Outcomes, evaluations and assessments

Adam Dubrowski
Overview

• Workshop 1: Conceptualizing simulation research
• Workshop 2: Methods used in simulation research
• Workshop 3: Outcomes, evaluations and assessments
• Workshop 4: Data collection and analyses
• Workshop 5: Dissemination
• Workshop 6: Funding
• Workshop 7: Becoming familiar with BCHSP’s Research process
Today

Lecture
• Evaluation models: Process and Outcomes 8:10 – 8:45
• Evaluation vs. Assessments 8:45 – 9:00
• Knowing a good assessment tool 9:00 – 10:00

Break 10:00 - 10:30

Independent work
• Searching for assessment tool 10:30 – 10:45
• Preparation of presentation 10:45 – 11:00

Presentations 11:00 – 12:00
  – Background
  – Question
  – Confounding factors
  – Study Design
  – Outcome Measures
Introduction

Evaluations

Process

Outcome

Assessments
Models and Frameworks

- Kirkpatrick’s Model of Evaluating Training Programs (Kirkpatrick, 1998)
- Miller’s Framework for Clinical Assessments (Miller, 1990)
- Moore’s Evaluation Framework (Moore et al., 2009)
Kirkpatrick’s Model

- Reaction
- Learning
- Behavior
- Results
Kirkpatrick’s Model

- **Reaction:** The degree to which the expectations of the participants about the setting and delivery of the learning activity were met
- **Questionnaires** completed by attendees after the activity
Kirkpatrick’s Model

- **Learning**: The degree to which participants recall and demonstrate in an educational setting what the learning activity intended them to be able to do.
- Tests and observation in educational setting.
Kirkpatrick’s Model

- **Behavior**: The degree to which participants do what the educational activity intended them to be able to do in their practices
- Observation of performance in patient care setting; patient charts; administrative databases
Kirkpatrick’s Model

- **Results**: The degree to which the health status of patients as well as the community of patients changes due to changes in the practice

- Health status measures recorded in patient charts or administrative databases, epidemiological data and reports
Kirkpatrick’s Model

• The importance of Kirkpatrick’s Model is in its ability to identify a range of dimensions that needs to be evaluated in order to inform us about the educational quality of a specific program.

• It focuses on the outcomes.

• However, it provides limited information about the individual learner.
Miller’s Framework

- Knows (Declarative Knowledge)
- Knows How (Procedural Knowledge)
- Shows How (Competence)
- Does (Performance)
Miller’s Framework for Clinical Assessments

- **Does**: Performance
- **Shows How**: Competence
- **Knows How**: Procedural Knowledge
- **Knows**: Declarative Knowledge

- **Knows**: Declarative knowledge. The degree to which participants state what the learning activity intended them to know.
- **Pre- and posttests of knowledge.**
Miller’s Framework for Clinical Assessments

- **Knows how:** Procedural knowledge. The degree to which participants state how to do what the learning activity intended them to know how to do
- Pre- and posttests of knowledge
Miller’s Framework for Clinical Assessments

- **Shows how**: The degree to which participants show in an educational setting how to do what the learning activity intended them to be able to do
- Observation in educational setting
Miller’s Framework

- **Does**: Performance. The degree to which participants do what the learning activity intended them to be able to do in their practices.

- Observation of performance in patient care setting; patient charts; administrative databases
Miller’s Framework

• The importance of Miller’s Framework is in its ability to identify learning objectives (what) and link them with appropriate testing contexts (where) and instruments (how).
Moore’s Framework

- Reaction
- Learning
- Behavior
- Results

Does
Shows How
Knows How
Knows
Assumption

• Moore’s and Kirkpatrick’s models assume a relationship between the different levels.

• If learners are satisfied, they will learn more, will be able to demonstrate the new skills, transfer them to the clinical setting, and consequently the health of patients and communities will improve!
Models assumptions are not met!

“I think you should be more explicit here in step two.”
What’s missing?

Dubrowski & Morin, 2011
Stufflebeam et al., 2002
CIPP

Context

• What is the relation of the course to other courses?
• Is the time adequate?
• Should courses be integrated or separate?
CIPP

Inputs

• What are the entering ability, learning skills and motivation of students?
• Are the objectives suitable?
• Does the content match abilities?
• What is the theory/practice balance?
• What resources/equipment are available?
• How strong are the teaching skills of teachers?
• How many students/teachers are there?
CIPP

Process

- What is the workload of students?
- Is there effective 2-way communication?
- Is knowledge only transferred to students, or do they use and apply it?
- Is the teaching and learning process continuously evaluated?
- Is teaching and learning affected by practical/institutional problems?
CIPP

Product

• Is there one final exam at the end or several during the course?
• Is there any informal assessment?
• What are the students’ knowledge levels after the course?
• How was the overall experience for the teachers and for the students?
CIPP

Methods used to evaluate the curriculum

• Discussions
• Informal conversation or observation
• Individual student interviews
• Evaluation forms
• Observation in class by colleagues
• Performance test
• Questionnaire
• Self-assessment
• Written test
CIPP

CIPP focuses on the process and informs the program/curriculum for future improvements.
Summary

• Outcome based models of program evaluation may be providing limited use.

• Need new, more complex models that incorporate both processes and outcomes.
Summary

- Assessment instruments, which feed these models are critical for successful evaluations.

- We need to invest efforts in standardization and rigorous development of these instruments.
Attributes of a Good Outcome Measure

• Choosing the appropriate measurement instrument for a specific purpose is a difficult and time-consuming task!

• Developing one is even more difficult!
Attributes of a Good Outcome Measure

The choice depends on a number of factors:

1. The proposed use of the instrument,
2. the concept to be measured,
3. the readability of the questions/scales,
4. the requirements and costs associated with the use of the instrument,
5. the burden on the participants, and
6. the measurement properties of the instruments (validity and reliability).
The SAC defined a set of eight key attributes of instruments that apply to measuring three properties:

– Distinguish between two or more groups,
– assess change over time,
– predict future status.
The SAC defined a set of 8 key attributes of instruments:
1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural And Language Adaptations
The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model

The concept to be measured needs to be defined properly and should match its intended use.
The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability

Reliability is the degree to which the instrument is free of random error, which means free from errors in measurement caused by chance factors that influence measurement.
Scientific Advisory Committee of the Medical Outcomes Trust

The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural and Language Adaptations

Validity is the degree to which the instrument measures what it purports to measure.
The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural and Language Adaptations

*Responsiveness is an instrument’s ability to detect change over time.*
The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural And Language Adaptations

**Interpretability is the degree to which one can assign easily understood meaning to an instrument's score.**
Scientific Advisory Committee
of the Medical Outcomes Trust

The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural And Language Adaptations

**Burden refers to the time, effort and other demands placed on those to whom the instrument is administered (respondent burden) or on those who administer the instrument (administrative burden)**
The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural And Language Adaptations

Alternative means of administration include self report, interviewer-administered, computer assisted, etc. Often it is important to know whether these modes of administration are comparable.
The SAC defined a set of eight key attributes of instruments:

1. The Conceptual and Measurement Model
2. Reliability
3. Validity
4. Responsiveness or sensitivity to change
5. Interpretability
6. Burden
7. Alternative Forms of Administration
8. Cultural and Language Adaptations or translations.

8. Cultural And Language Adaptations
Summary

• We strongly recommend that when selecting and using, or developing outcome instruments, researchers assess the 8 attributes carefully.

• There are a number of instruments in field of simulation which may be very good, but not many of these have undergone the process of being evaluated.

• Two criteria that **MUST** be paid attention:
  • validity and
  • reliability.
Use the following rating system when assessing the relevance of each global rating scale item and each checklist task:

1- Extremely Irrelevant
2- Irrelevant
3- Neutral
4- Important
5- Very Important

<table>
<thead>
<tr>
<th>Relevance of Global Rating Scale Items</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation for procedure (e.g., machine, machine placement, ultrasound gel, towels)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Patient Interaction</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Knowledge of Instrument/Instrument Handling</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Time and Motion</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Knowledge of Machine</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Image Interpretation</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Sterile Technique (If Applicable)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Knowledge of Procedure (If Applicable)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Troubleshooting</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Overall Performance</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relevance of Checklist Items</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informed Consent</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Proper Hand/Probe Hygiene</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Proper Patient Positioning</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Proper Patient Draping</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Landmarks Xiphoid</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Identifies Heart</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Identifies Spine</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Identifies Aorta</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Scans from Heart to Iliac Bifurcation Visualizing</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Aorta in Entirety</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Please indicate “Yes” for inclusion in the tool or “No” for exclusion.

Additional Items:

_________________________________________________________________________________________
_________________________________________________________________________________________
_________________________________________________________________________________________
_________________________________________________________________________________________
research rigor by any other name….

• Tools and measurements for qualitative inquiry address the same concerns as identified by SAC

• Work with a vocabulary that is different from quantitative approaches but reflect same principles
Clarification of terms

• Reliability = trustworthiness
• Bias = reflexivity
• Validity = authenticity
• Generalizability = transferability
Tools and Measurements

- Surveys
- Interviews - structured, semi structured, unstructured
- MCQs – forced choice
- Observations
- Focus groups
- Short answer questionnaires
- Checklists – global/binary
- Document analysis
Break

- Preparation of presentation
Searching for assessment tools

- Google scholar
- Pubmed
- MedLine
Preparation of presentation

- 20 minutes
- Use PPT
- 5+ slides:
  - Background
  - Question
  - Confounding factors
  - Study Design
  - Outcome Measures
    - Reliability
    - Validity
    - Sensitivity to change