Role of Assessment in the Future of Healthcare

Assessing Healthcare Outcomes:
Best Practices for Simulation-based Learning and Information Technology

Center for Advanced Medical Learning and Simulation (CAMLs)
Tampa, Florida

8-9 June, 2013

Richard M. Satava, MD FACS

Presenter Financial Disclosure Slide*

Financial Support: None (... but still hoping!)
Consulting: Ethicon Endosurgical
Grant Support Intuitive Surgical, Inc
Investment *InTouch Technologies, Inc

* There will be no discussion of products from these companies
Definitions *(OED*)

1. **Assessment:** the *measurement* of a learner’s potential for attainment, (or their actual attainment) . . . of performance

2. **Test:** a *procedure intended to establish* the quality, performance, or reliability of something, especially before it is taken into widespread use

3. **High Stakes Test:** a test with *important consequences* for the test-taker.

4. **Training:** the action of *teaching* a person a particular skill or type of behavior

---

1. **Outcomes Measures:** [*descriptive*] *quantifiable consequence(s)* of an action, set of actions, or procedure

2. **Metrics:** *a set of figures [numbers] or statistics* that measure results.
   a. Quantitative: *a value* or component that may be *expressed in numbers*
   b. Qualitative: *distinctive [unambiguous] attribute* or characteristic possessed by something

---

* Oxford English Dictionary
Outcomes Measures and Metrics

Outcomes Measures
- The prime determinant of the entire educational process
- Set by key stakeholders in training & certification (societies, boards, etc)
- Measures include correct and incorrect (errors)
- Must be unambiguous, measurable, relevant and practical

Metrics
- Must support an outcomes measure (“no measurement, no metric”)
- Useful both in training and assessment (formative feedback)
- Used in generating final results reporting (summative feedback)
- Applicable to high stakes testing
- Quantitative whenever possible – accurately measurable or binomial
- If Qualitative, unambiguously defined (for IRR $\geq 0.80$)

* Once determined, can be used by industry to build a simulator

Types of Assessment

1. Diagnostic: measures the learner's achievement on entry to their course of study [baseline] in order to identify their individual learning needs and their strength . . . [and] can be carried out through aptitude tests

2. Formative: ongoing process . . . which takes place throughout the learner's course of study and provides them with the [immediate] feedback and guidance necessary to enable them to improve their performance

3. Ipsative [iterative]: learners are assessed against their own previous level of attainment.

4. Summative: takes place at the end of their course of study, and measures the learner's attainment against the specified learning objectives [outcomes measures] of the syllabus or program
Fundamental Principles

Objective Metrics

and

Training to Proficiency

Fundamental Principles

Training and Assessment

are

Two sides of the Same Coin
**Setting Measures, Metrics and Assessment**

**Consensus Conference Methodology**

1. Subject matter Experts (SME) chosen on scholarly activity (literature review, publications, workshop participation in standards/curriculum, conference presentations, etc.

2. Consensus Conference convened as 2 day workshop

3. Use of Modified Delphi method (communal brainstorming) to identify elements (outcomes, tasks, etc) during consensus conference

4. Classic Delphi method (anonymous) afterwards (via email) to refine element matrix, with multiple iterations until consensus achieved

---

**Subject Matter Experts**

**Profession of the Participants**

1. Clinical nurses active in curriculum development

2. Medical Educators

3. Behavioral Psychologists, Psychometricians and Statisticians

4. Representatives of: Nursing societies, international nursing societies, certifying nursing boards, federal regulators in accreditation

5. Engineers, computer programmers, simulation experts

*(If building a device to support the training)*
Standardized Curriculum (Course)

Suggested template

- Goals and Metrics of the Simulation
- Anatomy
- Steps of the Procedures
- Errors

TEST
- Skills Training
- Outcomes Reporting

 Asset Curriculum Template

(Adopted for the FRS curriculum development)

- Needs Assessment (Requirement)
  - Inventory and Gap Analysis
- Goals and Objectives
- Outcomes Measures
  - Includes Societies, Boards, ACGME, etc input to determine outcomes measures
  - Benchmarked to criterion measures by experienced/expert surgeons/physicians
  - Criterion - unambiguously defined with quantitative measures
- Didactic Component (Cognitive skills)
  - Pre-didactic suggested readings
  - Anatomy or laboratory model description
    - Task analysis to develop procedural steps & assessment tools
    - Steps of Procedures
      - Skills and Techniques
      - Supplies, equipment and set-up
      - Errors - unambiguously described with quantitative measures
- PRE-TEST (before skills training begun)
- Technical Skills Training (Psychomotor skills AND team training skills)
  - Formative assessment with feedback
  - Summative Assessment with feedback
  - Competency - based set by repetition until benchmark criteria met – 2 consecutive trials
- Debriefing
  - Review errors and learning curve
  - Outcomes Measures Results Reporting
Errors

Definition
“the state or condition of being wrong in conduct or judgment” - OED
“a deviation from accuracy or correctness” (does not imply “fault”)

Characteristics
• Very difficult but important to unambiguously define
• May be individual or ‘system’ and overt (immediate) or latent
• Can be with or without “consequences” (James Reason “Human Error”)
• Not all errors are relevant – define minor and critical errors
• Mistake - a fault from misjudgment, carelessness, or forgetfulness

Application
• Arguably the most important measure of skill or judgment
• Fundamental principle behind patient safety (No errors)
• Must train student in most common errors (avoid, recognize, remediate)

Methodology for Consensus Conference

1. Modified Delphi Method (in person)
   a. Select subject matter experts (SME)

2. Classic Delphi Method (via email)
   a. Includes participants from each conference (cumulative)

3. Public Forum (when indicated)
Methodology for Full Life-cycle Curriculum Development

Criteria Consensus Conference

1. Outcomes Measures (Conf #1)
2. Curriculum Development (Conf #2.0 & 2.5)
3. Validation Study Design (Conf #3)
4. Validation Trials
5. High-stakes Testing (HST) (Conf #4 – Separate from Curriculum)

Full Life Cycle Development

The Metrics Drive the Process

<table>
<thead>
<tr>
<th>Who</th>
<th>How</th>
<th>What Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certifying Boards &amp; Specialty Societies</td>
<td>Consensus Conference</td>
<td>Outcomes Measures Curriculum Development Train &amp; Assess</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Curriculum Development</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulator Development</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation Studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High Stakes Testing for Certification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Issue Certification</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Issue Mandates And Certificates</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS certifier</td>
</tr>
</tbody>
</table>

Outcomes Measures: Curriculum Development Train & Assess
Curriculum Development:
- Standard Curriculum Template
- Engineering Physical Simulator
- Standard Validation Template

Validation Studies:
- Current Procedures

High Stakes Testing for Certification:
- HST independent Organization
- SAGES-FLS

Issue Certification: ABS certifier
Role for developing metrics

Each component has a different need for assessment and metrics

<table>
<thead>
<tr>
<th>DEVELOP CURRICULUM</th>
<th>VALIDATE</th>
<th>HST</th>
<th>CERTIFY</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUTCOMES MEASURES</td>
<td>DIDACTIC COGNITIVE</td>
<td>PSYCHOMOTOR SKILLS</td>
<td>TEAM TRAINING</td>
</tr>
</tbody>
</table>

Metrics need input by training & certifying or licensing bodies

Metrics developed for the outcomes need to be consistent across the different components of the curriculum development

Benchmarks for metrics are set by experts at beginning of trial

Metrics set by independent testing body

If HST results poor - feedback changes metrics

Outcomes Measures methodology

Defining the required skills and tasks

Task Deconstruction - Task Analysis

Methodology for identifying outcomes measures

<table>
<thead>
<tr>
<th>Cognitive Skills</th>
<th>Psychomotor Skills</th>
<th>Team Training Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify structures of Cauda Triangle</td>
<td>Laparoscopic Inserting</td>
<td>Position Instrument parts</td>
</tr>
<tr>
<td>Common Duct, Cystic Duct, Cystic Artery</td>
<td>Turn the screw, Does not break, Knot tight</td>
<td>Communication between surgeon, Assistant and nurse</td>
</tr>
<tr>
<td>Must identify all 3 structures</td>
<td>Surgeon should first throw</td>
<td>Assistant receives tool or from nurse</td>
</tr>
<tr>
<td>Not identifying all 3 structures</td>
<td>1st throw tight 2nd throw tight</td>
<td>Assistant confirms site with surgeon</td>
</tr>
<tr>
<td>Question/answer identify on image</td>
<td>Surgery breaks knot, type of Mg force applied</td>
<td>Tool cannot be used at any point of flow</td>
</tr>
<tr>
<td>Broken instruments, score one out of nature, apply 0.5 kg force</td>
<td>Measured Time Observed only</td>
<td>Observation time through video analysis</td>
</tr>
<tr>
<td>Common Errors</td>
<td>Assessment Tools</td>
<td>Assessment Tools</td>
</tr>
<tr>
<td>Tasks and Sub-tasks</td>
<td>Measurement Methodology</td>
<td></td>
</tr>
<tr>
<td>Outcomes Measures</td>
<td>Quantitative or Qualitative</td>
<td></td>
</tr>
<tr>
<td>Metrics</td>
<td>Common Errors</td>
<td></td>
</tr>
<tr>
<td>Assessment Tools</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Tools
Principles of psychomotor skills exercise design

1. Multiple learning objectives that incorporate multiple tasks
2. Agnostic to type of device
3. Easy to administer (to insure Inter-Rater Reliability (IRR))
4. Cost effective (multiple tests/module)
5. High fidelity or real physical models for test mode. Training can be lower fidelity
6. VR and physical models are derived from same CAD/CAM model.
7. Preference for tasks that have existing evidence of validity

<table>
<thead>
<tr>
<th>Metric</th>
<th>Running suture</th>
<th>Dome with four towers</th>
<th>Vessel dissection + clipping</th>
<th>UTSW 4 arm extraction + cutting</th>
<th>Energy and Mechanical cutting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye hand instrument coordination</td>
<td>S</td>
<td>P</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Needle holding</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atraumatic handling</td>
<td>S</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>Safety of operative field</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camera</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clutching</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissection (Fascia/Blunt)</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Docking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knot tying</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrument exchange</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutting</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy sources</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign body management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suture handling</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrist articulation</td>
<td>P</td>
<td>P</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Multi arm control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clip applying</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Psychomotor skills matched to psychomotor tasks

<table>
<thead>
<tr>
<th>Skill</th>
<th>Task</th>
<th>Running suture</th>
<th>Dome with four towers</th>
<th>Vessel dissection + clipping</th>
<th>UTSW 4th arm retraction + cutting</th>
<th>Energy and Mechanical cutting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye hand instrument coordination</td>
<td>S</td>
<td>P</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Needle Driving</td>
<td>P</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>Atraumatic handling</td>
<td>P</td>
<td>P</td>
<td></td>
<td></td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>Safety of operative field</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Clutching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Dissection (Fine/Blunt)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Docking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Knot tying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Instrument exchange</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Cutting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Energy sources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Foreign body management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Suture handling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Wrist articulation</td>
<td></td>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Multi arm control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Clip applying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
</tbody>
</table>

P = primary skill that task measures  S = secondary skill that task measures

### Developing the Didactic Section from the skills/tasks

*From outcomes measures to didactic portion of curriculum*
# Curriculum Development

## Didactic Knowledge Skills (Sample)

<table>
<thead>
<tr>
<th>Skill Title</th>
<th>Description</th>
<th>Desired Presentation Format</th>
</tr>
</thead>
</table>
| Trocars placement | • Correct choice of trocar  
• Correct alignment of robotic arm  
• Correct position of Ports placement  
**Errors**  
• trocar entrance injury  
• port-site injury  
• Not checking for injuries after placement  
• Tip of the trocar not visualized during insertion  
• Inserting trocar in areas of previous scars or incisions  
• incorrect position, spacing and location,  
• incorrect insertion depth | Video demonstrations of safe use of open cutdown, Veress needle, and Optiview techniques. Ideally video showing injuries occurring  
Video of arm collisions at the bedside due to inappropriate trocar placement  
Video or picture showing injury to port site when port not inserted appropriately  
Images of correct and incorrect port positions (outside view and inside) |

---

## Psychomotor Skills development

**Identical simulation for all methods**

- Real objects for skills on real robot
- Virtual objects for skills on VR simulator
- CAD/CAM
- "Backpack"

---
Psychomotor Multi-Skill Device Design

Includes same skill set as Fundamentals of Laparoscopic Surgery - FLS

Consists of 7 specific Tasks that include all 25 Skills

Unique dissection tasks

3-dimensional platform (Dome shape)

Includes all 7 tasks on single device

Real device created from virtual prototype

Very low cost, compatible with other simulations

Software version is open source

Team Training and Communication (Sample)

Checklist 1: Pre-operative
Checklist 2: Robotic Docking
Checklist 3: Intraoperative (see below)
Checklist 4: Undocking
Checklist 5: Debriefing

Checklist 3: Intraoperative Checklist (Pauses at Critical Steps in the Procedure and time-based - hourly)

• Is there “call-out” team communication concerning instrument usage and transfer?
• Are all foreign objects accounted for (i.e. white boarding) and removed?
• Are the periodic checks occurring to discuss case progression, team member continuity, and other issues?
• Has there been regular communication with anesthesia?
Validation Study Design Conference

<table>
<thead>
<tr>
<th>Choose “Validity” Types</th>
<th>Outcomes &amp; Metrics</th>
<th>Design Validation Study</th>
<th>Choose Participating Institutions</th>
<th>Determine Benchmark Metrics</th>
<th>Conduct Multi-institutional Trial</th>
<th>Analyze &amp; Publish Results</th>
</tr>
</thead>
</table>

- Design Validation Trial Criteria
  - Define the validation types (Face, content, concurrent construct, predictive)
  - Define the validation measures
  - Design the study (RCT?)

- Design Multi-Institutional Study
  - 10 independent sites (ACS-AEI accredited Institute)
  - Faculty in at least 3 specialties
  - Establish benchmark proficiency criteria by faculty
  - Conduct the trial
  - Analyze and publish results

Assessment Tools
Crucial Principles

- Choose assessment strategy based upon type of behavior and context for rating
- Design the assessment tool to match the fidelity of the training (simple vs complex)
- Training and assessment are two sides of the same coin
- Use formative assessment and feedback as much as possible

Task

Formative assessment and immediate feedback
Assessment Tools
Essential First Steps

• Definitions - identify and define which behaviors are to be measured
• Outcomes Measures - Describe unambiguously the DISCRETE behaviors (must describe beginning and end of each event)
• Metrics – quantify the behaviors/measures

Assessment Tools
Types of measurements

• Quantitative measures – have numerical measures (metrics)
  - Easy to score
  - Computer-based whenever possible
• Qualitative measures - have descriptive measures
  - using Likert or similar scales (eg, Global Rating, GEARs, etc)
  - must use unambiguous descriptors*
  - more difficult to score to obtain IRR ≥ 0.80
• Applies to both correct actions and errors

* Terms such as “poor-average-excellent” or “adequate-inadequate” are too ambiguous
Assessment Strategies
Direct Observation

• Observe/rate - in real time or video review
  Direct observation of behavior
  - Video is easier to review – ideal for sim lab setting and MIS/image guided procedures – can pause during assessment
  - Open procedures (esp on patients) not often video taped thus less flexibility for assessment

• Comparative observation - both Expert and Novice (3-5 of each)
  - difference will highlight critical behaviors (and different techniques of Expert & Novice) and will help define errors
  - may help infer judgment, decision making etc (cognition) - Eye tracking ?

Assessment Strategies
Frequency Measures

• Often binomial – either occurred or did not occur

• Discrete time frame - most valuable when events occur in short time frame (usually seconds) - [see discrete event measures]
  - Some difficulty for “ongoing” or repeated identical events (eg coagulation, dissection, suturing, etc)
  - Good for assessing progress over time (between trials, as in learning curve)
  - Can also give direct quantitative measure (amount of time, pressure, etc)
Assessment Strategies
Discrete Event Measures

• Entire event is discrete—(correct/incorrect performance of entire task
  - one score for entire event/task, not by minute or number of times
  - must have clear beginning and ending (choose correct suture, tying a knot,
    # of sutures/inch, etc)
• Very good approach for checklist

Assessment Strategies
Interval Recording Measures

<table>
<thead>
<tr>
<th>Behavior</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6 \ldots \textit{n} (\text{min})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply cautery</td>
<td>0</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>0</td>
<td>x</td>
</tr>
<tr>
<td>Burn wrong tissue</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clip cystic duct (prox)</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clip misapplied</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clip cystic duct (distal)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clip misapplied distal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cut between clips</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cut common duct</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
**How to Define/score Errors**

**Identify**
Consensus Conference of “experts”

**Define**
Unambiguous
Test for Inter-rater reliability \( \text{IRR} \geq 0.80 \)

**Observation**
Evaluate specific interval (eg q 1 min)
Score “X” for each error each minute
Insure \( \text{IRR} \geq 0.80 \) (repeat if necessary)

### ERROR

<table>
<thead>
<tr>
<th>ERROR</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>. . .</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-visualize Triangle of Calot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>7 . . .</td>
<td>Poor quality of dissection</td>
</tr>
<tr>
<td>Non dissect cystic a.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-isolate cystic d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>7 . . .</td>
<td>Jct of cystic-common duct not visualized</td>
</tr>
<tr>
<td>Clips on cystic duct crossed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>Dropped 1 clip, crossed the double clip</td>
</tr>
</tbody>
</table>

**Assessment Strategies**

### Challenges

- **Hawthorne Effect**
  - if learner knows they are being assessed, it affects their performance

- **Assessment drift**
  - over time, assessor ‘drifts’ from original criteria, thus need to retrain the assessor

- **Assessor bias**
  - if a trainer, lenient (teaching mindset), but if an assessor, more stringent (criteria mindset)

- **Assessor objectivity**
  - must assess what they see, not what they think learner is capable of

**Although training and assessment are two sides of the same coin, there are different requirements and mindsets for trainer (mentor) and assessor (proctor)**

**NOTE:** During the Pilot Trial of assessment of performance of tasks – use group who created the initial task analysis and definitions
Setting the Benchmark

Benchmark criterion

Setting Benchmark Criteria for Any Curriculum

* Applies to all aspects of curriculum including didactic, team training and psychomotor skills

Dreyfus and Dreyfus model
New Tools and Techniques

Inferring Judgment

Can we understand what you are thinking?
**Synthetic Cadaver**

- Realistic synthetic tissues, organs etc
- Vascular system which can pump ‘blood’
- CR, MRI, US and fluoroscopy compatible

**Uses**
- Anatomy and physiology teaching
- Procedural training
- Procedure rehearsal

**Properties**

---

**Quantitative Measures**
Pre-operative Warm-up

In-situ Care

ER Hand-off OR Hand-off ICU
Choreography

From the Musical "Company"

Courtesy Anne Marie Hunter, BFA 2010

Notional diagram of choreography for Laparoscopic Cholecystectomy

Courtesy Richard Satava, MD FACS 2010

Intelligent Feedback System

EEG amplifier

Feedback monitor
Multisensory

- Envirodine Scent System
- Gunpowder
- Cordite
- Body Odor
- Garbage
- Burning Rubber
- Diesel Fuel
- Iraqi Spices

Neuropsychophysiology

- Bass Shaker Platform
- Biopac System

Immersion

And Night Vision HMD Rig...
Questions

References

Anthony Gallagher et al
K. Anderson Ericsson, et al
BREAK

Questions & Answers