McMaster University

Family Medicine

David Braley Primary Care Research Collaborative

DESIGNING AND ADMINISTERING SURVEYS

Dr. Michelle Howard

Presenter's Twitter Handle
@McMasterFamMed

Michael G. DeGroote SCHOOL OF MEDICINE

Review: Main Components

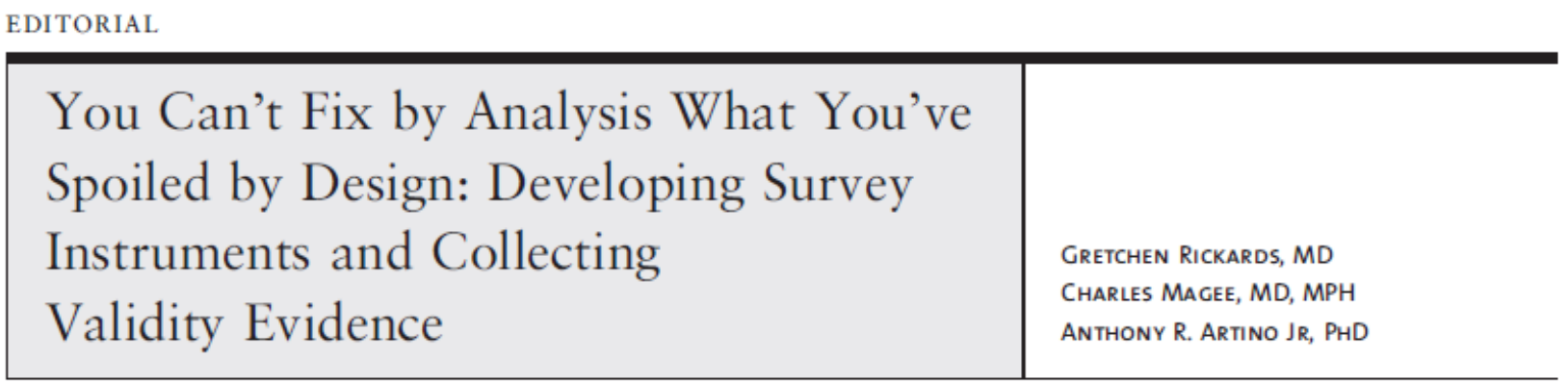
- Engage stakeholders
- Describe the educational innovation/program
- Choose evaluation framework
- **Gather credible evidence**
- Justify conclusions
- Ensure use of evaluation findings

Learning objectives

At the completion of this module, you will be able to:

1. Articulate the steps in developing a survey instrument, or knowing what to look for when using a survey
2. Compare the pros and cons of different methods of survey administration
3. List the considerations in determining how many people to survey

Outline



EDITORIAL

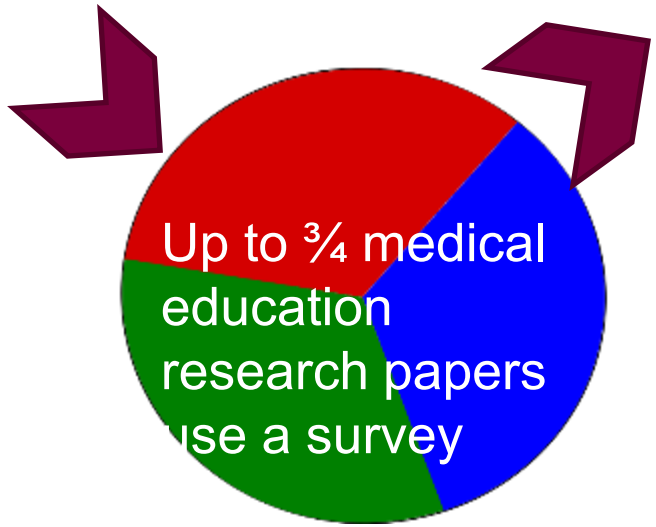
You Can't Fix by Analysis What You've Spoiled by Design: Developing Survey Instruments and Collecting Validity Evidence

GRETCHEN RICKARDS, MD
CHARLES MAGEE, MD, MPH
ANTHONY R. ARTINO JR, PHD

Survey Use in Evaluation of HSE

UBIQUITOUS e.g.

- Evaluations of courses
- Satisfaction with program
- Perceived learning



CAUTIONS

- Poorly designed survey leads to measurement error i.e.
  - Does not capture what is intended

Survey Terminology

Construct = A hypothesized concept, model, idea or theory that is believed to exist but cannot be directly observed

The following items are about activities you might do during a typical day. Does **your health now limit you** in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
3. **Vigorous activities**, such as running, lifting heavy objects, participating in strenuous sports	1	2	3
4. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
5. Lifting or carrying groceries	1	2	3
6. Climbing **several** flights of stairs	1	2	3
7. Climbing **one** flight of stairs	1	2	3
8. Bending, kneeling, or stooping	1	2	3
9. Walking **more than a mile**	1	2	3
10. Walking **several blocks**	1	2	3
11. Walking **one block**	1	2	3
12. Bathing or dressing yourself	1	2	3

Survey Design and Use

- Is a survey an appropriate tool?
- How have others addressed this construct in the past?
- How to develop survey items?
- Are survey items clearly written and relevant to the construct of interest?
- Will respondents interpret items in the manner intended?
- Are the scores obtained reliable and do they relate to other measures as hypothesized?

Item Generation

1. **Literature Review**
   a) Identify existing survey measures
   b) If a new survey measure needed, to define the construct from previous research- how is it the same or different from other constructs? (e.g. depression versus anxiety)
2. **Interviews or focus groups**
   a) People who represent population of interest
   b) Experts on the topic

Item Generation Cont'd

Write out the items (i.e. questions)

- Accurately reflect the construct
- Use language of respondents
- Start with more items and reduce
- 'rule of thumb'- 25 questions to address a research question, approx. 5 per domain

Item Reduction

Why?

- To make as short as possible without losing necessary information (keeping desired constructs)
  - Eliminate questions that duplicate the same idea
  - Eliminate or re-word questions that are difficult to answer
  - Eliminate questions that everyone answers the same way

How?

- Expert input
- Pilot testing
- Statistical analysis

Testing the Survey

Pre-testing sensibility testing

Focus on Individual Items

- Expert review
- Language clarity, readability, cultural issues
- Interpretation of individual questions

Pilot testing

Focus on entire survey

- Actual respondents
- Penultimate draft of entire package (pre-amble, introductory letter etc)
- Ensure consistent interpretation, accurate responding
- Flow, ease, burden
- May use for statistical testing of reliability and validity (n=100+)

Clinical

Focus on big picture

- Asks about whether the survey reflects the issue of interest - comprehensive, on topic

Validity

Validity = the degree to which evidence and theory support a measure's intended use

- Not a 'property' of the survey - depends on context and population
- Accumulation of evidence across time, settings, samples to build an argument
- Face, content, construct, criterion

Reliability Types

Reliability = extent to which the scores produced by a the survey are consistent and reproducible

Test Re-test

- To assess whether the same question posted to the same individuals gives consistent results at different times

Internal consistency

- To assess whether different items in the same ‘construct’ are correlated with each other (versus items from other constructs)

Re-Cap

- There are many steps to good survey design
- Use existing survey where possible – look for elements of good design
- Pilot test in your population
- Allow time and resources if developing a new survey

Methods of Administration

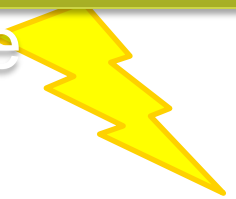
- **Self-administered – Electronic/Online**
  - OK for captive audience (e.g. recent graduates, course evaluation)
  - Inexpensive
  - Generally low response rate
  - Difficult to track respondents, more 'incompletes'
  - Easy to ignore (survey overload)
  - More likely to complete if it comes from a respected source, relevant to respondent

- **Self-administered – Paper based**
  - Likely remains best option
  - Works for people who don't use computer
  - More personal – send a pre-notice, reminders
  - Easier to track responses

- **Interviewer administered**
  - Time intensive, expensive
  - Difficult to schedule
  - Interviewer effects
  - Use sparingly, as needed e.g. small target population, detailed survey

Achieving good response rate

A good response rate is 70% for generalizability
Lower than 50% is questionable



Improving response rate:

- Pre-notice
- Shorter survey
- Reminders (up to 2 after initial)

What is the right sample size?

Some ‘rules of thumb’

1. Surveying ALL or SOME program participants?
   - Survey ALL if possible (e.g. if 100 or fewer)
   - Survey a sample of SOME if all not possible

2. If using a sample – confidence in estimate
   - A common standard is 95% confidence, with a sampling error of +/- 5%

3. Allow for non-response – inflate sample size
   - Expect at least 20% non-response

Diminishing returns (5% margin of error)

Need to sample:
80 from population of 100
278 from population of 1000
370 from population of 10,000

- Methods of survey administration affect response rate
- In many program evaluations it is reasonable to survey everyone rather than a sample
- If a sample must be drawn, do a sample size calculation so the confidence in the estimate is known
- Inflate the sample size to account for non-response

McMaster University

Family Medicine

David Braley Primary Care Research Collaborative

Department of Family Medicine
Michael G. DeGroote School of Medicine
Faculty of Health Sciences

fammedmcmaster.ca
@McMasterFamMed