



## How to Develop and Utilize Surveys in Research

CDR Anthony R. Artino, Jr., PhD  
Medical Service Corps, U.S. Navy  
Associate Professor, PMB

LTC Kent J. DeZee, MD, MPH  
Medical Corps, U.S. Army  
Associate Professor, MED



February 16, 2012

## Learning Objectives

- Recognize how to use a systematic, 7-step process as a framework for survey design
- Demonstrate how to develop an appropriate set of items to characterize the construct being measured
- Identify common item-writing pitfalls in survey design



## Consider this...

- The puppy problem

- The poodle has 9 puppies. - The collie has 5 puppies. - How many more puppies does the poodle have?	<ul style="list-style-type: none"> <li>• Students' common response... "None"</li> <li>• Why? "It said she had 9 puppies, but it didn't say she had any more, so it's none."</li> </ul>	Revised item... - The poodle has 9 puppies. - The collie has 5 puppies. - How many more puppies does the poodle have <i>than the collie?</i>
---	--	---



## And this...

Your opinion is that the high cost of health care is the *second* most important issue in America today.

The high cost of health care is the most important issue in America today.

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

vs.

How important is the issue of high healthcare costs in America today?

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



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

- When creating a survey, it's important to get it right the first time
- We often use surveys to measure "fuzzy" constructs (e.g., beliefs, attitudes, opinions), which is already hard enough
  - *Bad data only makes it that much harder*



## Outline

- Survey Background
  - Uses of surveys
  - Survey language
- Survey Design
  - 7-Step Process
  - Small-group activity
- Common Pitfalls
- Debrief



### Good for...

- Abstract ideas/ concepts
  - Opinions
  - Beliefs
  - Attitudes
- Behaviors
  - Assuming that observing behaviors is impractical
  - And that people might reasonably report on their own behaviors

- In 2009, **56% of 141** research articles published in *Academic Medicine* used survey research methods (Gehlbach et al., 2010)
- From 2008 to 2011, **30% of 116** research papers presented at annual meetings of the Association for Medical Education in Europe (AMEE) used survey research methods (Artino et al., in press)



### Not so good for...

- Clinical outcomes
  - Better to measure perceptions (ex: health questionnaires)
- Populations with cognitive impairment, severe disease
  - ICU patients may have difficulty with surveys
- Tasks of high cognitive load/burden
  - "How many hours did you use the Internet last year?"
  - "What did you eat for dinner on Wed, 11 Jan 2012?"



### Survey Language

- Construct:
  - A model, idea, or theory
    - e.g., resilience, confidence, patient satisfaction, motivation, perceived barriers, interest, procrastination, health-related stigma
- Items (or "indicators"):
  - Individual questions/statements on the survey
- Scale:
  - 3 or more items intended to measure a construct



### Survey Language

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

for example

not at all important	slightly important	moderately important	quite important	extremely important
never true	rarely true	sometimes true	often true	true nearly all of the time
almost never	once in a while	sometimes	often	almost all the time

- Satisficing:
  - Occurs when respondents compromise standards to expend less energy
    - i.e., they don't put forth effort to answer truthfully or thoughtfully



ITEM(S)
CONSTRUCT
RESPONSE ANCHORS

Customer Service Questionnaire  
Please help us improve our services by answering the questions below. We are interested in your opinions about the recent service you received from the Reading Fire Department

Please indicate your response

	Very Satisfied	Satisfied	Adequate	Unsatisfied	Very Unsatisfied
1. Were our personnel polite and courteous?	5	4	3	2	1
2. Did our personnel take care of you in a professional manner?	5	4	3	2	1
3. Did we explain the services you needed in an understandable manner?	5	4	3	2	1
4. Did we answer all of your questions in an understandable way?	5	4	3	2	1
5. Overall, how satisfied were you with the service you received from us?	5	4	3	2	1

A. Did you have any other concerns, related to your emergency, that you felt was not addressed by our personnel?  
\_\_\_\_\_

B. Please tell us the single most important action we took that made you feel better.  
\_\_\_\_\_

C. What could we have done differently that might have made your experience more positive?  
\_\_\_\_\_

Reading Fire Department

SCALE

### 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

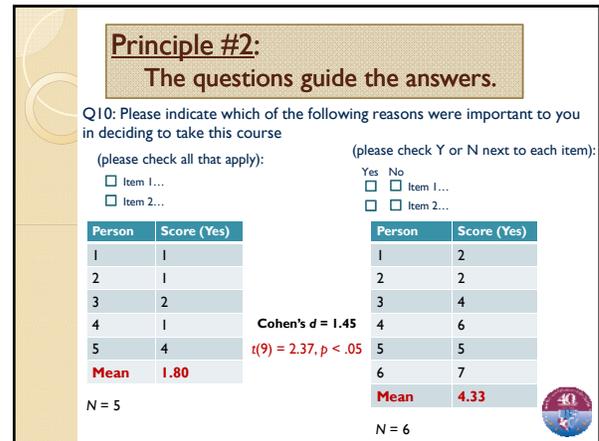
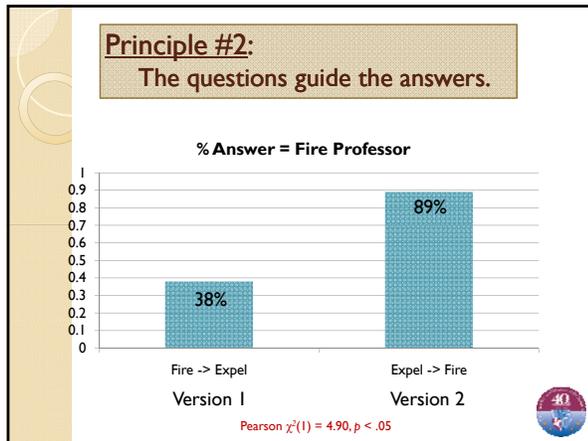
Ver 2: Small Amount of Space  
(1 line)  
(N = 6)

- 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





- 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 1: Literature Review**
- Goal: Ensure the construct is *relevant* in the field
- Critically evaluate the literature
    - How is the construct defined in prior studies?
    - Has the construct been evaluated sufficiently?
  - Identify existing scales
    - What items/scales currently exist?
    - Appraise quality

- Step 2: Interviews & Focus Groups**
- Goal: Ensure construct is what "real" people experience
- Interview experts
  - Create focus groups from target population
  - Apply open-ended questions
    - Avoid yes/no, multiple-choice questions

- Step 2: Interviews & Focus Groups**
- Example: *Educational quality* of a "redesigned" medical school curriculum
- Interview *experts*
    - How do "experts" define educational quality?
      - Experienced medical educators
      - Educators/administrators who have gone through redesign
  - Focus groups from the *target population*
    - How does the *target population* define educational quality?
      - Current students and recent graduates

### Step 2: Practical exercise

Group discussion:

- If I wanted to create a survey to measure the...  
*Educational quality of a redesigned medical school curriculum*

What aspects of the curriculum would I need to include/address/ask about?



### 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, order, other)



### Step 4: Develop Items (examples)

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

- How important was it for you personally to perform well in this course?
- How important were the practical applications of the information provided in this course?
- How important was the content of this course?
- 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
----------------------	--------------------	----------------------	-----------------	---------------------



### Step 4: Develop Items (examples)

Perceived Barriers to Mental Health Care (a belief; the full scale = 12 items)

Rate each of the possible concerns that might affect your decision to receive mental health counseling or services if you ever had a problem:

- I don't trust mental health professionals.
- I don't know where to get help.
- I don't have adequate transportation.

response anchors

strongly disagree 1	disagree 2	neutral 3	agree 4	strongly agree 5
------------------------	---------------	--------------	------------	---------------------

(from Hoge et al., 2004)



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

4) To what extent do you favor or oppose USU's 'Maximally Accessible Materials' (MAM) policy to make all printed materials at the school available upon request in enlarged font form for the visually impaired?

Strongly Oppose	Moderately Oppose	Slightly Oppose	Neither Favor nor Oppose	Slightly Favor	Moderately Favor	Strongly Favor
1	2	3	4	5	6	7

Total N = 17  
Mean = 5.1

8 said "neither favor nor oppose"  
6 said "slightly favor" or "moderately favor"  
3 said "strongly favor"!!

NO MISSING DATA!!!!



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

A note about providing a reason  
"because..." or "so that..."

Total N = 17

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

Yes = 12.5%

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

Yes = 55.6%

Pearson  $\chi^2(1) = 3.44, p = .06$

## Today's Small-Group Activity

- Form groups of 3 to 4
- Using a construct definition, develop a scale of 3 items to measure the construct

**Construct = Basic Clinical Skills Self-Efficacy**

- Definition:
  - A student's confidence in his/her ability to perform the basic clinical skills expected of a graduating medical student

## Today's Small-Group Activity

- Construct:
  - Basic Clinical Skills Self-Efficacy
- Definition:
  - A student's confidence in his/her ability to perform the basic clinical skills expected of a graduating medical student
- Tell us your items:
  - \_\_\_\_\_
  - \_\_\_\_\_
  - \_\_\_\_\_
- Tell us your response anchors:
 

?	?	?	?	?
---	---	---	---	---

## Common Pitfalls

- Creating double-barreled items
  - Example Item: "How effective is the inpatient and ambulatory teaching?"
    - What if one is good and the other is bad?
  - Solution: split into two items
    - "How effective is the inpatient teaching?"
    - "How effective is the ambulatory teaching?"

## Common Pitfalls

- Creating double-barreled items
  - Construct = Elaboration

	Q2_A_30	Q2_A_36	Q2_A_37	Q2_A_40	Q2_A_41	Q2_A_50
Q2_A_30	1.000	.094	.439	.277	.616	.297
Q2_A_36	.094	1.000	.741	.255	.087	.727
Q2_A_37	.439	.741	1.000	.217	.356	.683
Q2_A_40	-.371	-.765	-.717	1.000	.700	-.168
Q2_A_41	.616	.087	.356	.700	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

	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	25.88	13.076	.379	.745	.461
Q2_A_37	25.36	12.814	.608	.704	.261
Q2_A_40	26.84	17.004	-.176	.373	.733
Q2_A_41	26.14	12.028	.559	.644	.380
Q2_A_50	26.50	10.643	.698	.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)
    - "I can't stop thinking about the war in Iraq" (rarely-often)
  - Solution: make sure "yes" means "yes" and "no" means "no"
    - "In an average week, how often do you start rounds on time?"
    - Iraq:
      - Slight Improvement: "I keep thinking about the war in Iraq"
      - Better: "How often do you think about the war in Iraq?"

### Common Pitfalls

- Using statements instead of questions
  - Example Item: "I am confident I can do well on this rotation."

Not at all true	A little bit true	Somewhat true	Mostly true	Completely true
-----------------	-------------------	---------------	-------------	-----------------

- Statements are not very "conversational"
  - People are more practiced at answering questions
- Solution: use questions
  - "How confident are you that you can do well in this rotation?"
  - Use construct-specific confidence response anchors



### Common Pitfalls

- Using agreement response anchors
  - Example Item: "The high cost of health care is the most important issue in America today."

Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
-------------------	----------	---------	-------	----------------

- Agreement response anchors don't just measure the construct of interest
  - Confounded by how "agreeable" respondents are
  - Respondents often "agree" just because
- Solution: avoid agreement response anchors; maintain focus on construct by using construct-specific anchors
  - "How important is the issue of high healthcare costs in America today?"

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

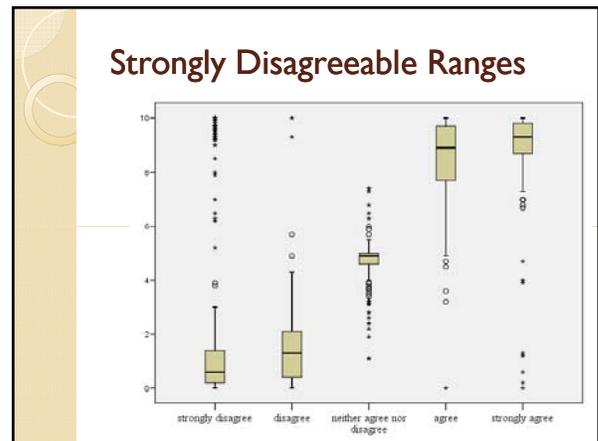
### Common Pitfalls

- 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.

100% Disagreement	_____	100% Agreement
-------------------	-------	----------------

### 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

### Common Pitfalls

- Using too few or too many response anchors

AAMC Graduation Questionnaire

6. - 1. Do you believe that your instruction in the following areas was inadequate, appropriate, or excessive?

Clinical Decision Making and Clinical Care

	Inadequate	Appropriate	Excessive
1. Patient interviewing skills	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Physical examination skills	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Diagnosis of disease	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



## Common Pitfalls

- **Using jargon and/or unclear language**
  - Instead, use language that is simple, direct, comprehensible, and unambiguous
  - Instead of **exhausted**... consider **tired**
  - Instead of **leisure time**... consider **free time**
  - Instead of **due to the fact that**... consider **because**
  - Instead of **at this point in time**... consider **now**



## Today's Small-Group Activity Debrief

Would anyone like to share some particularly good (or bad) items?



## 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



## Step 6: Cognitive Interviewing

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

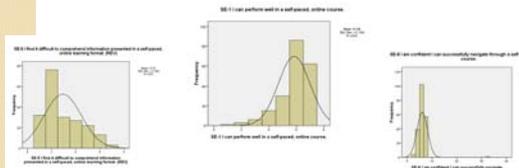
- Recruit members of the targeted population
  - e.g., students, teachers, patients, etc.
- Conduct one-on-one interviews, in “laboratory” or other location
- THEN: Make informed decisions, with cognitive interview as one source of input



## Step 7: Pilot Testing

**Goal:** Collect evidence of survey's reliability (score reproducibility) and validity (are you measuring what you intend to measure?)

- Collect data from a small sample
- “Get to know” your descriptive statistics
  - Are the item scores normally distributed?



## Step 7: Pilot Testing

- “Get to know” your descriptive statistics
  - Do individual survey items “hang together”?
  - Does your survey scale relate to other, more established scales?
    - You would expect ratings of “training quality” to be...
      - (+) correlated with “training satisfaction”; (-) correlated with “boredom”

		Correlations				Avg. Course Grade	NSMC_Overall
Task/Value		Satisfaction	Enjoyment	Anxiety	Boredom		
Satisfaction	Pearson Correlation	1.000	-.389**	-.506**	-.083	-.283**	-.267**
	Sig. (2-tailed)		.000	.000	.309	.001	.005
	N	109	109	109	109	109	109
Enjoyment	Pearson Correlation	-.399**	1.000	-.449**	-.168	-.222**	-.174
	Sig. (2-tailed)	.000		.001	.003	.011	.011
	N	109	109	109	109	109	109
Anxiety	Pearson Correlation	-.506**	-.449**	1.000	-.237**	-.300**	-.164
	Sig. (2-tailed)	.000	.000		.005	.000	.057
	N	109	109	109	109	109	109
Boredom	Pearson Correlation	-.083	-.168	-.237**	1.000	.026	-.248**
	Sig. (2-tailed)	.396	.000	.005		.762	.004
	N	109	109	109	109	109	109
	Pearson Correlation	-.267**	-.174	-.300**	-.248**	1.000	-.162



## Final Example

19) Parking is the most important issue at USU today.

group: group 2 positive (n)

19) How important is the issue of parking at USU today?

group: group 1 negative (n)

**Cohen's d = 1.16**  
 $t(15) = 2.31, p < .05$

strongly disagree	disagree	neutral	agree	strongly agree
not at all important	slightly important	moderately important	quite important	extremely important

## Questions?

If you remember nothing else, remember...

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

**Principle #2:** The questions guide the answers.

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

CDR Anthony R. Artino, Jr., PhD  
[anthony.artino@usuhs.mil](mailto:anthony.artino@usuhs.mil)

LTC Kent J. DeZee, MD, MPH  
[kent.dezee@usuhs.mil](mailto:kent.dezee@usuhs.mil)