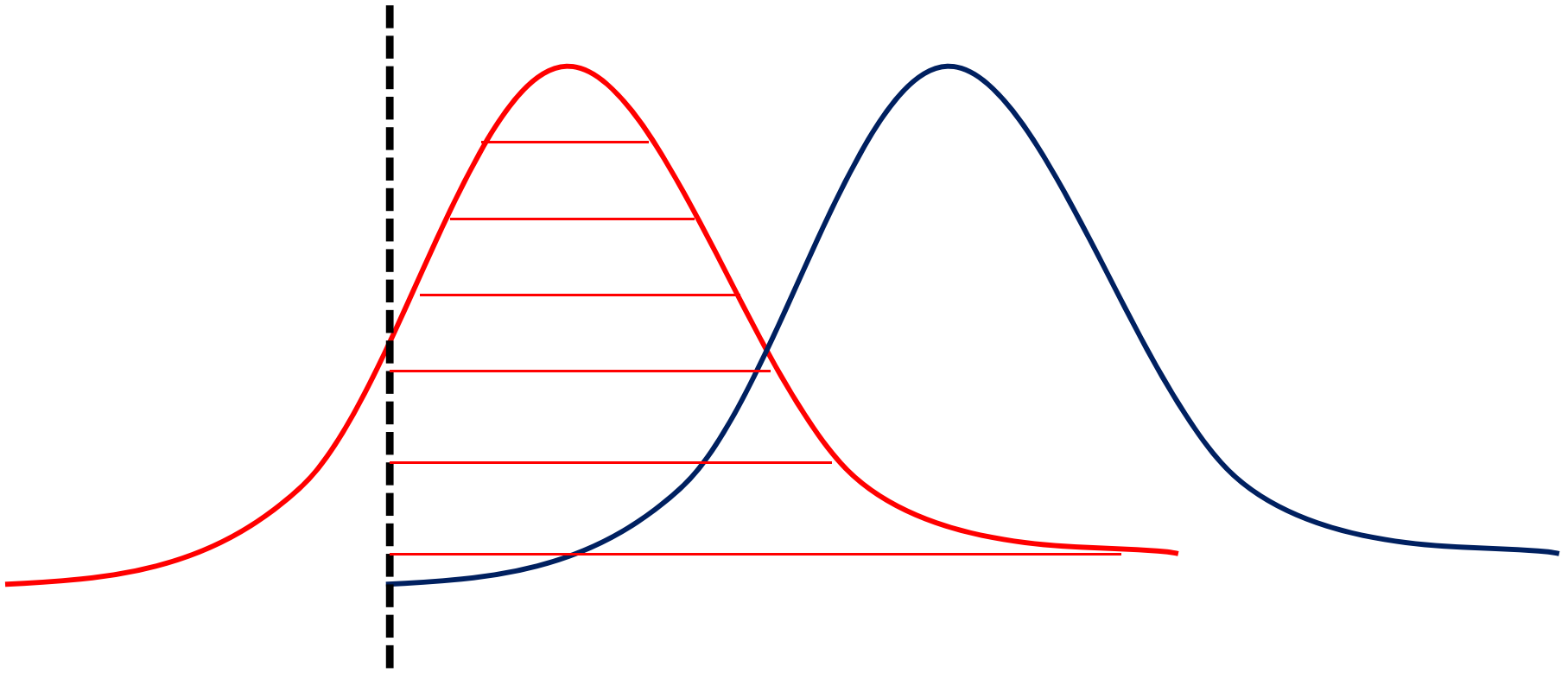
# Classification Error (2)

More **'True Pass'** Will Fail



# Classification Error (3)

More **'True Fail'** Will Pass



# "Borderline" Performance (1)

- Cut scores based on "borderline" performance
- Minimally competent!

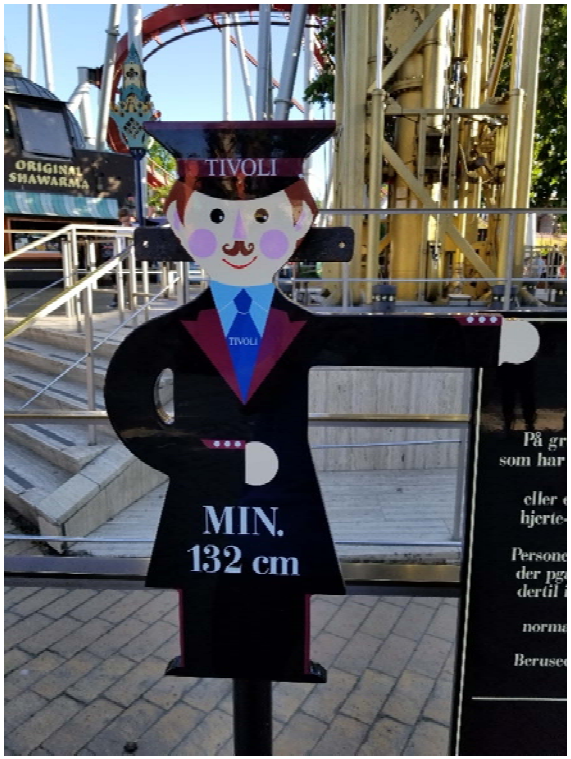
- Best performance that is still basic
- Worst performance that is still proficient

Borderline / Minimally Competent Student

→ at least know the following…

→ types of mistakes that are forgivable errors…

# "Borderline" Performance (2)





- Worst surgeon who should be licensed
- Worst pilot who should be allowed to fly
- Shortest child who is tall enough to ride the roller coaster
- Worst barber who can cut your hair

# Angoff Method (1)

- Probability Angoff
  - What is the probability a borderline student would accomplish each item?
  - 0 to 1 scale (e.g., 0.30, 0.60)
  - Sum probability across items = score needed to pass case
- Other variants also exist

# Item-Based Method: Angoff Method(2)

What is the probability a borderline student would accomplish this item?

| Item | Probability Angoff |
|---|---|
| 1 | .25 |
| 2 | .75 |
| 3 | .50 |
| 4 | .50 |
| Sum = 2/4 = passing score=50% | |

# Item-Based Method: Angoff (3)

## Judge 1

| Item | Probability Angoff (What is the probability a borderline student would accomplish this item?) |
|---|---|
| 1 | .25 |
| 2 | .75 |
| 3 | .50 |
| 4 | .50 |
| Sum = 2/4 = passing score=50.0% | |

## Judge 2

| Item | Probability Angoff (What is the probability a borderline student would accomplish this item?) |
|---|---|
| 1 | .40 |
| 2 | .60 |
| 3 | .30 |
| 4 | .20 |
| Sum = 1.5/4 = passing score=37.5% | |

## Judge 3

| Item | Probability Angoff (What is the probability a borderline student would accomplish this item?) |
|---|---|
| 1 | .30 |
| 2 | .70 |
| 3 | .70 |
| 4 | .80 |
| Sum = 2.5/4 = passing score=62.5% | |

Passing Score = 50%

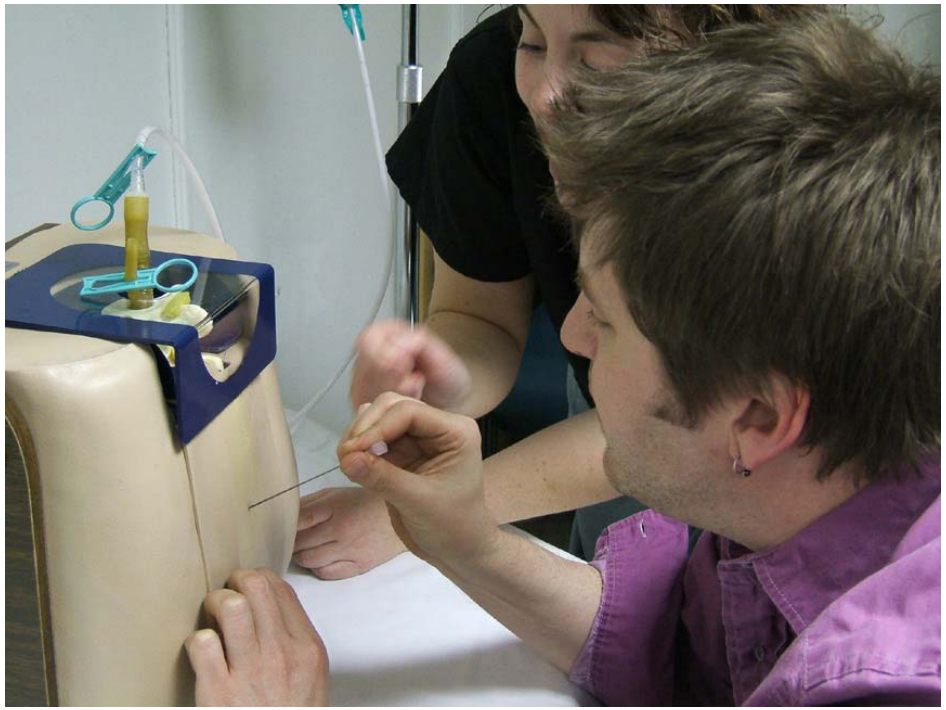
THE UNIVERSITY OF ILLINOIS COLLEGE OF MEDICINE
CHICAGO PEORIA ROCKFORD URBANA

# Pass Rates: USMLE 2014-2015

- Step 1 Basic Sciences: 85%
  - US/Canada: 94% (58% test takers)
  - Non-US/Canada: 72% (42% test takers)
- Step 2 Clinical Knowledge (CK): 85%
  - US/Canada: 94% (61% test takers)
  - Non-US/Canada: 71% (39% test takers)
- Step 2 Clinical Science (CS): 88%
  - US/Canada: 96% (58% test takers)
  - Non-US/Canada: 78% (42% test takers)
- Step 3: 94%
  - US/Canada: 98% (67% test takers)
  - Non-US/Canada: 85% (33% test takers)

Reference: http://www.usmle.org/performance-data/

# Passing score before doing a Lumbar Puncture on your child?



**50%**

**70%**

**90%**

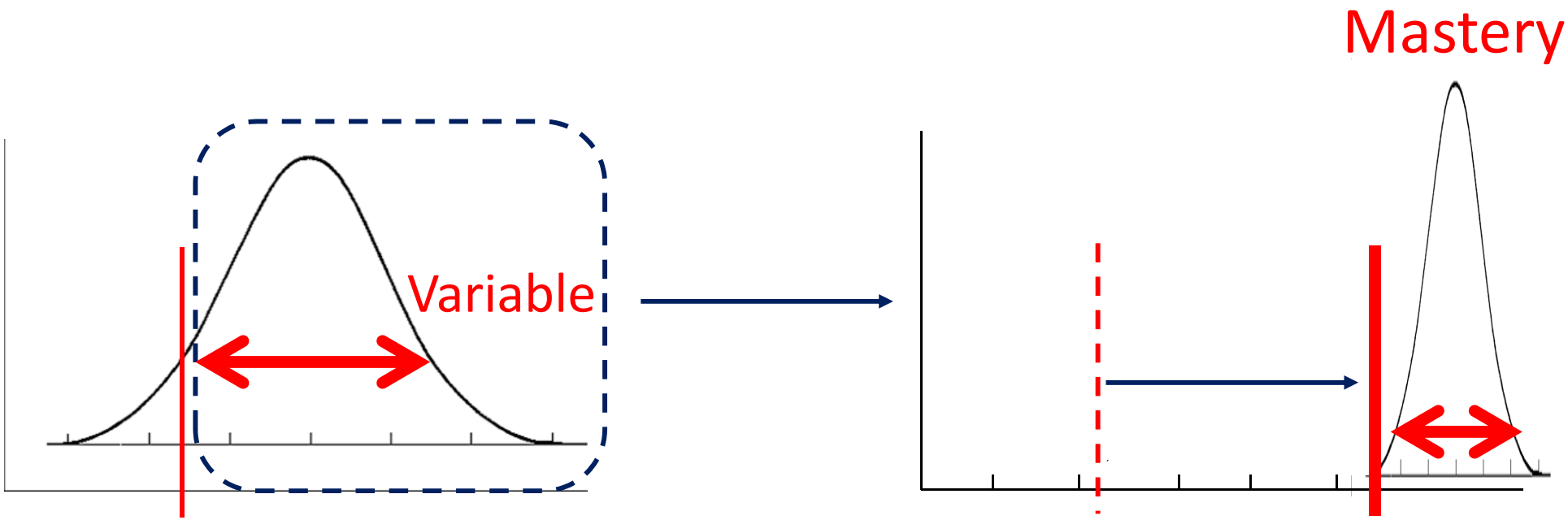
**100%**

Do you want a "borderline" student?

Do you want a "master" (well prepared) student?

"Borderline" → "Mastery"
Well Prepared

# *Adapt* Angoff for Mastery Learning standards?



# Inferences:
# Well prepared to do what???

To perform the procedure **tomorrow, with supervision and coaching**?

To perform the procedure **in 3 months, with no supervision**?

# Mastery Angoff Calculation

| Item | Probability of a ***well prepared*** student accomplishing this item after repeated practice |
|---|---|
| 1 | .90 |
| 2 | .95 |
| 3 | 1.00 |
| 4 | .90 |
| Test | Sum = 3.75/4 = passing score = 94% |

# Patient Safety Considerations: Are there Critical Items?

| Item | Impacts patient or provider safety? | Impacts patient comfort? | Impacts procedure outcome? |
|---|---|---|---|
| 1 | X | | |
| 2 | X | | X |
| 3 | | | |
| 4 | | X | |

**Critical items**

# Set standards separately and conjunctively for critical and non-critical items

## For example…

**Must pass 98% of critical items**

Example:
Items 1, 5, 7, 9, …

**AND**

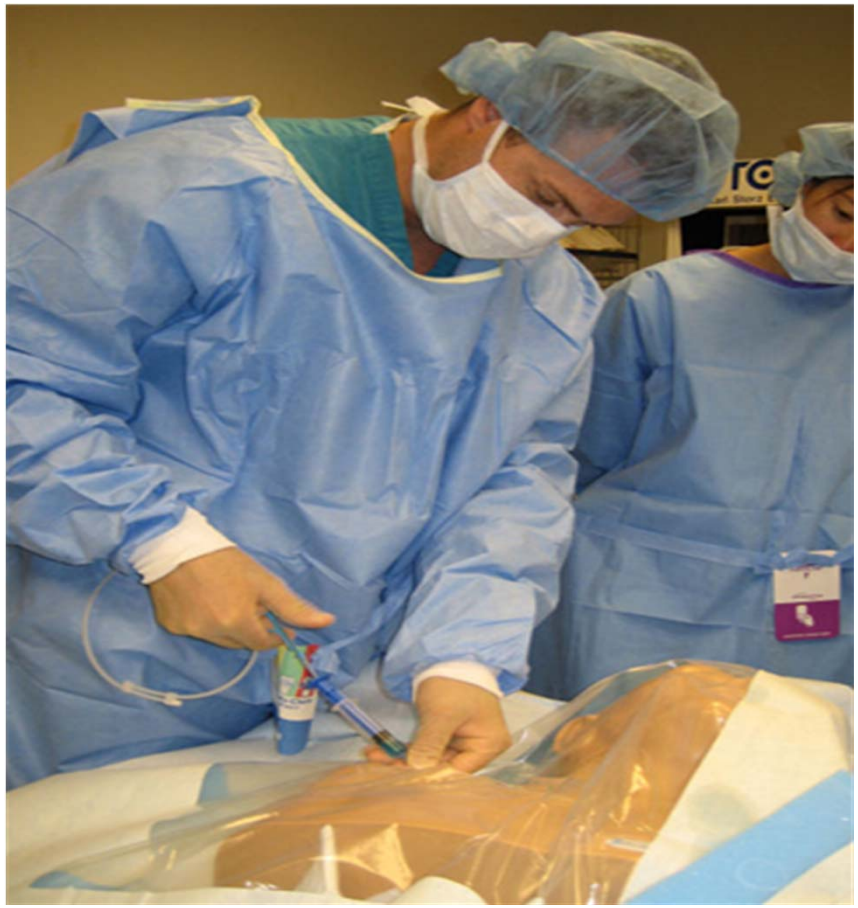
**Must pass 75% of non-critical items**

Example:
Items 2, 3, 4, 6, 8, …

# Does it make a difference?

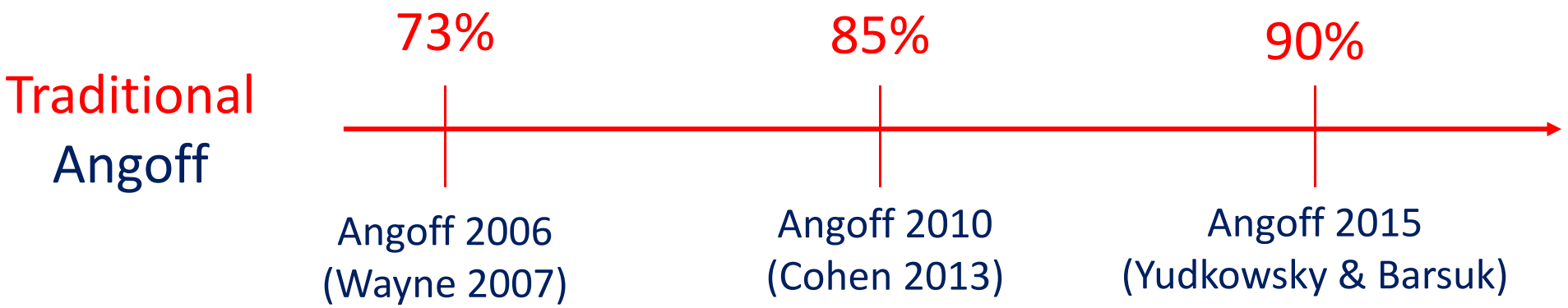
## Central Line Insertion

**29 item checklist**



# Traditional Angoff change over time?

Mastery Angoff (2015) Pass Score = 98%




Traditional Angoff – USMLE Step 1 Score Example

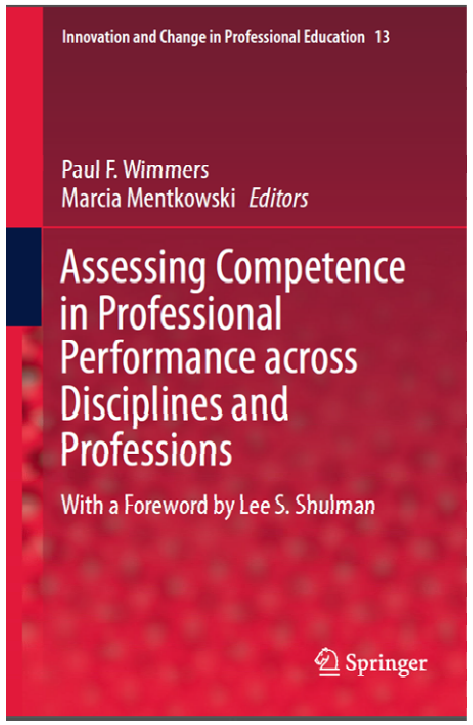
- Year 2014: 192
- Year 2010: 188
- Year 2007: 185
- Year 2001: 182

Increase in Step 1 Passing Scores Over Time

# Innovative Curricular Applications:

1. EPAC Program
2. Orthopedic Surgery

# Paradigm Shift in Medical Education


**Chapter 19**
**Evaluating the Paradigm Shift from Time-Based Toward Competency-Based Medical Education: Implications for Curriculum and Assessment**

Yoon Soo Park, Brian D. Hodges and Ara Tekian

**Abstract** In the early twentieth century, most curricula were based on a concept of fixed time. Students who successfully completed a program were judged to be competent. However, a paradigm shift toward competency-based education occurred at the end of the twentieth century, allowing only students who are judged "competent" to move forward in a professional school curriculum. There are significant implications to this paradigm shift, particularly for curricular design, performance assessment, faculty development, and resources. Educators may find challenges addressing individual learning differences—some students are able to progress easily in some subject areas, while some may continue to struggle. Learners can also progress at different rates in competency-based education programs. While it is relatively easy to develop competencies in areas of knowledge

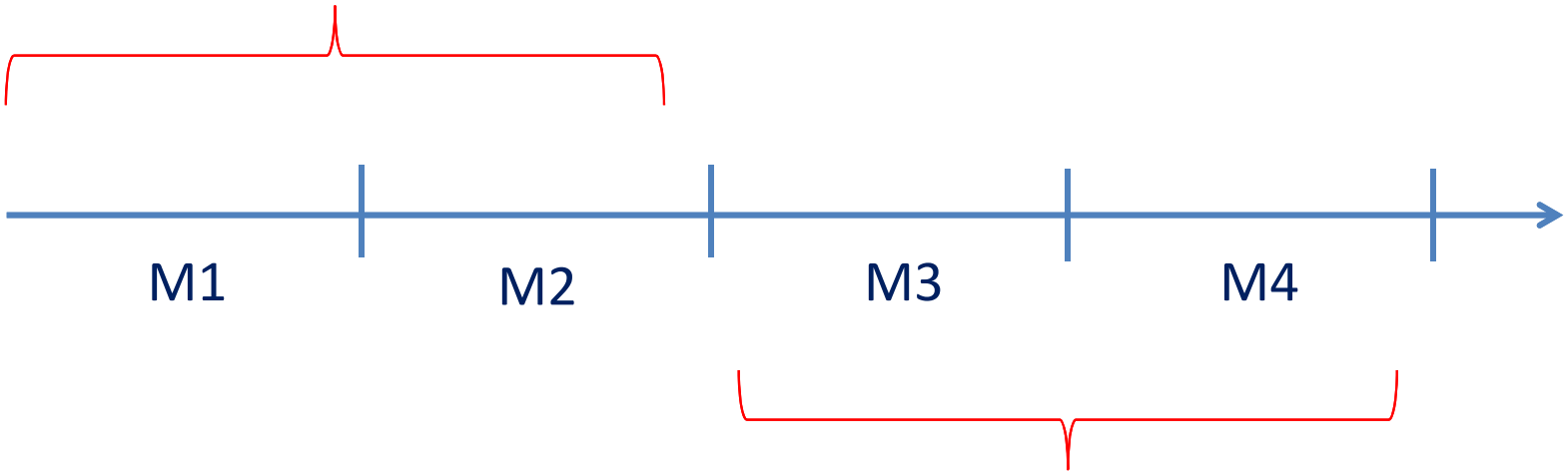
# Current Curriculum: Time-Based Model

**Preclinical and Clinical Years**

Basic Science,
Clinical Knowledge courses

USMLE Step Examinations

- Step 1
- Step 2 Clinical Knowledge (CK)
- Step 2 Clinical Skills (CS)



# Is "Time-Based" Model *still* appropriate?

Flexner Report (1910)

→

100+ Years Later??

The NEW ENGLAND JOURNAL of MEDICINE

REVIEW ARTICLE

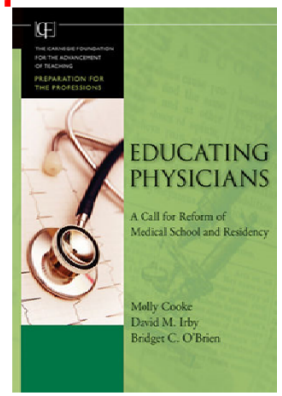
MEDICAL EDUCATION

Malcolm Cox, M.D., and David M. Irby, Ph.D., Editors

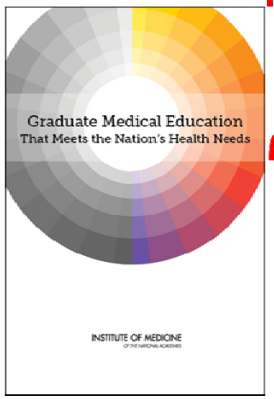
American Medical Education 100 Years after the Flexner Report

Molly Cooke, M.D., David M. Irby, Ph.D., William Sullivan, Ph.D., and Kenneth M. Ludmerer, M.D.

American Reactions




2010




2014

Canadian Reactions


2010

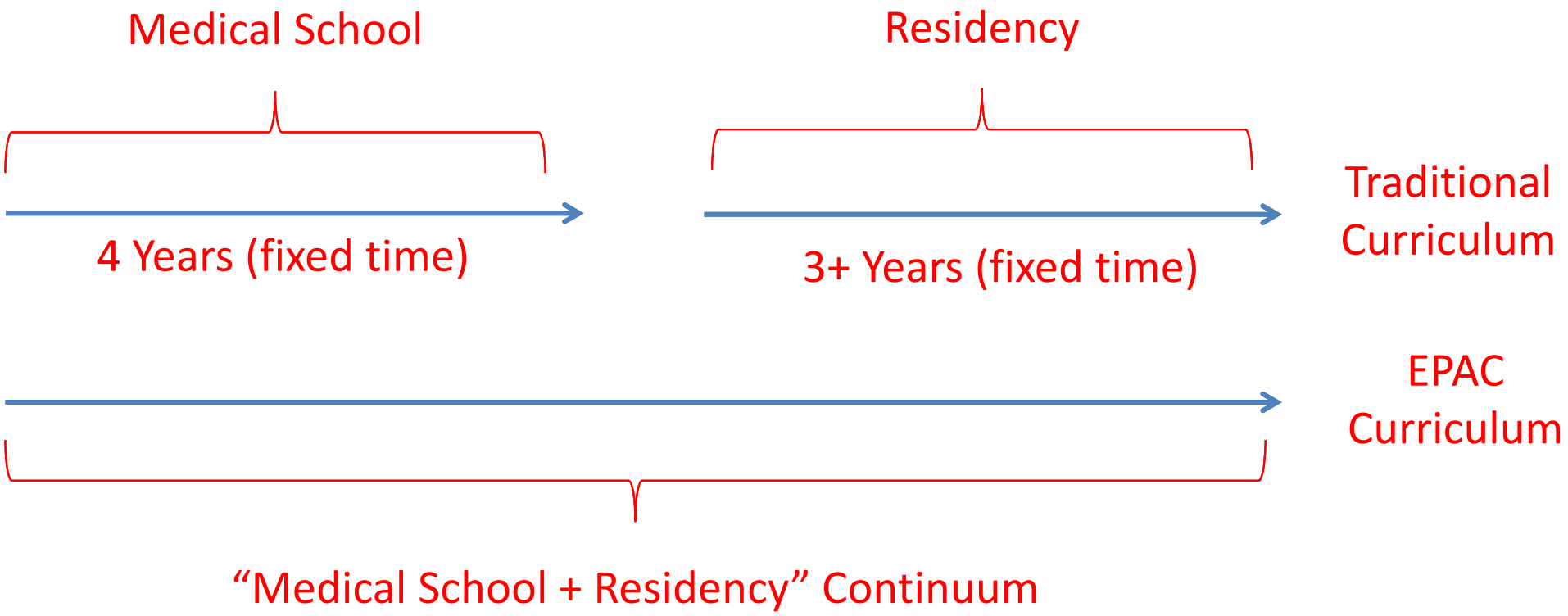

2012

# Education Continuum: EPAC (1)

**Education in Pediatrics Across the Continuum (EPAC): First Steps Toward Realizing the Dream of Competency-Based Education**

Andrews, John S. MD; Bale, James F. Jr. MD; Soep, Jennifer B. MD; Long, Michele MD; Carraccio, Carol MD, MA; Englander, Robert MD, MPH; Powell, Deborah MD; for the EPAC Study Group

Academic Medicine: Post Acceptance: October 11, 2017
doi: 10.1097/ACM.0000000000002020



# Education Continuum: EPAC (2)

- Competency-based, time-variable progression

- Five medical schools
  1. UC San Francisco
  2. University of Colorado
  3. University of Maryland
  4. University of Minnesota
  5. University of Utah

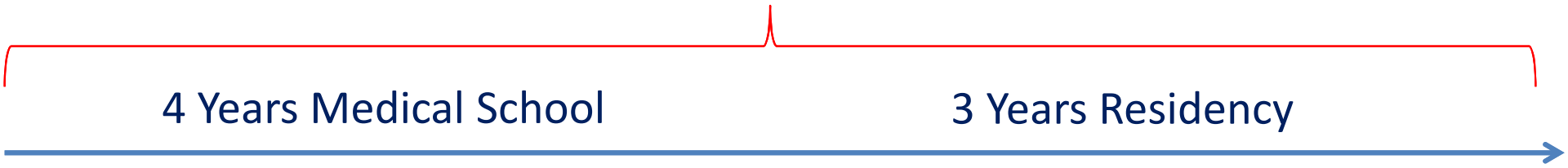
Four cohorts per school
Years 2013 – 2016

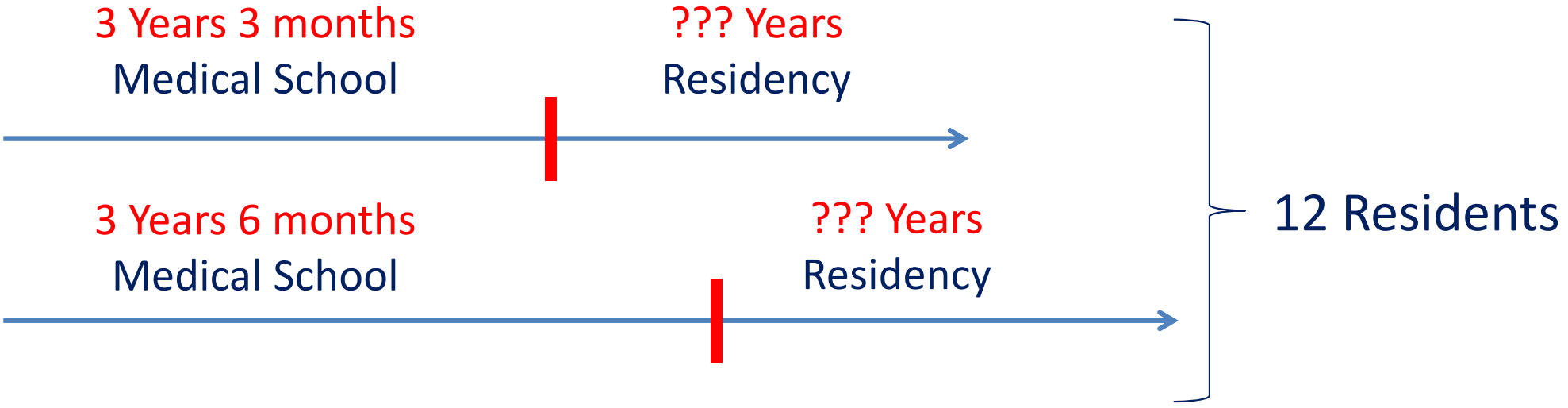
EPAC ($n$ = 48)
- 4 withdrew
- 12 in residency

- LCME approval in 2011
- Assessment based on entrustable professional activity (EPAs)
  - Unit of work, observable, and measurable
  - Example: *Enter and discuss orders/prescriptions* (EPA-4)

# Education Continuum: EPAC (3)

"Medical School + Residency"
Traditionally 7 Years in Pediatrics



# Residency – Orthopedic Surgery

**Training tomorrow's surgeons: what are we looking for and how can we achieve it?**

Teodor P. Grantcharov* and Richard K. Reznick†
*Division of General Surgery, St Michael's Hospital, and
†Department of Surgery, University of Toronto, Toronto, Ontario, Canada

ANZ J Surg 79 (2009) 104–107

Three-Year Experience with an Innovative, Modular Competency-Based Curriculum for Orthopaedic Training

Peter C. Ferguson, MD, FRCSC, William Kraemer, MD, FRCSC, Markku Nousiainen, MD, FRCSC, Oleg Safir, MD, FRCSC, Ranil Sonnadara, PhD, Benjamin Alman, MD, FRCSC, and Richard Reznick, MD, FRCSC
*Investigation performed at the University of Toronto, Toronto, Ontario, Canada*

**J Bone Joint Surg Am.** 2013;95:e166(1-6) • http://dx.doi.org/10.2106/JBJS.M.00314

- Competency-Based Residency at University of Toronto
- Began in July 2009
- Recruited 14 residents (2013)

Two residents completed in 4 Years





Canada RCPSC Orthopedic Surgery Program

5 Years Fixed Time

THE UNIVERSITY OF ILLINOIS
COLLEGE OF MEDICINE
CHICAGO PEORIA ROCKFORD URBANA

Yoon Soo Park, PhD
email: yspark2@uic.edu