

Measurement & Evaluation Issues

2016 Dissemination and Implementation Short Course

**October 10-11, 2016
Madison, WI**

**Ross C. Brownson
Washington University in St. Louis**

Objectives

1. Understand the importance of measurement & evaluation issues.
2. Consider measurement issues in designing studies.
3. Learn about resources for D&I measures.

Why worry about evaluation & measurement?

Inspirational quotes:

“What gets measured, gets done.”¹

“Not everything that counts can be counted, and not everything that can be counted counts.”²

“Without a standard, there is no logical basis for making a decision or taking action”³

- Some D&I specific challenges/considerations and efforts on coming slides

¹Attributed to Peter Drucker, Tom Peters, Edwards Deming, Lord Kelvin and others ²Albert Einstein ³Joseph Juran

Core evaluation principles

- Evaluation should be considered an ongoing process, rather than a one-time, post hoc activity
- Best results are obtained when you plan for evaluation early
 - Needs assessment
 - Formative evaluation
 - Process evaluation
 - Summative evaluation (impact, outcome)

Core evaluation principles (particular to D&I research)

- Start early with evaluation activities
- Be transparent in reporting
- Focus on external validity issues to increase relevance
- Have explicit multi-level focus
- Include costs and resource issues
- Include mixed methods

Numerous useful evaluation processes/frameworks

- Pragmatic-Explanatory Continuum Indicator Summary (PRECIS)
- PRECEDE-PROCEED
- Realist Evaluation
- Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM)

Why and to whom do measures matter?

Align your research question and evaluation with the appropriate outcomes

Engenders agreement and respect within the scientific community

Have practical utility to stakeholders

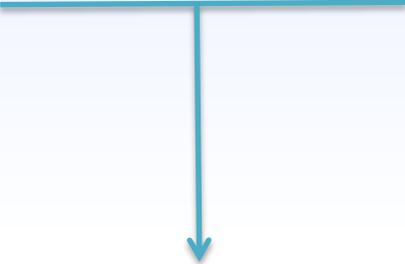
Reassure decision makers of the validity of the approach and motivate use of the EBP

Designing your study

Considering measurement & evaluation issues

What question are you asking?

What type of evidence do you need?



What design & evaluation options do you have?

Design and evaluation options

- **Actual**
 - Retrospective
 - Reliance on secondary data and recall for primary data
 - Observational (longitudinal or cross-sectional)
 - Primary and secondary data
 - Experimental or quasi-experimental
 - Primary and secondary data
- **Modeling/Simulation**
 - Agent-based modeling
 - System dynamics modeling
 - Social network analysis

Considering measurement & evaluation issues

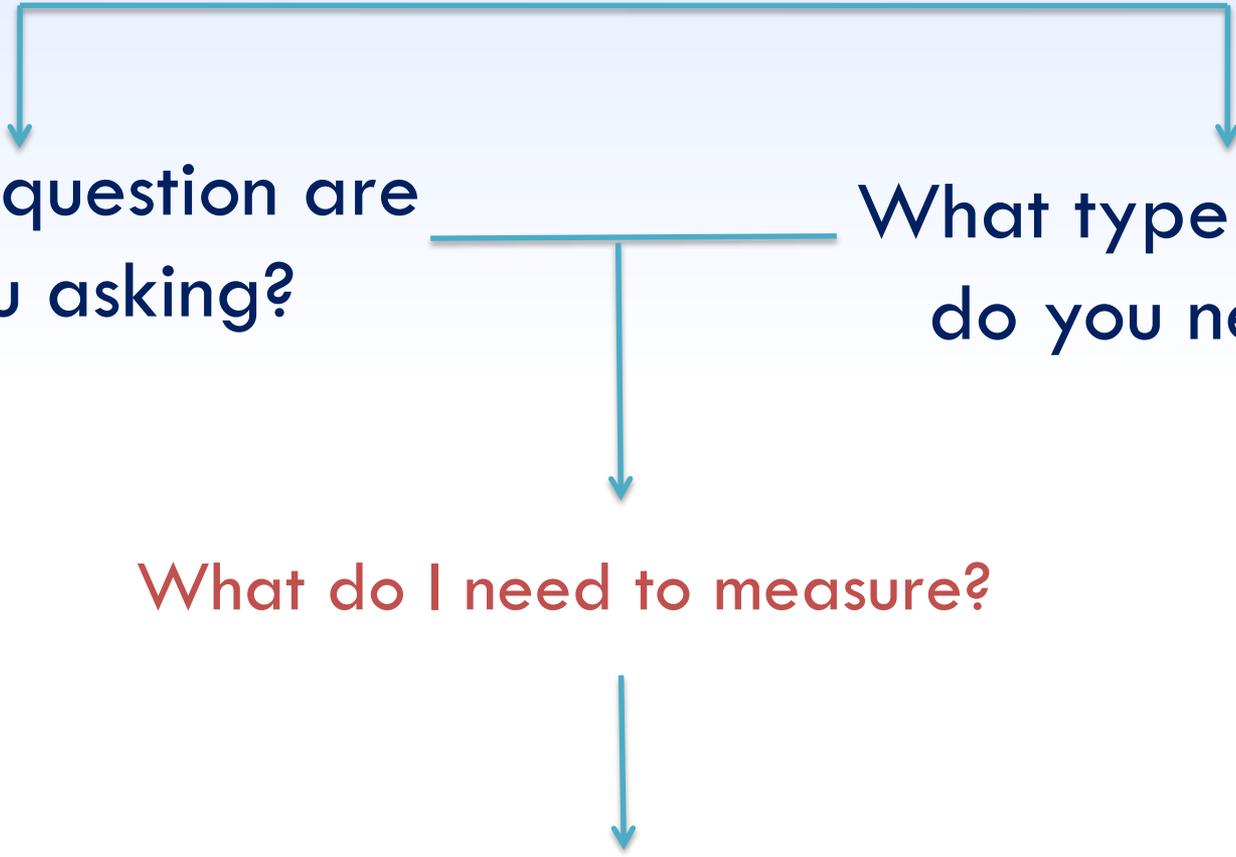
THEORETICAL FRAMEWORK/MODEL

What question are you asking?

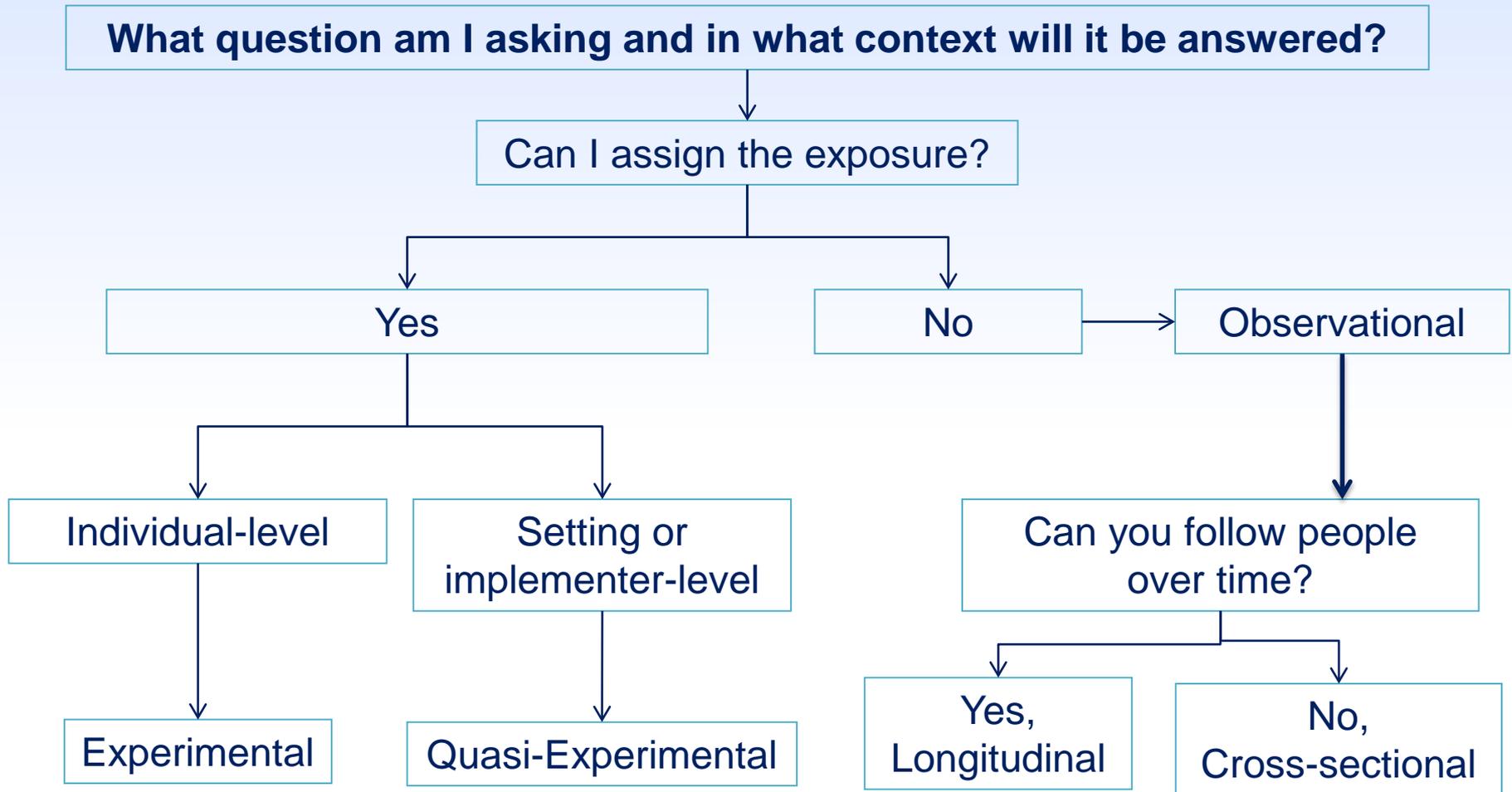
What type of evidence do you need?

What do I need to measure?

What design & evaluation options do you have?



Our flow chart



Some key measurement issues for investigators

- **Maturity of the evidence-based for a practice or policy**
 - May suggest a hybrid design
- **When to use existing measures *versus* developing new measures *versus* adapting measures**
 - Where to turn for existing measures
 - What can be shown in pilot studies
- **How to tie the measures with a conceptual framework**
 - There are over 60 to choose from!!

Consolidated Framework for Implementation Research



INNER SETTING

Combined

Culture

Implementation Climate

IC: Tension for Change

IC: Compatibility

IC: Relative Priority

IC: Organizational Incentives & Rewards

IC: Goals and Feedback

IC: Learning Climate

Networks & Communications

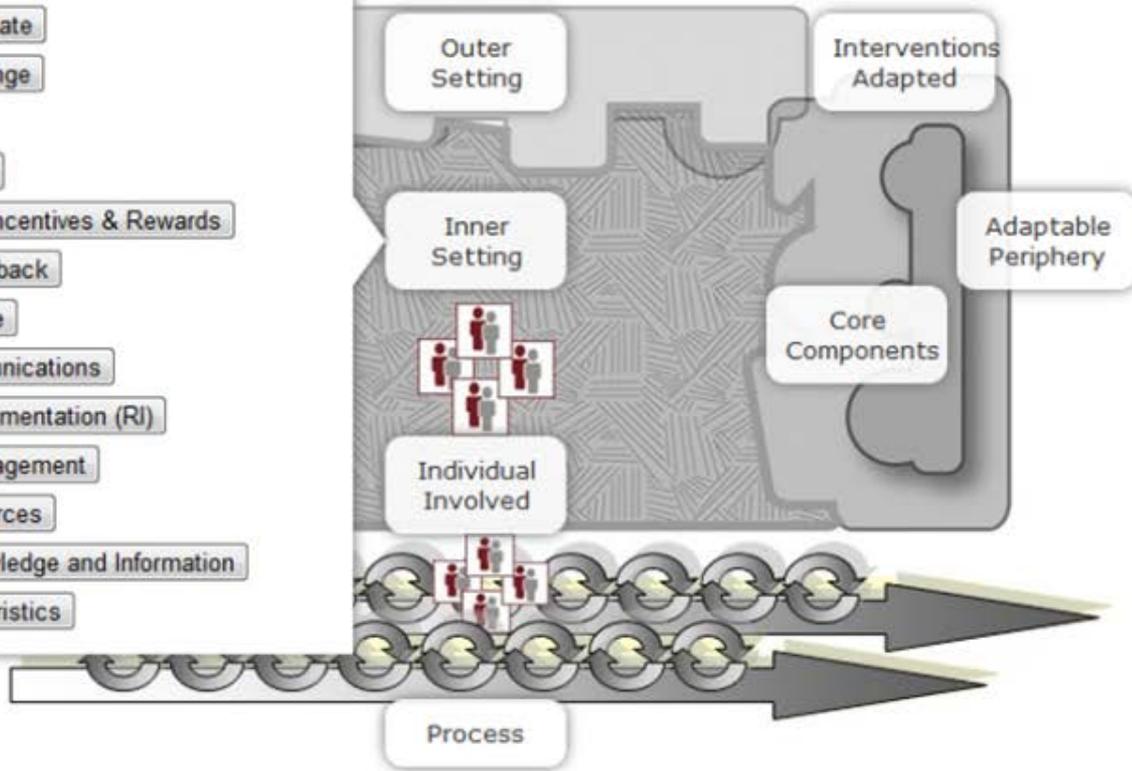
Readiness for Implementation (RI)

RI: Leadership Engagement

RI: Available Resources

RI: Access to Knowledge and Information

Structural Characteristics



Measurement issues D&I research

Measurement in D&I: how is it different?

**Does this program/intervention/treatment work under optimal conditions?
(commonly efficacy/effectiveness measures)**

VERSUS

**When, where, how, with whom, under what circumstances, and why does this program/intervention/treatment work?
(commonly contextual or organizational variables)**

Proposed criteria for rating dissemination and implementation measures for scientific soundness and practicality

GOLD STANDARD MEASURE RATING CRITERIA - For Primary Research Focus

Reliable: Especially test-retest (less emphasis on internal consistency)

Valid: Construct validity, criterion validity, performed well in multiple studies

Broadly Applicable: Available in English and Spanish, validated in different cultures and contexts; norms available; no large literacy issues

Sensitive to Change (if applicable): Longitudinal use, for performance tracking over time

Public Health Relevance: Related to Healthy People 2020 goals, key IOM objectives or national priorities

PRACTICAL MEASURE RATING CRITERIA - For Real-World Application¹

Feasible: Brief (generally 2 to 5 items or less); easy to administer/score/interpret

Important to Practitioners and Stakeholders: Relevant to health issues that are prevalent, costly, challenging; helpful for decision makers or practice

Actionable: Based on information, realistic actions can be taken, e.g., immediate discussion, referral to evidence-based on-line or community resources

User Friendly: Patient interpretability; face valid; meaningful to clinicians, public health officials, and policy makers

Low Cost: Publicly available or very low cost to use, administer, score, and interpret

Enhances Patient Engagement: Having this information is likely to further patient engagement

Do No Harm: Can likely be collected without interfering with relationships, putting respondents at risk, or creating unintended negative consequences

Measurement issues at the macro level (now relates to NIH transparency & reproducibility)

- **Common data elements**
 - Include in terms of award?
- **Shared infrastructure**
 - Platforms needed to carry out the studies (e.g., EHRs, claims records, registries)
- **Data pooling**
 - Expectations to pool data across studies?
- **Harmonization**
 - How to address key D&I issues (e.g., context)?

Measurement resources

Earlier work, focusing on GEM

Rabin *et al.* *Implementation Science* 2012, **7**:119
<http://www.implementationscience.com/content/7/1/119>



METHODOLOGY

Open Access

Advancing the application, quality and harmonization of implementation science measures

Borsika A Rabin^{1*}, Peyton Purcell², Sana Naveed², Richard P Moser², Michelle D Henton¹, Enola K Proctor³, Ross C Brownson^{4,5} and Russell E Glasgow²

Abstract

Background: The field of implementation science (IS) encompasses a broad range of constructs and uses measures from a variety of disciplines. However, there has been little standardization of measures or agreement on definitions of constructs across different studies, fields, authors, or research groups.

Methods: We describe a collaborative, web-based activity using the United States National Cancer Institute's (NCI) Grid-Enabled Measures (GEM) portal that uses a wiki platform to focus discussion and engage the research community to enhance the quality and harmonization of measures for IS health-related research and practice. We present the history, process, and preliminary data from the GEM Dissemination & Implementation (D&I) Campaign on IS measurement.

Results: The GEM D&I Campaign has been ongoing for eight weeks as of this writing, and has used a combination of expert opinion and crowd-sourcing approaches. To date it has listed definitions for 45 constructs and summarized information on 120 measures. Usage of the website peaked at a rate of 124 views from 89 visitors on

Overview

Measures

Datasets

Uploads

Reports

GEM-Dissemination and Implementation Initiative (GEM-D&I)

Related Constructs & Measures

Constructs ▲	Measures	Ratings
Acceptability	Acceptability of Decision Aid Scale	not yet rated Rate This!
Acceptability	Evidence Based Practice Attitude scale	★★★★☆ (average) Rate This!
Acceptability	Suicide Prevention Program Rating Profile	★★★★☆ (average) Rate This!
Acceptability	Therapists' attitudes toward treatment manuals	not yet rated Rate This!
Acceptability	Treatment Evaluation Inventory	★★★★☆ (average) Rate This!
Acceptability	Residents incorporation of EBM in practice	★★★★☆ (average) Rate This!
Acceptability	Atkinson's Perceived Attributes of eHealth Innovations	★★★★☆ (average) Rate This!
Acceptability	Van Schaik's Technology acceptance	★★★★☆ (average) Rate This!
Adaptability	CFIR Adaptability	not yet rated Rate This!
Adherence	Morisky 8-Item Medication Adherence Scale	★★★★☆ (average) Rate This!
Adherence	Morisky 4-Item Self-Report Measure of Medication-Taking Behavior (MMAS-4)	★★★★☆ (average) Rate This!
Adherence	Questionnaire on the Quality of Physician-Patient Interaction	★★★★☆ (average) Rate This!
Adherence	Medical Outcomes Study (MOS) Measures of Adherence	not yet rated Rate This!
Adjustment process	Meaning in Life Scale	★★★★☆ (average) Rate This!
Adopter attitudes	Steckler's Perception of Innovation "Rogers's Adoption Questionnaire"	★★★★☆ (average) Rate This!
Adoption	Roman's measure of adoption in substance abuse treatment (application of the many-facet Rasch model)	not yet rated Rate This!
Adoption	Li, Simon, Bodenheimer, Gillies, Casalino, Schmittiel, & Shortell, adoption	★★★★☆ (average) Rate This!
Adoption	Menachemi, Saunders, Chukmaitov, Matthews, & Brooks	not yet rated Rate This!
Adoption	Noonan, Emshoff, Mooss, Armstrong, Weinberg, & Ball, Adoption, Adaptation, and Fidelity	not yet rated Rate This!
Adoption	Zaidi, Marriot, & Nation, adoption, acceptability and	not yet rated

Society for Implementation Research Collaborative (SIRC): The Instrument Review Project

Leads: Cara C. Lewis, Cameo Borntrager, Ruben Martinez, Phil Fizur, & Kate Comtois

The SIRC Instrument Review Project

Three primary outcomes for this project series include:

- (1) a comprehensive library of D&I instruments measuring the implementation outcomes identified by Proctor and colleagues (2010) and organized by **the Consolidated Framework for Implementation Research** (CFIR; Damschroder et al., 2009) for SIRC members;
- (2) a rating system reflecting the degree of empirical validation of instruments, adapted from the **evidence-based assessment (EBA) work** of Hunsley and Mash (2008) and Terwee et al (2012);
- (3) a **consensus battery of instruments** decided upon by D&I expert task force members using the EBA criteria to guide future D&I research efforts.

An update on SIRC Instrument Review Project Progress

We have now identified nearly 450 instruments to be included in our repository. The instruments assessing Implementation Outcomes (Proctor et al., 2009) have all been twice rated using our evidence-based assessment rating criteria. Results are available through the IRP portion of the website. The repository is available to all SIRC members

www.societyforimplementationresearchcollaboration.org

Evidence-Based Assessment Criteria

EBA

0 = None

1 = Minimal

2 = Adequate

3 = Good

4 = Excellent

Psychometric Properties

- ◆ Norms
- ◆ Reliability Information: Inter-rater, Test-retest, Internal
- ◆ Validity Information: Criterion, Content, Construct
- ◆ Validity Information: Predictive, Discriminant, Convergent, Concurrent
- ◆ Usability

SIRC's Instrument Repository

- Organized by
 - Consolidated Framework for Implementation Research (39 constructs)
 - Damschroder, et al. 2009
 - Outcomes for Implementation Research (17 constructs)
 - Proctor, et al. 2011
 - Construct reviews resulted in over 400 measures

SHORT REPORT

Open Access



Measurement resources for dissemination and implementation research in health

Borsika A. Rabin^{1,2*}, Cara C. Lewis^{3†}, Wynne E. Norton⁴, Gila Neta⁴, David Chambers⁴, Jonathan N. Tobin⁵, Ross C. Brownson^{6,7} and Russell E. Glasgow²

Overview

- NIH consensus meeting of 23 experts in 2013
- Snowball sampling to identify 17 measurement resources
 - 12 static, 5 web based (“living”)
 - 13 from health
 - 11 used database reviews
 - 14 organized around some theory/framework
 - 10 were QUANT focused
 - 6 reported psychometric properties

GEM D&I https://www.gem-measures.org/public/wsoverview.aspx?wid=11&cat=8&aid=0	359	895	Yes	Author/date, instrument citation, language, purpose/definition, total number of items	Reliability: internal consistency Validity	Yes	Consumer, organization, Provider, System, Team
---	-----	-----	-----	---	---	-----	--

Table 2 Summary data of measures information

Resource	# of constructs	# of measures	Measures included	Measure metadata	Psychometric information	Pragmatic rating	Analysis level
Ready, set change! Online decision support tool http://www.ncbi.nlm.nih.gov/pubmed/24886072	4	9	Yes	Author/date, instrument citation, instrument link, instrument setting, total number of items	NS	Yes	Organization
Proxy measures of clinical behavior http://www.ncbi.nlm.nih.gov/pubmed/19575790	1	15	No	Author/date, instrument citation	Reliability: inter-rater; test-retest Validity: external	No	Consumer, organization, Provider, System, Team
Organizational readiness for change measures http://mcr.sagepub.com/content/65/4/379.abstract	1	43	No	Author/date, instrument citation, instrument setting	Reliability: inter-item; inter-rater; parallel forms; test-retest Validity: concurrent; convergent; discriminate; face/content; predictive	No	Organization

Key challenges and opportunities for assessment of D&I research

Challenge

Challenge 4: Qualitative information matters (at least) as much as quantitative data

No standardized guidance on assessment of qualitative instruments for quality

Use of common instruments is even more challenging

Opportunities

Need to identify unobtrusive, observational approaches

A few examples/guidance for use and integration of qualitative data:

OBSSR: Best Practices for Mixed Methods Research in the Health Sciences

My Own Health Record project and AHRQ context instruments (Kurt Stange, Russ Glasgow)¹

CFIR qualitative wiki²

NCI Implementation Science group-led

QUALRIS: Qualitative Research in Implementation Science

Challenge 5: Need for practical/feasible and actionable measures

Need to balance traditional gold standard criteria with the criteria for practicality/feasibility for use in real world settings especially when measuring multiple behaviors, multiple people, multiple conditions

Need to ensure that what is measured does matter to end-users (actionable, in-line with their priorities, measured at the right level)

Examples and guidance:

GEM D&I Initiative criteria for feasibility ³

Glasgow & Riley paper on practical measures ⁴

EMR brief measure campaign using GEM ⁵

Jonathan Tobin - CTSA D&I measures effort

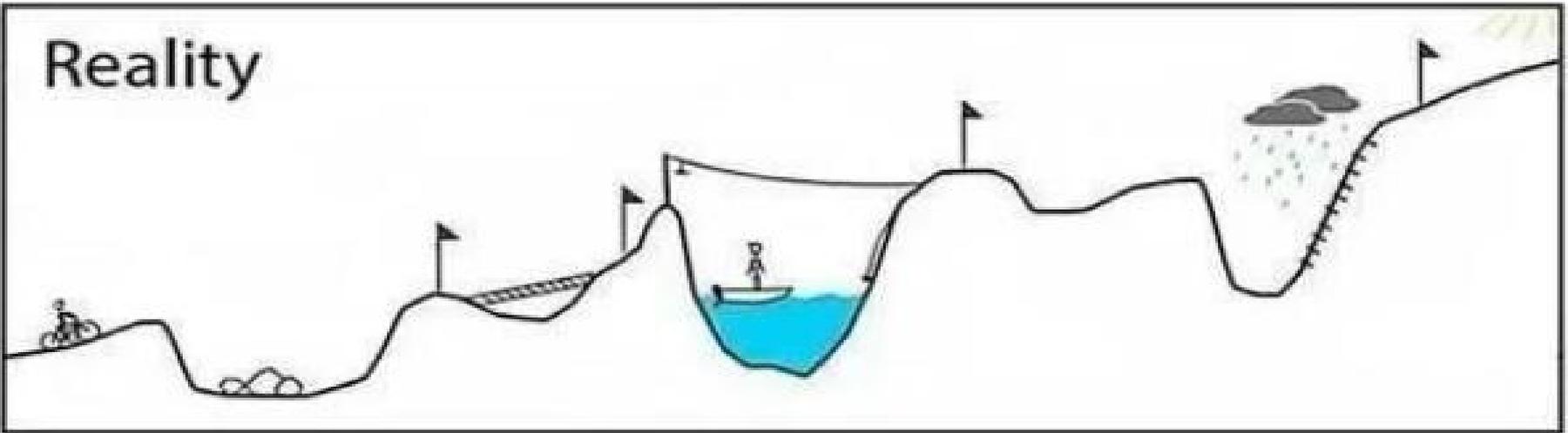
¹ Krist AH, et al. Designing a valid randomized pragmatic primary care implementation trial: the my own health report (MOHR) project *Implementation Science* 2013, 8:73; ² CFIR wiki; ³ Rabin et al. Advancing the application, quality and harmonization of implementation science measures *Implementation Science* 2012, 7:119; ⁴ Glasgow RE, Riley WT. Pragmatic measures: what they are and why we need them. *Am J Prev Med.* 2013; 45(2):237-43.; Glasgow RE, Brownson RC, Kessler RS. Thinking about health-related outcomes: what do we need evidence about? *Clin Transl Sci.* 2013;6(4):286-91. ⁵ Estabrooks PA, et al. Harmonized patient-reported data elements in the EHR: Supporting meaningful use by primary care action on health behaviors and key psychosocial factors. *J Am Med Inform Assoc* 2012 Jul 1;19(4):575-82.

Expect the Unexpected (Plan B Evaluation)

Your plan



Reality



Take home points

1. Since D&I research asks different questions than effectiveness research, the measurement and evaluation approaches also differ.
2. The research questions, design and matched theoretical framework should guide measure selection.
3. There is a growing set of resources and measurement tools to support D&I research.

THANKS to David Chambers,
Bridget Gaglio, Christine
Hunter, Gila Neta!!