The Role of Measurement in System and Human Performance Evaluation

Sara J. Czaja Ph.D.
University of Miami Miller School of Medicine
Overview

• Provide a brief overview of the topic
  Caveat - complex and evolving topic

• Discuss the increasingly important role of measurement in the evaluation of
  • System design and evaluation of system performance
  • Design and training interventions
  • Human performance evaluation

• Types of measures & methods for data collection.

• Criteria for measure selection.
What is Outcome Measurement

• **Systematic** determination and evaluation of the results of an activity, plan, program, process, system component, or element of behavior.

• The comparison with strategic goals, the intended or projected results, or other relevant findings (e.g. similar programs).
Role of Outcome Measurement in HSI

- Increasingly important given the increased complexity of systems and the increased demand for evidenced-based intervention protocols and system decisions.

- **Measurement is critical to understanding:**
  - Efficiency, effectiveness of a system/areas for improvement
  - Cost effectiveness
  - System safety
  - Causality of system breakdowns or analysis of accident events
  - Impact of design decisions (e.g. interface, workflow)
• **Role of Outcome Measurement in HSI**
  
  • Measurement is critical to understanding:
    • Skill levels, attitudes, user preferences.
    • Sources of human error.
    • Usability of technology/design elements.
    • Process or mechanism of action - why something works.
    • Efficiency/effectiveness of an intervention - such as training.
Impact of Outcome Measurement

• Design of systems and systems components - e.g., choice of control mechanisms, work schedules, interface.

• Ability of system to achieve strategic goals.

• Human performance - e.g., choice of performance metrics, frequency of measurement.

• Ability to provide guidance for strategic decisions - e.g., staffing, areas for improvement, choice of training program.

• Ability to detect changes associated with an intervention.
Impact of Outcome Measurement

• Ability to perform comparative analyses.
• Ability to test study hypotheses.
• Understand why something works or doesn’t work or for whom.
• Data collection protocols.
  – Data source
  – Sample size
Example

- **Comparative analysis** - evaluate the impact of two worker training programs: standard classroom vs. multimedia approach (real time webinars - lectures and animated power point, online assignments)

- **Findings** - classroom approach superior

- **Measurement strategy:**
  - Ratings by instructors - ease of implementation, etc.
  - Student ratings
  - No student performance evaluation
  - No measure of transfer of training or sustainability
  - No measure of cost or feasibility (e.g. technology)
Categories of Measures

• Wide variety of measures
• Varied ways to categorize measures
  • Role of measure
    – Screening
    – Outcome
    – Moderator
    – Mediator
  • Subjective vs. Objective, quantitative, qualitative
• Lack of consensus on optimal choice of measures
  • Cognitive abilities
  • Usability
  • Satisfaction
  • Cost effectiveness
Types of Outcomes Measures

• System Level
  – Resource utilization
  – Changes in processes (e.g., reduced wait time)
  – Time to an event (e.g., system failure)
  – Frequency of an event occurrence
  – Reductions in event occurrence
  – Near misses
  – Cost measures
  – Network metrics (e.g., communication patterns)
  – Tasks accomplished
Types of Outcomes Measures

Individual

• Performance metrics
  • Time
  • Errors
  • Types of errors
  • Skills/abilities, learning outcomes
  • Attendance records
  • Behavioral patterns

• Biological/Health metrics
  – Cortisol, EEG, Heart Rate, MRI, Galvanic Skin Response, caloric expenditure
  – Weight, BMI, Cholesterol
Types of Outcomes Measures

• Individual
  – Ratings
    • Satisfaction
    • Usability
    • Quality of life
    • Skills/abilities
    • Self-efficacy, support
Study Overview

• Cross-site Randomized Field Trial
  – PRI SM-C condition
  – Binder condition (control)

• Target Population
  – Older adults aged 65 + yrs. who live alone and at risk for isolation
  – 300 participants
    • 150 per condition
Welcome to PRISM, Michael

It is Wed, Nov 30, 2011, 09:34 AM

PICTURE OF THE DAY: IMG_0457.JPG

TODAY’S QUOTE:

"The fact is, that to do anything in the world worth doing, we must not stand back shivering and thinking of the cold and danger, but jump in and scramble through as

MIAMI, FL WEATHER BY: Yahoo! News

CURRENT CONDITIONS:
Partly Cloudy. 68F

TODAY’S FORECAST:
Mostly Sunny. High: 77 Low:53

Forecast at Yahoo! Weather (provided by The Weather Channel)
Measurement Battery

• Screening measures
  – MMSE - cognitive status
  – Vision
  – Prior Computer Experience

• Potential Moderators
  – Age
  – Technology attitudes
  – Cognitive abilities (working memory, executive function, processing speed)
  – Health status
Measurement Battery

Potential Mediators
- Perceived social support
- Size of social network
- Satisfaction with social support
- Self-efficacy

Primary Outcome Measures
- Quality of Life
- Loneliness
- Isolation

Secondary Outcome Measures
- Computer proficiency
- Attitudes
Measurement Battery

Other measures
Real time usage
Usability
Acceptability
Usefulness
Prediction of Community Activities

Neg Sx
Processing Speed
Verbal Memory
Working Memory
Executive Functioning

SSPA

UPSA

Community Activities

R^2=0.48

*p<.01
Intrusive Thoughts Partially Mediate the Effects of Psychological Suffering on Depression

**p < .01

1 Multivariate betas controlling for CG gender, ethnicity, age, education and CR MMSE, ADL/IADL
Choice of Measures

• Objectives?
  • Evaluation of a new system
  • Understanding user skill requirements

• Resources available?
  – Analytic & technical
  – Monetary

• Stage of analysis?
  – Initial design
  – Sustainability analysis

• Time constraints?

• Availability of data source(s) for required information?

• Availability of required sample?
Choice of Measures

- Stakeholders and intended use of evidence?
- Characteristics of target population?
  - Age/cohort
  - Ethnicity, race, culture
  - Experience, literacy
- Currency/relevance of the measure?
- Burden?
- Ease of Administration?
- Measurement context?
- Adherence to measurement criteria?
Choice of Outcome Measures

• Psychometric Properties
  – Measurement level (ordinal; interval)
  – Reliability
    • Test-retest (consistency in results)
    • Internal consistency (all items address the same underlying concept)
    • Inter-rater reliability (important for observational measures – two raters give the same score)
  – Sensitivity – does the measure detect differences in the intended outcome – between groups or over time within an individual
    • Ceiling effects – too easy
    • Floor effects – too difficult
Choice of Outcome Measures

- **Validity** – degree to which we are measuring what is intended in a defined population
  - Face validity (test appears to measure what it is intended to measure)
  - Content validity (includes all items necessary to represent the concept being measured)
  - Concurrent validity (agreement between measures of the same construct)
  - Construct validity (degree to which measure is capturing concept of interest)
  - Ecological validity (extent to which the findings of a research study are able to be generalized to real-life settings).
University of Miami  Computer-Based Functional Assessment Battery

• Encompasses a wide range of real world activities:
  • Medication Management
  • ATM/banking
  • Prescription refill via voice menu
  • Kiosk Ticket Purchase and way-finding
  • Online Forms Completion
  • Shopping

• High ecological validity – simulations are based on real world systems and tasks

• Computer-based/multimedia format
Functional Assessment Battery

- Real time performance data
- Remotely deliverable
- Can be easily translated for other cultures and languages
- Can be easily adaptable for alternative forms
- Does not require clinician for administration
Make a selection from the options below:

Buy EASY Card
- Use the card for three years.
- Load up to $150
- Load with cash or with a MoneyPass
- Can be used at any ATM

Buy EASY Ticket
- Use for one year.
- Can be used at any ATM

You need to provide some basic information about yourself by completing a form using the computer. Please click on the button labeled "Complete Form in English" to complete the form in English.

Click Image To Enlarge
Methods for Data Collection

Data mining
Objective measurement
Observation
Task analysis, cognitive task analysis
Verbal protocols
Questionnaires/rating scales
Observer ratings
Sensing systems/wearable technologies
Simulation
Other Factors to Consider

- Timing and frequency of measurement
- Number of measures
- Data Quality
- Data storage, security
- Data integration
- Interpretability of results
- Analytic strategy
- Establishment of cut-offs or criteria
- Weighting of measures
- Definition of constructs – what do we mean?