



# You Can't Fix by Analysis What You've Spoiled by Design:

## Survey Design Tenets Applied to Evaluations

**Jeffrey LaRochelle, MD, MPH, FACP**  
Associate Professor of Medicine  
Director, Integrated Clinical Skills



# Disclosures

- None

The views expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, Department of Defense, nor the U.S. Government.

# Learning Objectives

- *Recognize* a systematic framework for survey design
- *Demonstrate* how to apply survey design tenets to the development of evaluations
- *Identify* common item-writing pitfalls
- *Define* the purpose of expert validation, cognitive interviews, and pilot testing

# Consider this...

- The puppy problem

- The poodle has 9 puppies.  
 - The collie has 5 puppies.  
 - How many more puppies does the poodle have?

- Students' common response...  
 "None"
- Why?  
 "It said she had 9 puppies, but it didn't say she had any more, so it's none."

## Revised item...

- The poodle has 9 puppies.  
 - The collie has 5 puppies.  
 - How many more puppies does the poodle have *than the collie*?

# And this...

Your opinion is that the global economy is the *second* most important issue in the world today.

The global economy is the most important issue in the world today.

strongly disagree	disagree	neither agree nor disagree	agree	strongly agree
-------------------	----------	----------------------------	-------	----------------

How important is the issue of the global economy in the world today?

not at all important	slightly important	moderately important	quite important	extremely important
----------------------	--------------------	----------------------	-----------------	---------------------

# Evaluations as Surveys

- Survey
  - Abstract concepts
    - Attitudes
    - Opinions
    - Beliefs
  - Create action items
    - Improve attitudes
    - Understand deficits
  - Conversation
    - With respondents
- Evaluations
  - Abstract concepts
    - Communication skills
    - Clinical reasoning
    - Professionalism
  - Create action items
    - Improve skills/abilities
    - Provide feedback
  - Conversation
    - With faculty

Principle #1: You can't fix by analysis what you've spoiled by design.

- Critical to get it right at the design phase
  - Measure multi-dimensional concepts
  
- Poorly written items lead to “bad data”
  - Inconsistent results – fairness issues
  
- Difficult to develop actionable items
  - Appropriate feedback

# Common Language

- Construct or Concept
  - Communication Skills
- Dimension or Facet
  - Major components of your *construct*, e.g.
  - Communication Skills
    - Opening the discussion, Gathering information, etc...
- Items (or “indicators”)
  - Individual questions/statements
- Scale
  - > 3 items intended to measure a *construct/dimension*



# Common Language

- Response anchors (aka, “response options”):
  - All the named points along the response scale

for example

Clear fail	Borderline	Clear pass	Exceeds expectations	Exceptional
never true	rarely true	sometimes true	often true	true nearly all of the time
Poor FOK	Limited FOK	Solid FOK	Outstanding FOK	Not observed

- Satisficing:
  - Respondents compromise standards
    - Don't put forth effort to answer thoughtfully

# ITEM(S)

# CONSTRUCT

## Common In-Patient Milestones

1. Manage a hospitalized patient with common infections
2. Discharge a patient safely by coordinating care with the primary care manager.
3. Assess the risk for venous thromboembolism for a hospitalized patient and develop an appropriate prevention plan.
4. Diagnose the cause of altered mental status in a hospitalized patient.
5. Effectively ensures that their patients comprehend the plan of care by avoiding the use of medical jargon.

# RESPONSE ANCHORS

1. Resident *cannot* perform this skill *even with assistance*
2. Resident should perform this skill under *direct supervision* of a senior resident or fellow
3. Resident can perform this skill under *indirect supervision* of the attending
4. Resident can perform this skill *independently*
5. Resident can act as an *instructor or supervisor* for this skill (*aspirational*)

SCALE



# 7-Step Design Process

- Step 1: Literature Review
- Step 2: Interviews & Focus Groups
- Step 3: Synthesize
- Step 4: Develop Items
- Step 5: Expert Validation
- Step 6: Cognitive Interviewing
- Step 7: Pilot Test

# Step 1: Literature Review

- Critically evaluate the literature
  - How is the construct defined in prior studies?
  
- Identify existing scales
  - What items/scales currently exist?
  - Appraise quality

## Step 2: Interviews & Focus Groups

- Goal
  - Identify initial dimensions of the construct
  
- Interview experts
  - Local faculty
  
- Apply open-ended questions
  - Avoid yes/no, multiple-choice questions

## Step 3: Synthesize Literature & Interviews

Goal: Arrive at consensus/agreement



# Step 4: Develop Items

Goal: Develop items using vocabulary your target population can understand

- Considerations

- Vocabulary and wording
- Response anchor selection
  - Ratings vs. rankings; Likert-scale items; yes/no items?
- Item formatting
  - Visual design, item order, instructions, etc.

# Step 4: Develop Items (examples)

## Communication Skills – Standardized Patient Encounter (full scale = 7 items)

Rate the student on the following communication skills:

1. The student introduced themselves properly
2. The student treated you with respect
3. The student used appropriate, open-ended questions
4. The student listened intently, and let you tell your story
5. The student showed interest in your symptoms and concerns

response anchors

Poor	Fair	Good	Very Good	Excellent
1	2	3	4	5



# Step 4: Develop Items (examples)

## Course Importance (a belief; the full scale = 6 items)

1. How important was it for you personally to perform well in this course?
2. How important were the practical applications of the information provided in this course?
3. How important was the content of this course?
4. How important was it for you to learn the material in this course?

### response anchors

not at all important	slightly important	moderately important	quite important	extremely important
----------------------	--------------------	----------------------	-----------------	---------------------

## Principle #2:

The questions guide the answers.

9) What topic(s) of study are you most interested in pursuing while at USU? (Total  $N = 11$ )

Ver 1: Lots of Space  
(5 lines)

( $n = 5$ )

- Financing of health care
- Global health, joint operations
- Policy development with regard to military and operational
- Health policy, health economics
- Health care admin and policy

Total Word Count = 25  
Mean Word Count = 5.0

Ver 2: Small Amount of Space  
(1 line)

( $n = 6$ )

- Public health
- International health
- (blank)
- (blank)
- Health insurance
- Policy

Total Word Count = 7  
Mean Word Count = 1.2

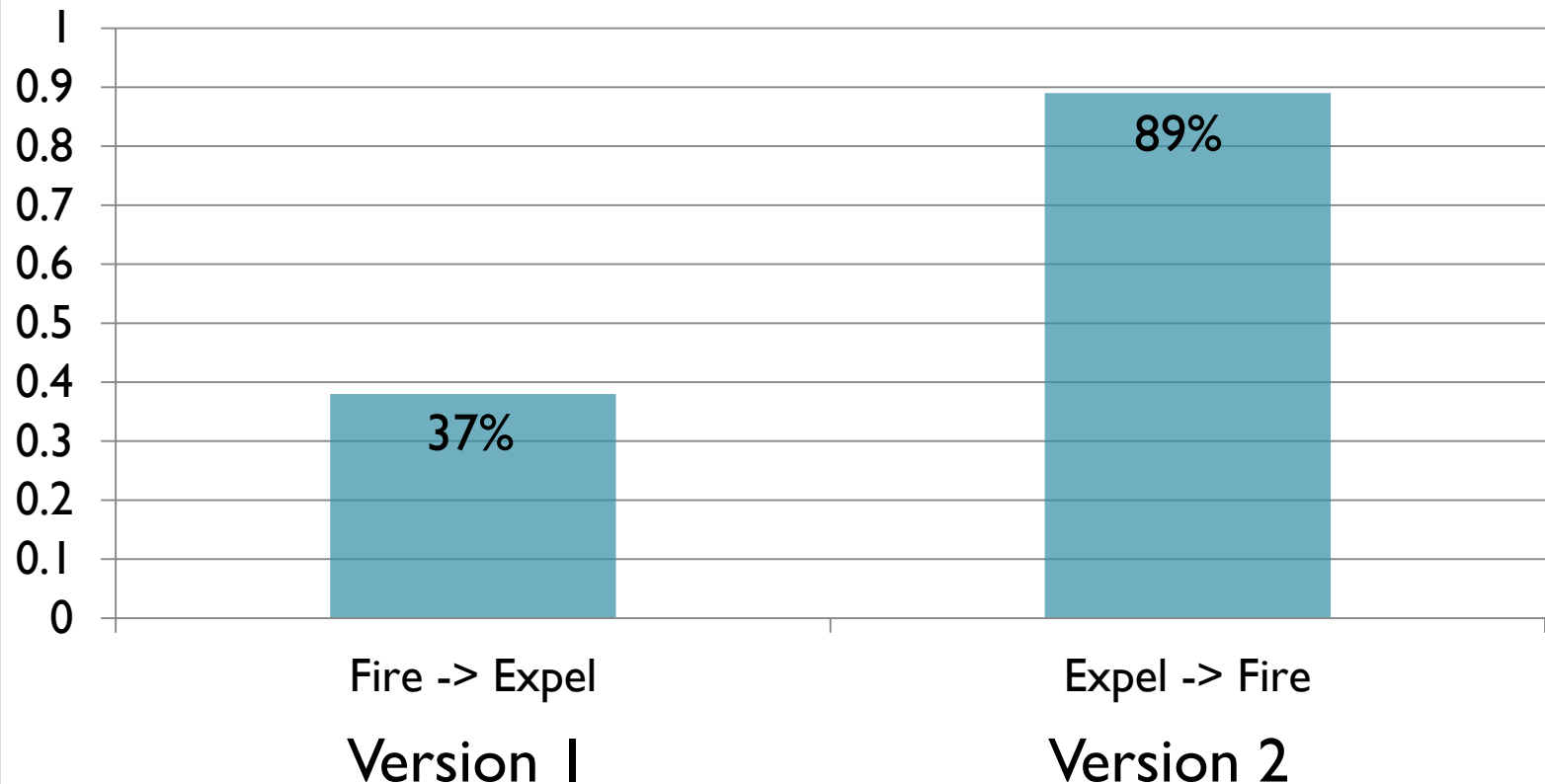
**Cohen's  $d = 2.62$**   
 $t(9) = 4.63, p < .001$

## Principle #2:

The questions guide the answers.

N = 91 Faculty

% Answer = Fire Professor



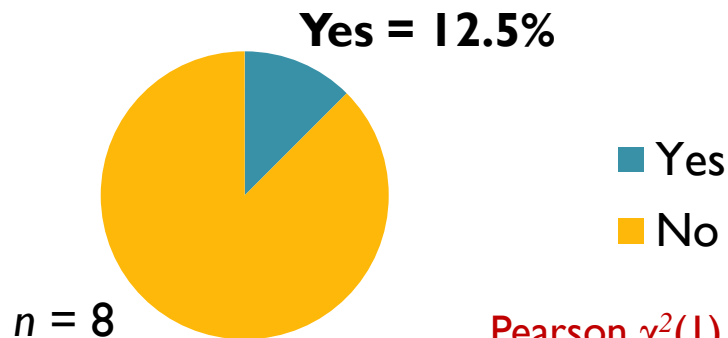
Pearson  $\chi^2(1) = 4.90, p < .05$

## Principle #3: A survey is a conversation between you and your respondents.

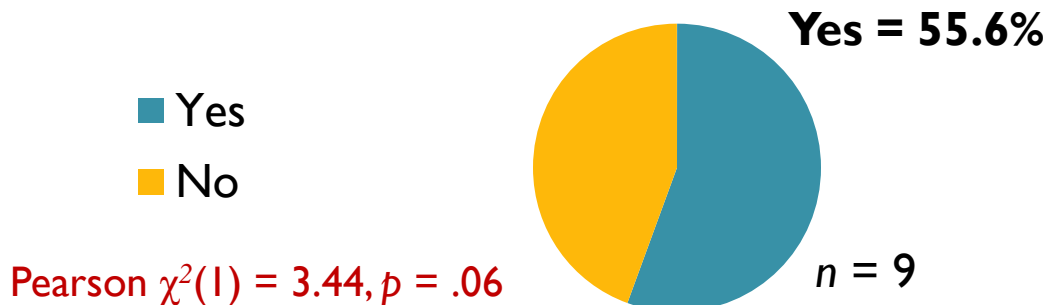
A note about providing a reason

**“because...” or “so that...”**

8) As some of you may know, the university is debating whether to move some parts of the university to a new section of campus in Rockville. Do you think the university should move to Rockville?



8) As some of you may know, the university is debating whether to move some parts of the university to a new section of campus in Rockville. Do you think the university should move to Rockville so that the school can have more space?



# Common Pitfalls

- **Creating double-barreled items**
  - Example Item: “Assess the risk for venous thromboembolism for a hospitalized patient and develop an appropriate prevention plan?”
    - What if one is good and the other is bad?
  - Solution: split into two items
    - “Ability to assess risk...?”
    - “Ability to develop a prevention plan...?”
    - Create a double-barreled response anchor??

# Common Pitfalls

- **Creating double-barreled items**

- Construct = *Elaboration*

Inter-Item Correlation Matrix

	Q2_A_30	Q2_A_36	Q2_A_37	Q2_A_40	Q2_A_41	Q2_A_50
Q2_A_30	1.000	.084	.439	-.271	.616	.297
Q2_A_36	.084	1.000	.741	-.255	.087	.727
Q2_A_37	.439	.741	1.000	-.217	.356	.683
Q2_A_40	-.271	-.255	-.217	1.000	.200	-.168
Q2_A_41	.616	.087	.356	.200	1.000	.445
Q2_A_50	.297	.727	.683	-.168	.445	1.000

- Item 40. When I study for this course, I write brief summaries of the main ideas from the readings and online discussions

Cronbach's alpha = 0.546

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q2_A_30	25.68	15.180	.285	.623	.512
Q2_A_36	25.86	13.076	.379	.745	.461
Q2_A_37	25.36	12.814	.609	.704	.391
Q2_A_40	28.64	17.004	-.178	.373	.797
Q2_A_41	26.14	12.028	.558	.644	.380
Q2_A_50	26.50	10.643	.598	.686	.329

# Common Pitfalls

- **Creating negatively worded items**

- Unnecessary cognitive burden
- Promotes satisficing
  - “In an average week, how often are you unable to start rounds on time?” (rarely-often)
- Solution: make sure “yes” means “yes” and “no” means “no”
  - “In an average week, how often do you start rounds on time?”

# Common Pitfalls

- **Using statements instead of questions**
  - Example Item: “The learner is respectful to patients based on their gender.”

Not at all  
true

A little bit  
true

Somewhat  
true

Mostly  
true

Completely  
true

- People are better at answering questions
- Use questions with construct-specific anchors
  - “How often is the learner respectful to patients based on their gender?”
    - Rarely to Always response anchors (frequency)

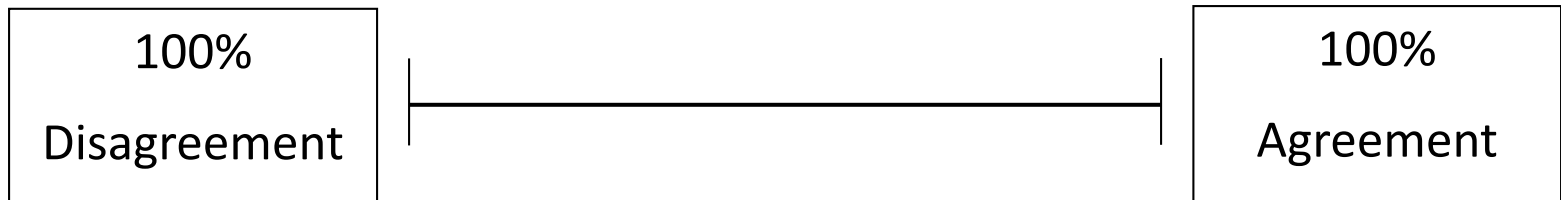


# Common Pitfalls

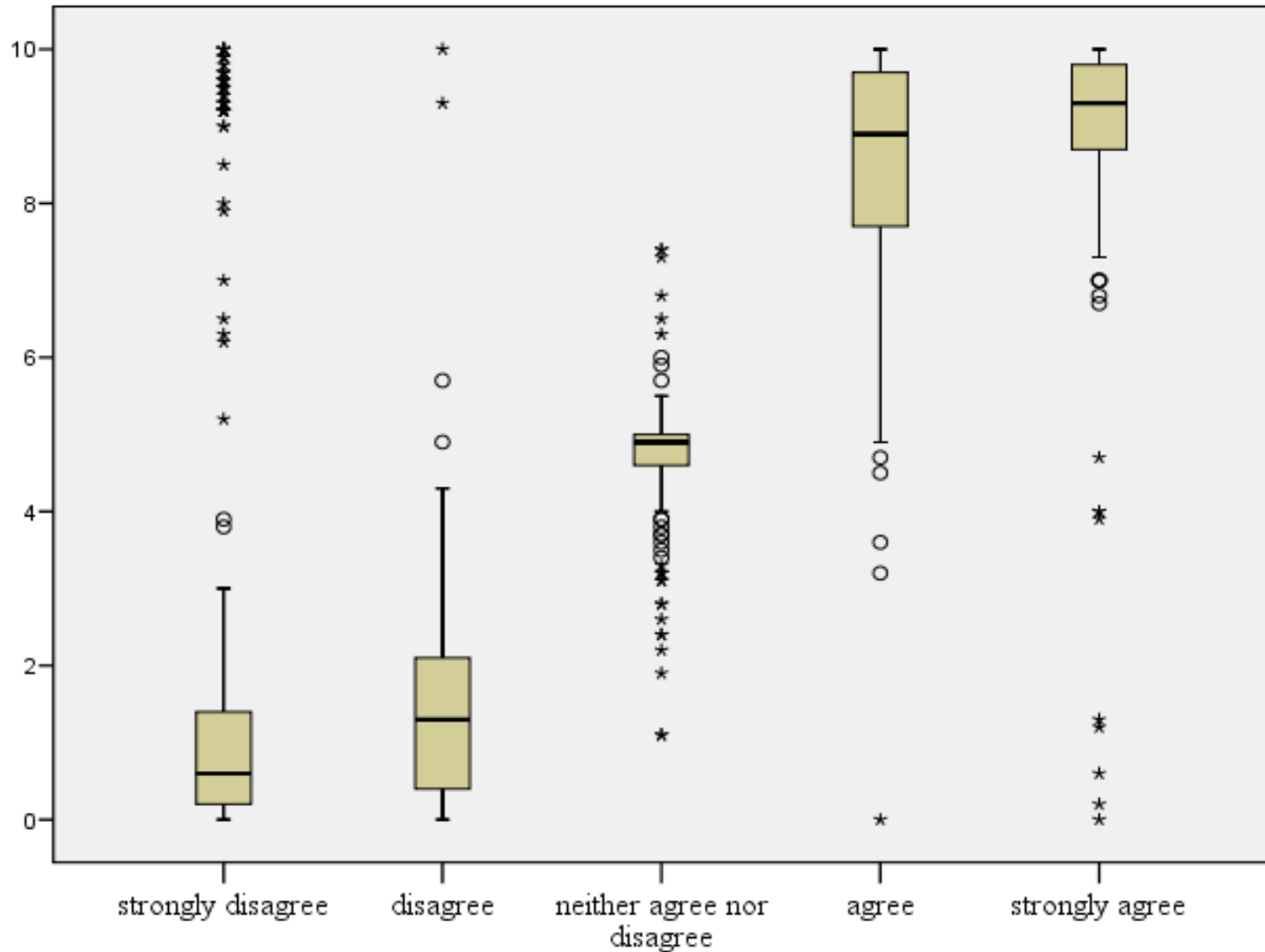
- **What does it mean to “strongly agree” anyway?**

**Section II:** In this section, each question will ask you to indicate how you understand a commonly used phrase by marking an “X” at the appropriate place on the line.

25) When you say that you “strongly agree” with somebody else, what do you mean? Indicate on the line below where “strongly agree” is by marking an “X” on the line.



# Strongly Disagreeable Ranges



# Common Pitfalls

- **Using too few or too many response anchors**
  - Influences reliability within a set of survey items
    - Too few (<4) → less reliable
    - Too many (>7-9) → diminishing return; false impression of precision
- Example Item: “How useful was the rotation in emergency medicine?”

Not at all useful	Moderately useful	Very useful
-------------------	-------------------	-------------

?

Not at all useful	Slightly useful	Moderately useful	Quite useful	Extremely useful
-------------------	-----------------	-------------------	--------------	------------------

Not at all useful	Minimally useful	Slightly useful	Somewhat useful	Moderately useful	Quite useful	Very useful	Extremely useful
-------------------	------------------	-----------------	-----------------	-------------------	--------------	-------------	------------------

# Survey Design: 7-Step Process

- Step 1: Literature Review
- Step 2: Interviews & Focus Groups
- Step 3: Synthesize
- Step 4: Develop Items
- **Step 5: Expert Validation**
- **Step 6: Cognitive Interviewing**
- **Step 7: Pilot Test**

# Step 5: Expert Validation (aka, content validation)

**Goal:** Make sure the items “ring true” to experts

- Depending on your needs, experts can consider the following for each of your survey items...
  - Clarity
  - Construct relevance
  - Language level
  - Missing facets/aspects
  - Difficulty

# Step 6: Cognitive Interviewing

**Goal:** Make sure respondents understand the items as intended by you (the developer)

- Recruit members of the targeted population
  - e.g., students, teachers, patients, locals, etc.
- Conduct one-on-one interviews, in “laboratory” or other location
- THEN: Make informed decisions, with cognitive interview as one source of input

# Step 6: Cognitive Interviewing

## Example

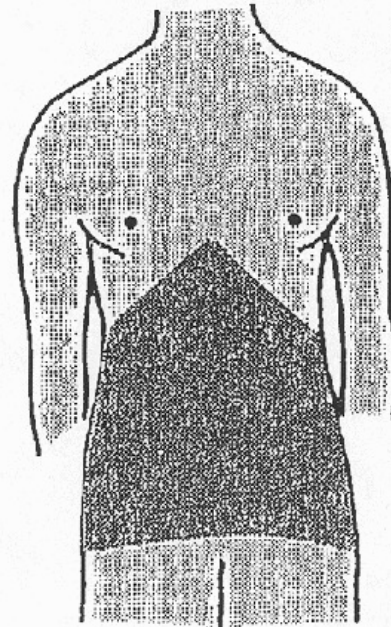
**“In the last year, have you been bothered by pain in the abdomen?”**

- What problems do you anticipate?
  - What time period are you thinking about?
  - What does “bothered by pain” mean to you?
  - Where is your “abdomen?”

# Step 6: Cognitive Interviewing

## Example

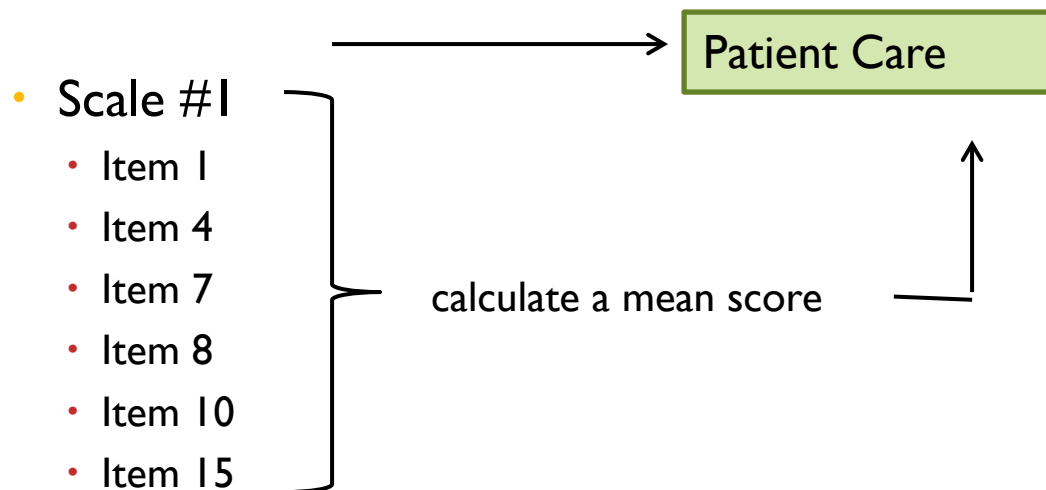
**“Please look at this diagram. During the past 12 months, have you had pain in this area (the area shaded on the diagram)?”**





# Step 7: Pilot Testing

- “Get to know” your descriptive statistics
  - Do individual survey items “hang together”?
    - Factor analysis and reliability analysis



- Relation to other dimensions as you would expect?
  - (+) correlations with Clinical Reasoning
  - (-) correlations with Clinical Exam Skills

# Questions?



If you remember nothing else, remember...

Principle #1: You can't fix by analysis what you've spoiled by design.

Principle #2: The questions guide the answers.

Principle #3: A survey is a conversation between you and your respondents.

Jeffrey La Rochelle, MD, MPH

[jeffrey.la-rochelle@usuhs.edu](mailto:jeffrey.la-rochelle@usuhs.edu)