



Survey Development 101

UNC Center for AIDS Research

August 30, 2013

Carol Golin

Catherine Grodensky



Catherine Grodensky, MPH
Core Manager



Carol E. Golin, MD
Core Director



Meheret Mamo, BA
Research Assistant



**Kemi Amola, PhD, and Steve
Bradley-Bull, LPC**
Counseling Trainers



Social and Behavioral Sciences Research Core



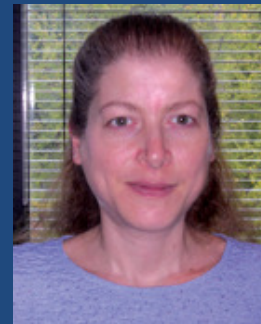
David Wohl, MD
Criminal Justice Expert



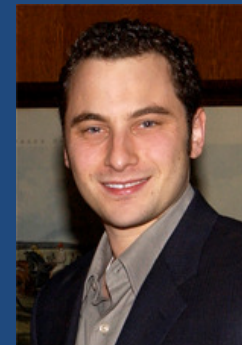
Randall Teal, MA
Qualitative Research,
Intervention Development
Specialist



David Rosen, MD, PhD
Epidemiologist,
Prison Database
Designer



Jennifer Groves, MBA
ACASI Programmer



Seth Noar, PhD
Health Communication
Specialist

What is a survey?

- A gathering of data from a sample of individuals considered to be representative of a whole (vs. census).
- The intent is not to describe the particular individuals who, by chance, are in the sample but to obtain a composite profile of the population.
- A data collection tool commonly used in public health to gather such information, especially self-report data from study participants:
 - Factual information (e.g., date of birth, health status)
 - Knowledge, attitudes or opinions (e.g., Perceived HIV risk)
- Collects information by standardized procedures so that every individual is asked the same questions in more or less the same way.

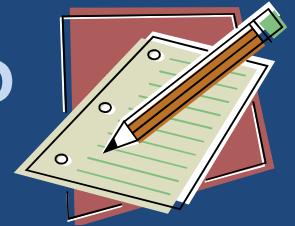
A Survey Starts with a Research Question

- **DO** Ask yourself, **before** you begin to design your survey:
 - **What** you want to know?
 - **Why** do you want to know this information?
 - **How** you will use the information you collect?
 - **How will this contribute** important knowledge?
- **DO** use your Research Questions to guide your survey.
 - How many people in County “X” have high HIV knowledge?
 - How does antiretroviral adherence affect viral load?
 - Is participant retention worse among patients with more structural barriers to HIV care?
 - Do participants with lower levels of HIV knowledge have higher levels of HIV-related stigma?
- **Do NOT** include items
 - That would just be “interesting to know”
 - On the off chance that they may come in handy
 - Because they are available

A Survey Starts with a Research Question

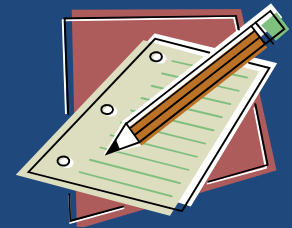
- **Independent variables:** “predictor variables” — their values or categories may be used to predict the values or categories of dependent variables.
- **Dependent variables:** those the researcher is interested in explaining and predicting
- **Potential Confounding Factors:** those that the researcher wants to control for that may confound the relationship between the IV/DV.

When are surveys useful?



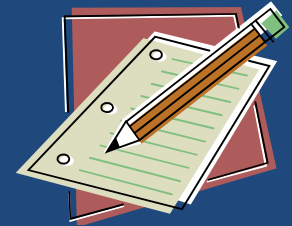
- You need a quick and efficient way of getting information.
- You need to reach a large number of people.
- You need statistically valid information about a large number of people.
- The information you need isn't readily available through other means (e.g., BP vs. symptoms).

How are surveys useful?



- In HIV research, useful to assess many domains:
 - Demographics
 - Medical (e.g., comorbidities, symptoms, health service utilization)
 - Psychosocial (e.g., emotional state, social support, attitudes)
 - Behavioral (e.g., med adherence, dietary, condom use)
 - Many others . . .
- Survey quality → Data quality

How do surveys work?

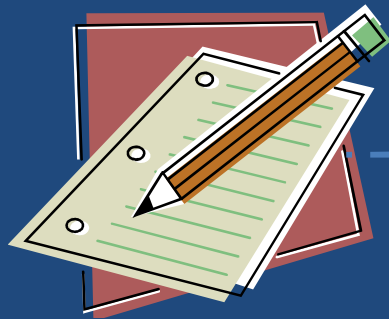


- “Typically, researchers are interested in *constructs* rather than items or scales per se.”
- “The underlying phenomenon or construct that a scale is intended to reflect is often called the *latent variable*.”

How do surveys work?

SURVEY
MEASURE
(scale,
item)

LATENT VARIABLE

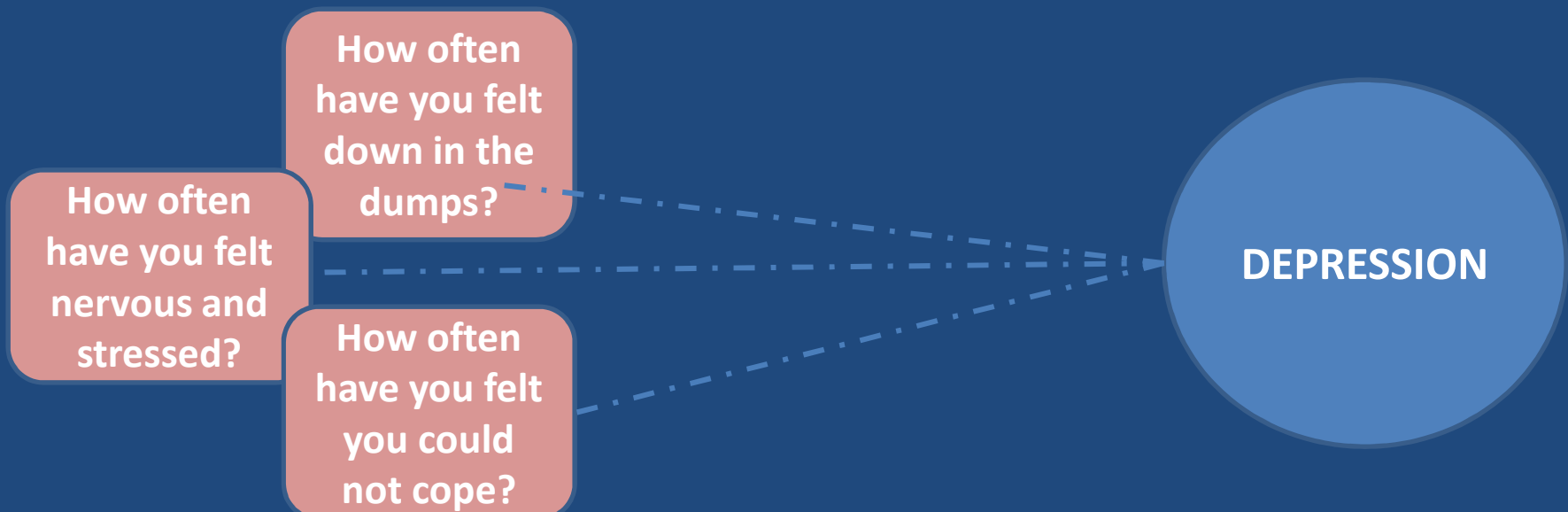


Why so many questions?

- Sometimes you can measure a construct with just one item (question)

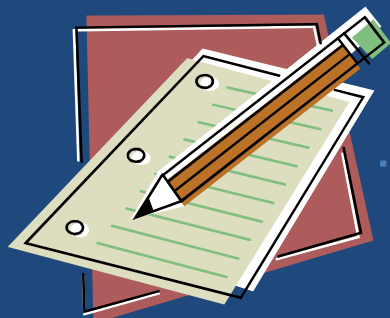
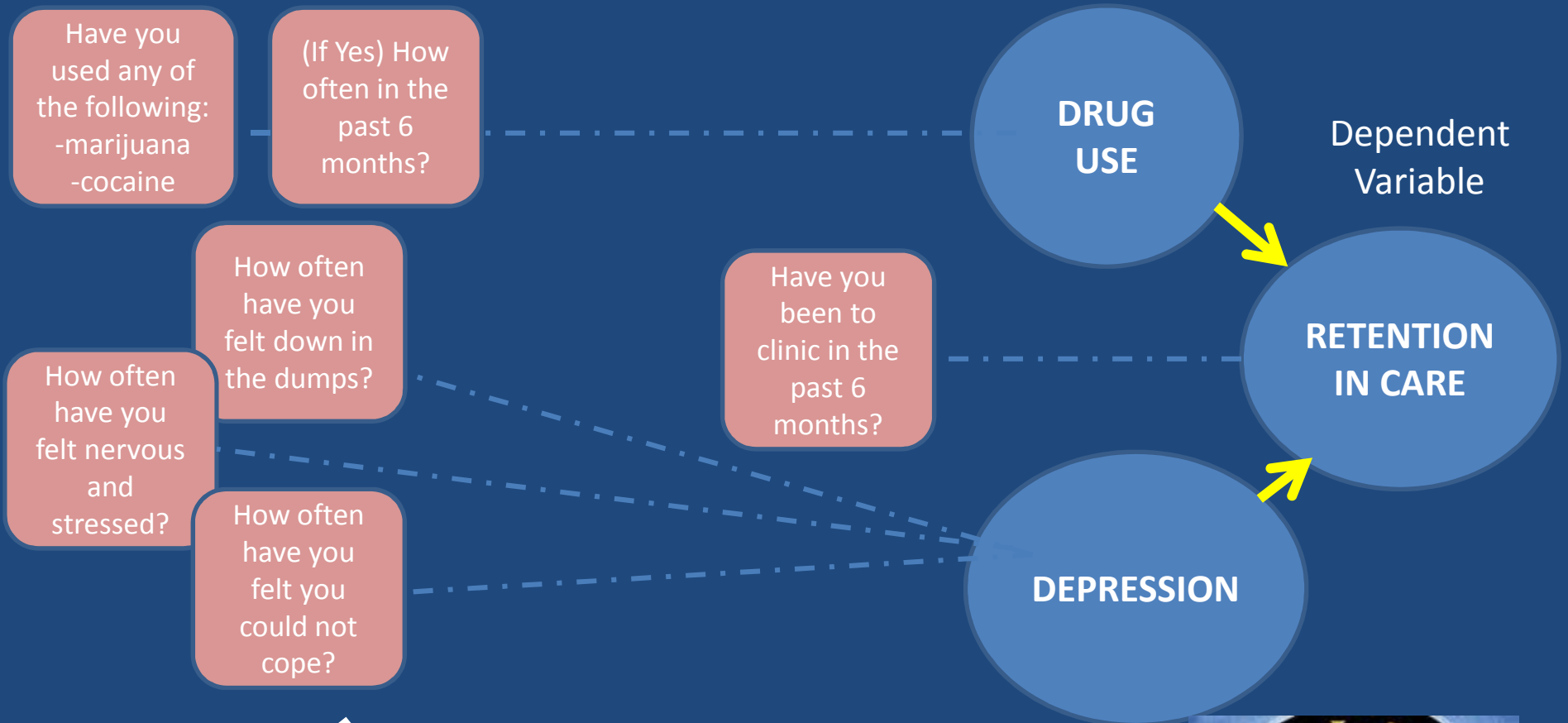


- For more complex constructs, the *precision* of measurement is increased with more items

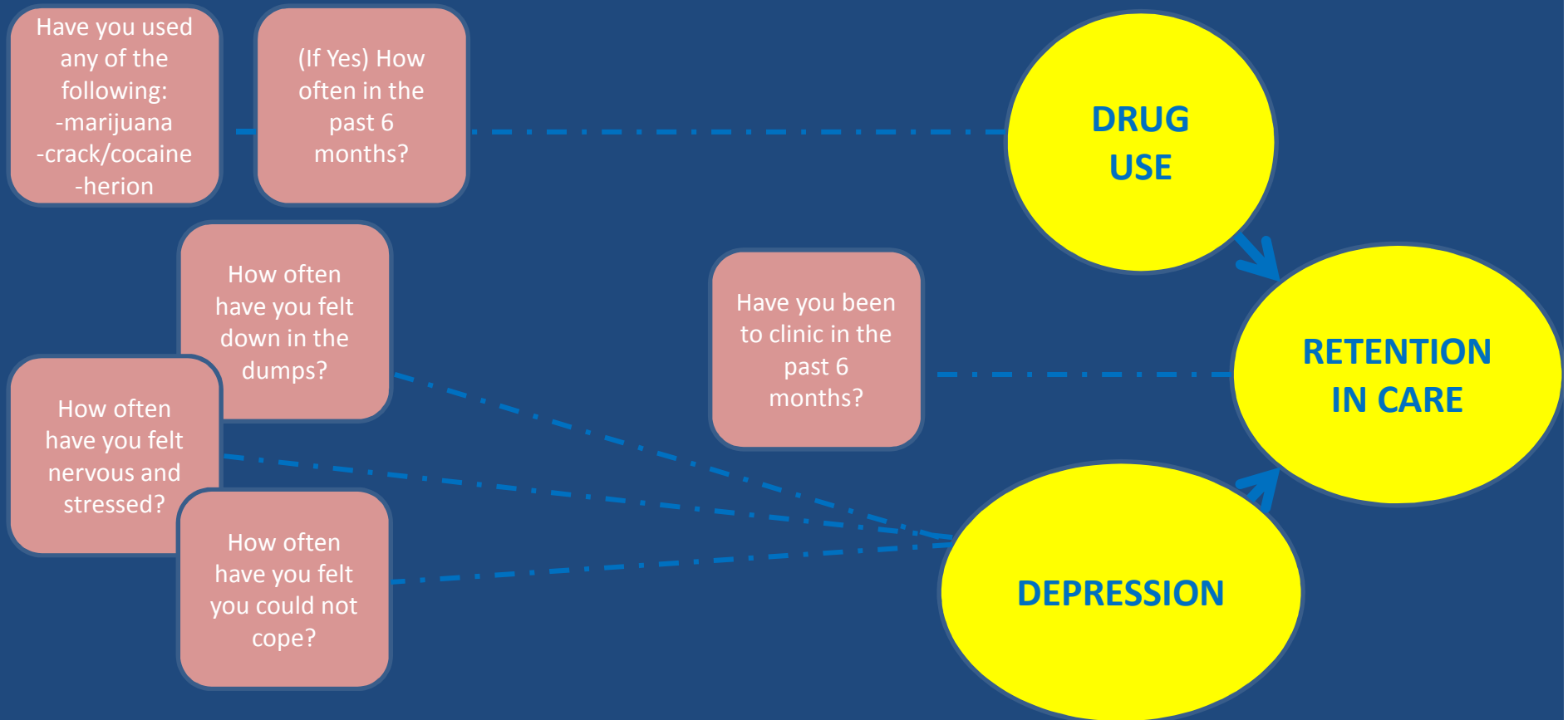


Many questions, many constructs

Independent Variables

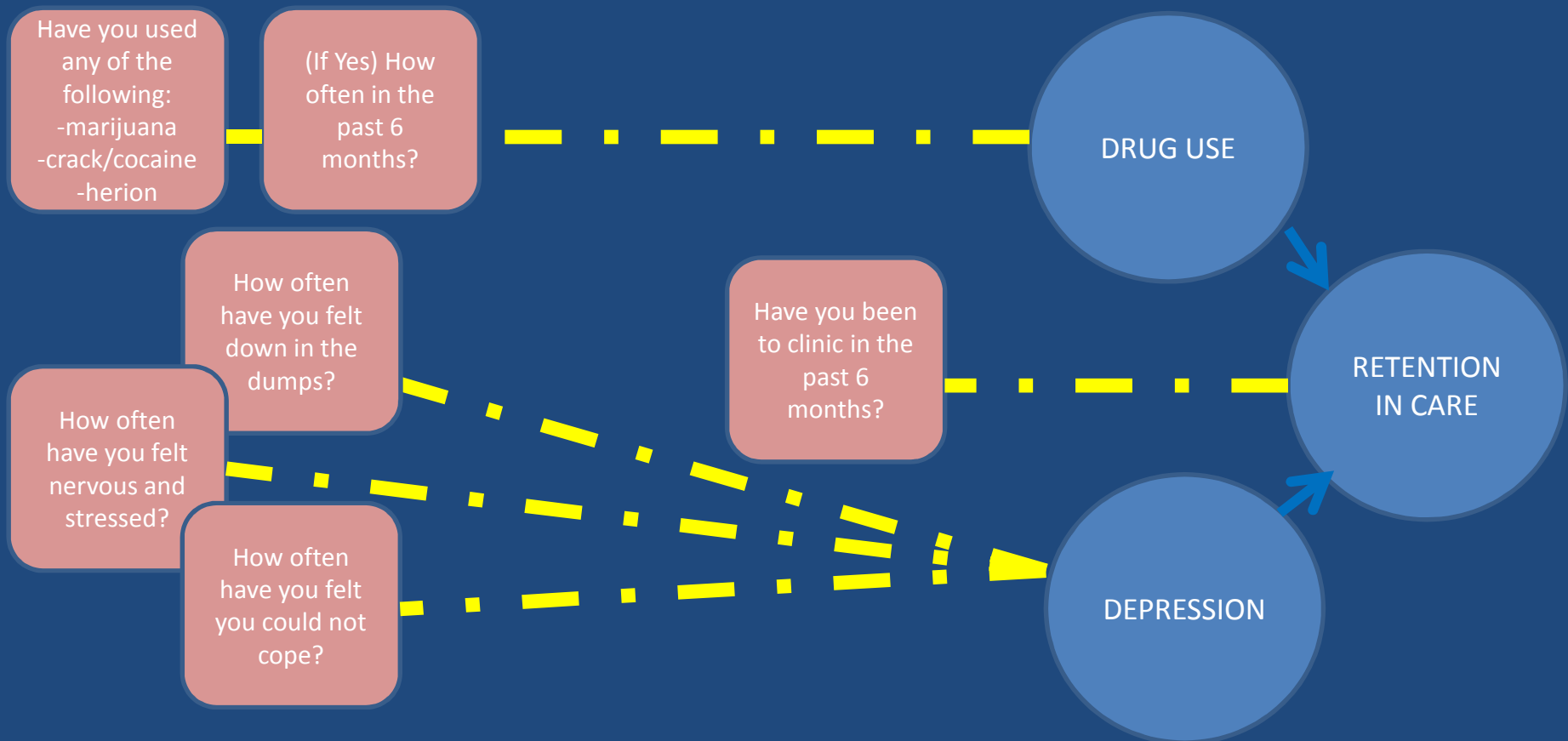


Survey Quality = Data Quality



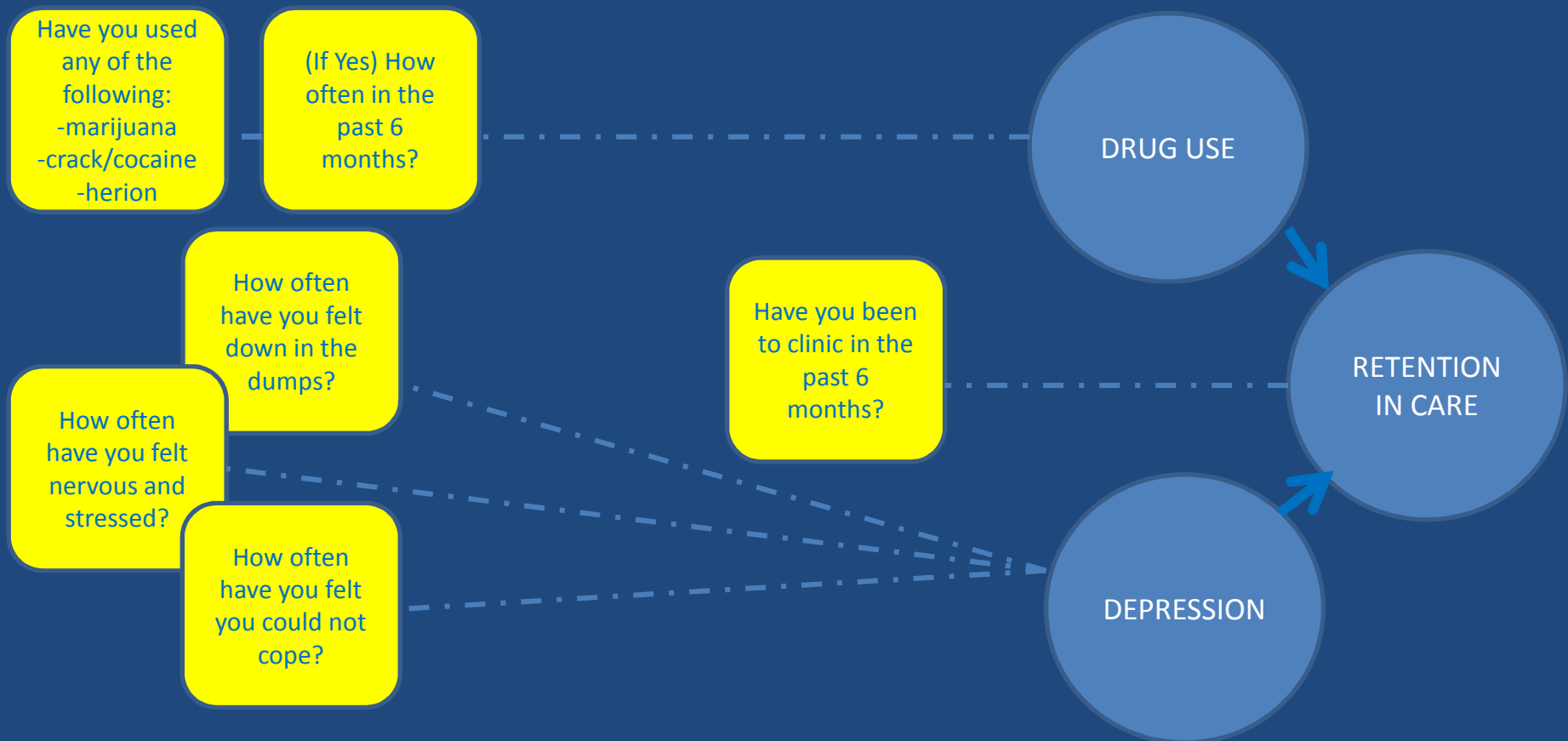
- Are these the most relevant latent variables for answering your research questions?
- Are you asking about all aspects relevant to the latent variables of interest?

Survey Quality = Data Quality



- Do the questions have any relationship with the latent variables you want to measure?
- Are the latent variables the SOLE cause of variation in responses to the survey items or is there error?

Survey Quality = Data Quality



- Does the respondent understand what the questions are asking?
- Does the respondent want to give you correct information?
- Is the respondent too tired / confused / offended to answer the questions correctly?

How do you develop a survey?

- Questions to consider:
 - **What constructs do you want to measure?**
 - How will you measure those constructs?
 - How will you administer the survey?
 - How will you test the survey in your target population?

What should you measure?

- “For any given research problem, the researcher can observe and measure only a **few** of the many potentially relevant properties.”
 - Research participants have limited time and energy to spend on a survey.
 - If the survey is burdensome, you may get missing or inaccurate data.
 - To get the best data possible, limit the constructs you measure in your survey to those that are ***relevant, reliable, and rare.***

What should you measure?

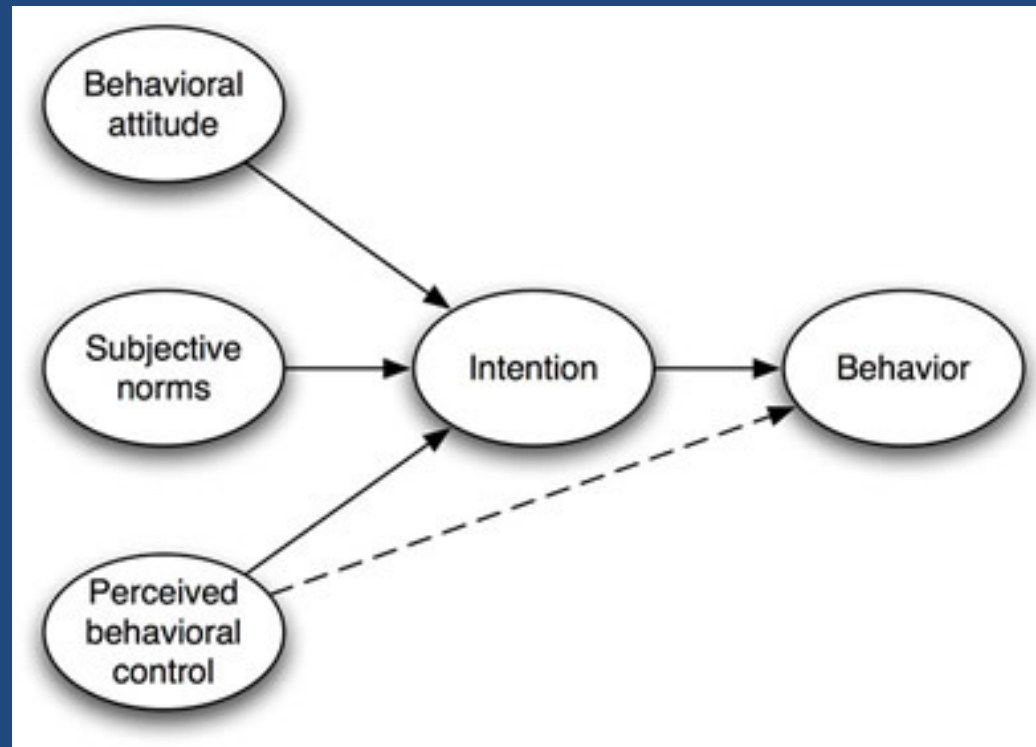
- Several sources can guide you in deciding what to measure:
 - Research questions
 - Theory
 - Conceptual Model
 - Previous research

What should you measure?

- Research questions:
 - How does antiretroviral adherence impact viral load?
 - Is participant retention worse among patients with more structural barriers to HIV care?
 - Do participants with lower levels of HIV knowledge have higher levels of HIV-related stigma?

What should you measure?

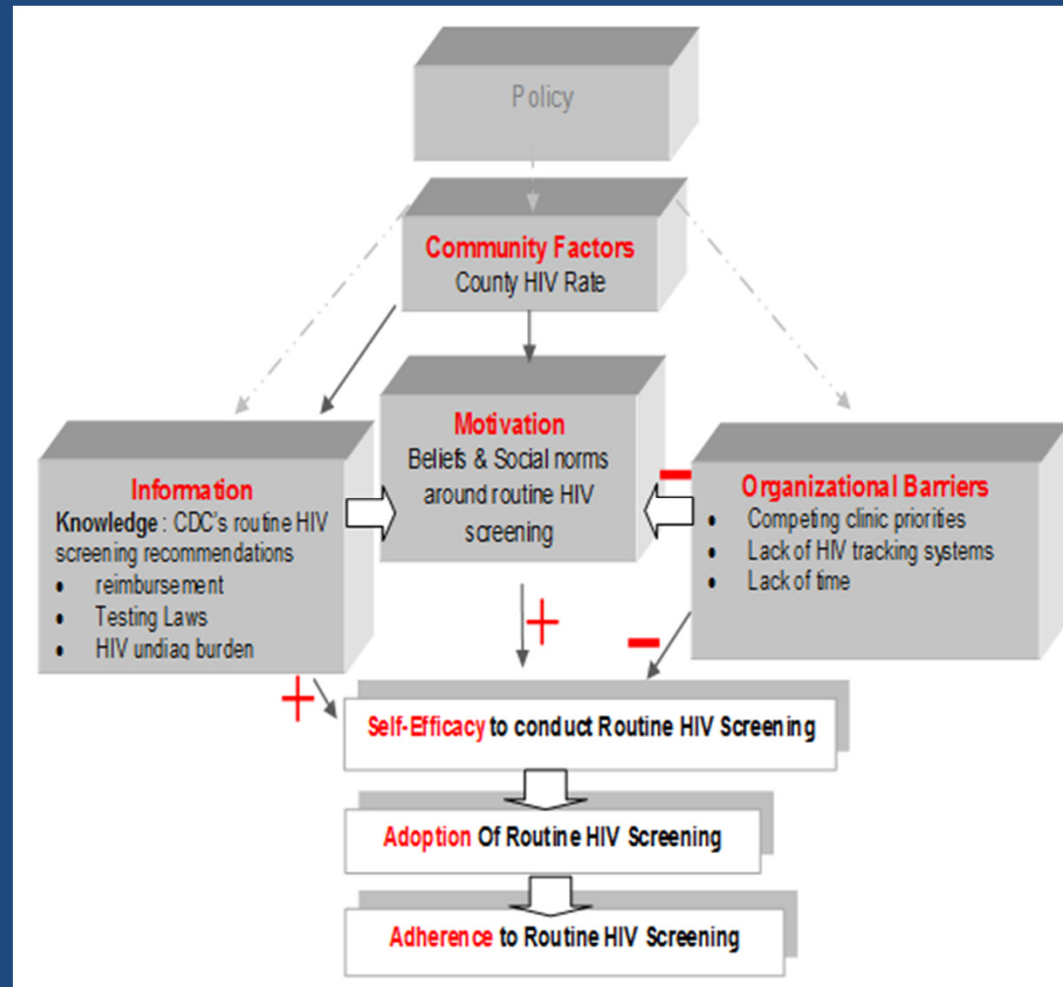
- Theory:



Theory of Planned Behavior (Ajzen 1991)

What should you measure?

- Conceptual model:



Courtesy of Dr. Becky White, 2013

What should you measure?

Prior research:

- You may want to focus on measuring factors that are *not* already well-demonstrated in the literature (why reinvent the wheel?)
- Focus instead on novel relationships, or on exploring novel dimensions of well-demonstrated relationships (i.e., mediation, moderation, dose-response, etc.)

How do you develop a survey?

- Questions to consider:
 - What constructs do you want to measure?
 - **How will you measure those constructs?**
 - How will you administer the survey?
 - How will you test the survey in your target population?

What measures should you use?

- You will need a measure for each construct (variable) you want to assess.
- Depending on your research question and the variable you seek to measure, sometimes **one item** may suffice. For example:
 - How old are you?
 - Last week, how many patients did you screen for problem drinking?
 - How satisfied were you with the program?

What measures should you use?

- For more complex questions, you may need **multiple items**; for example:
 - Drinking
 - Last week, how many patients did you screen for problem drinking?
 - How did you screen for problem drinking?
 - What resources are available for patients who screen positive for problem drinking at your clinic?
 - Satisfaction
 - How satisfied were you with the program as a whole?
 - How satisfied were you with the booklets you received?
 - How satisfied were you with your counselor?

What measures should you use?

- For more complex constructs, you may need a **scale or index**, such as:
 - Medication adherence (sum all items):
 - Do you ever forget to take your medications? (Y/N)
 - Are you careless at times about taking your medications? (Y/N)
 - When you feel better, do you sometimes stop taking your medications? (Y/N)
 - Sometimes, if you feel worse when you take your medications, do you stop taking them? (Y/N)

Morisky, DE, Green, LW, Levine, DM. (1986). Concurrent and predictive validity of a self-reported measure of medication adherence.

What measures should you use?

- **Scales and Indexes:** composite measures containing multiple items that reflect a common underlying construct (e.g., depression, social support, quality of life)
- Can include subscales; for example:
 - Stigma: internalized, discrimination, etc.
 - Quality of life: emotional, physical, etc.

Scales

- Multiple items that are combined into one or more summary scores
- Each item is an indicator of the **same** variable of interest, with multiple items triangulating in on that variable and increasing precision
- Because items assess the same variable, they should correlate with one another
- Answers to items are “caused” by the same (latent) variable, e.g., perceived risk of infection

Indexes

- Multi-item composites based on items **sharing a common effect rather than a common cause**. (e.g., number of risky behaviors a person engages in → greater HIV risk)
- Items are similar in their consequences but don't necessarily share a common cause
- Items typically are not correlated with one another

Contrast between Scale and Index

- “Perceived risk scale” is an attribute of the person.
E.g.: “I am concerned about contracting HIV,” “I avoid people who may have HIV out of fear of catching it.”
Sum score would quantify how the respondent perceives risk.
- “Risk index” quantifies the number of risky behaviors.
E.g.: “I inject IV drugs.” “I have multiple sex partners.”
“I frequently have unprotected sex.” Each contributes a quantum of risk, but they may not be correlated nor do they necessarily arise from a common underlying characteristic of the respondent.

What measures should you use?

- Options for measures include:
 - Using an existing measure
 - Adapting an existing measure
 - Developing a new measure (de novo)

Finding existing measures

- To find options for existing measures, you can:
 - Ask an expert in the field
 - Conduct a literature search on the construct you wish to measure, and read the methods for measurement (often the measure itself is included in the article or a cited source article)
 - Search the UNC CFAR Social and Behavioral Instruments (SABI) Database at www.cfarsbcdirectory.unc.edu/sabi/index.asp
 - Search another measurement repository (e.g., UCSF CAPS <http://caps.ucsf.edu/resources/survey-instruments/>, PROMIS www.nihpromis.org/, etc.)

Evaluating existing measures

- Don't assume that the name of a scale accurately reflects what it measures
- Use empirical criteria to determine if a scale is suitable
- Examine the evidence supporting the claims made for the measure

Evaluating existing measures

When evaluating a measure for your study, it is important to consider:

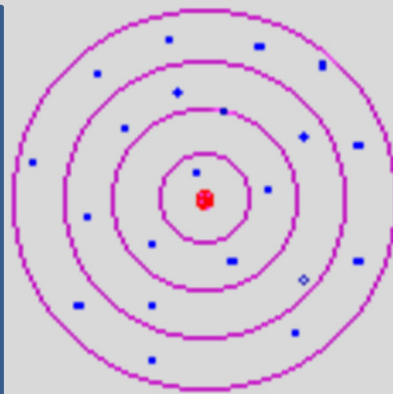
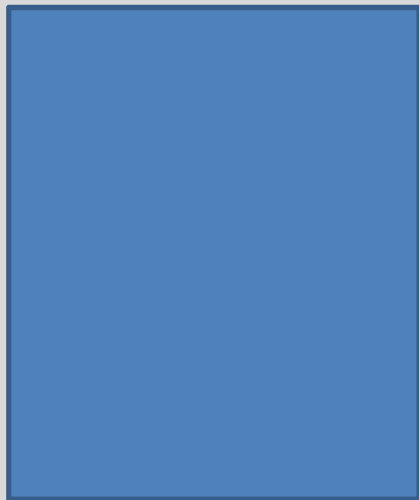
- How has the measure performed in past research?
 - **Reliability**: The extent to which a group of items are capturing the same construct. More technically, reliability is the extent to which the items vary as a function of the same latent variable.
 - **Validity**: The extent to which a group of items measures what it claims to measure – the extent to which the latent variable measured is the one intended.
- How does it perform in your target population?

Evaluating existing measures

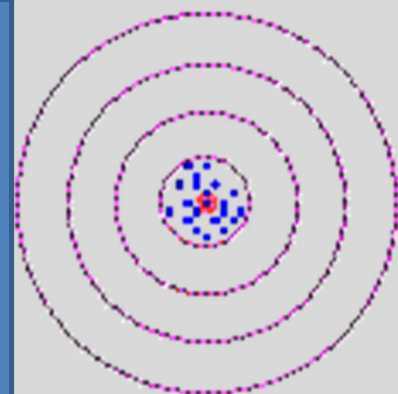
Reliability: produces similar results under consistent conditions (absence of random error). For example,

- The extent to which results **remain consistent over time** (temporal stability, or test-retest reliability). This is only meaningful if you can safely assume no change in the levels of the latent variable over the time interval.
- The **homogeneity of items within a scale** (Internal consistency reliability—Cronbach's alpha). Alpha is driven by (a) how closely the items correlate to each other and (b) the number of items.

Reliability



Not Reliable



Reliable

Evaluating existing measures

Validity: measures what it claims to measure
(accuracy; absence of systematic measurement error)

- **Criterion Validity**--Has an empirical association with some criterion or “gold standard”

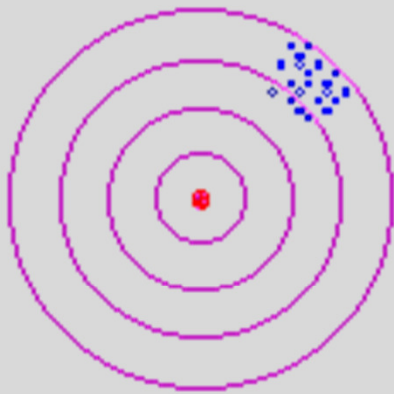
E.g. Arterial Line BP criterion for BP cuff

- **Content Validity**--Reflects all (and only) relevant aspects of the domain.

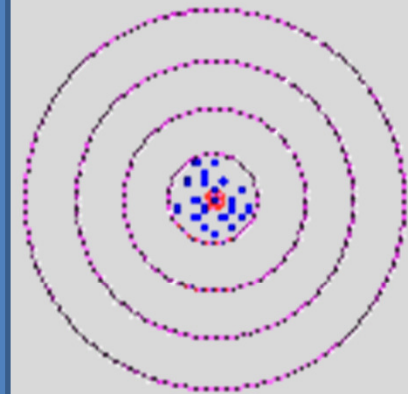
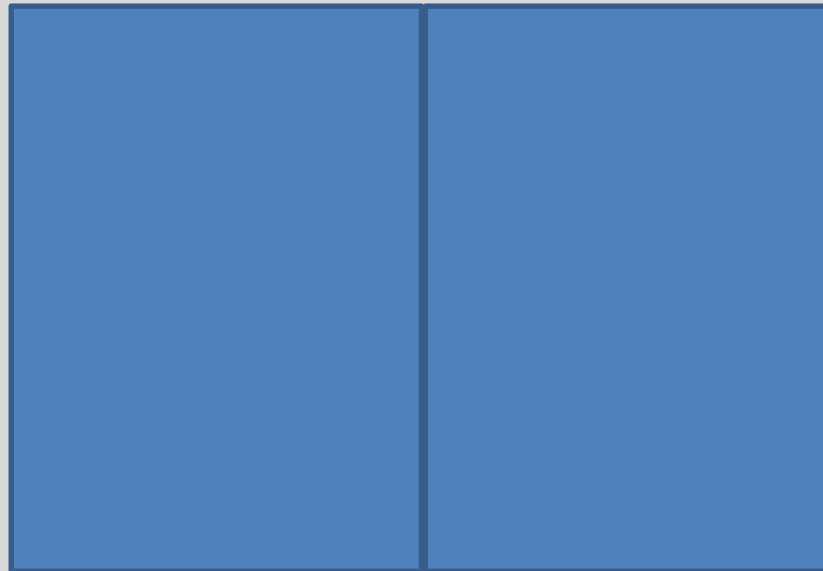
- **Construct validity**--Is related to other variables in a way that relevant theory would predict

E.g. Lower HIV Stigma associated with “knowing someone living with HIV”

Validity

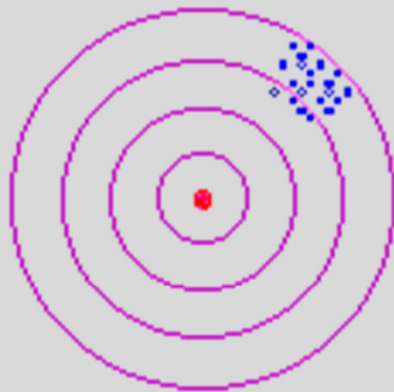


Not Valid

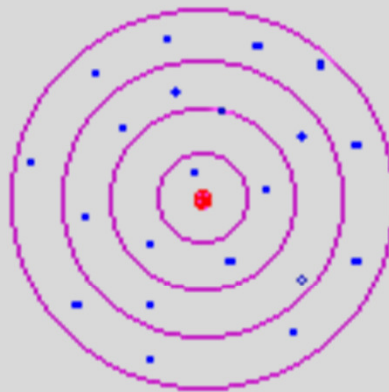


Valid

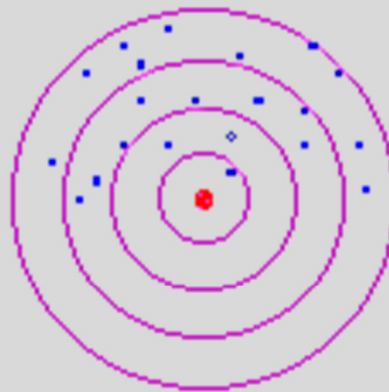
Reliability and Validity



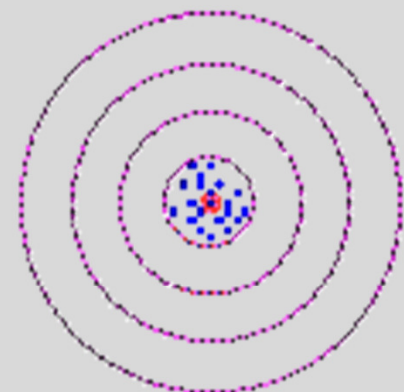
**Reliable
Not Valid**



**Valid
Not Reliable**



**Neither Reliable
Nor Valid**



**Both Reliable
And Valid**

Evaluating existing measures

- Evaluate how well your measure has performed in the past:
 - Search for other studies that measured the construct
 - Literature searches
 - SABI (Social and Behavioral Instruments) Database
 - Pay attention to what the studies report about the measure's reliability and validity (usually reported in methods, or in results if the purpose of the study is to evaluate the scales, or in a “methods paper” cited.
 - Consider how similar the study population and setting are to your own—the more similar they are, the greater the likelihood that the measure will perform similarly in your study

When existing measures won't work

- Sometimes (especially if you're really onto something new!) you will not find any existing measures that fit with your research questions. In this case, you can either:
 - Adapt an existing measure
 - Develop a new measure (de novo)

Adapting existing measures

- You may want to adapt a measure if it:
 - Does not capture exactly what you want
 - Contains too many items (but be careful!)
 - Contains components that seem irrelevant or redundant
 - Contains language that will not suit your target population

Adapting existing measures

- Be warned!
 - When you change an existing measure in any way, you cannot count on it to be reliable and valid
 - The scale development process taps into aspects of the construct that may not be readily obvious to end users

Adapting existing measures

- Shortening scales is particularly risky
 - Scale precision increases as a function of the number of indicators (items)
 - A scale that is too short will be error prone (just as a study with a too small sample is error prone)
 - Unless a scale can provide reliable and valid information, it is of no use
 - Convenience is not a sufficient reason for shortening a scale. (Most studies would be more convenient if the sample were cut in half, but precision would be compromised!)
 - If a survey is too long, it may be better to measure fewer variables than to try shortening one or more scales

Adapting existing measures

- If you are going to adapt a measure, you can:
 - Search the literature for existing variations (i.e., short forms, different health topic / behavior focus, etc.)
 - Look for studies that have reported subscales or done factor analysis on the scale to determine if there are item groupings
 - Edit the items yourself (this is the most risky!)
- You should pilot and/or evaluate the reliability and validity of your adapted measure before reporting its use

Creating new measures

- When creating a measure, keep in mind:
 - Response scales
 - Question wording
 - Question “loading”
 - Recall accuracy

Creating new measures: Types of response scales

What are your concerns about your daughter receiving the HPV vaccine?

- Open-ended
-

- Closed-ended

- She may become sexually active
- She may have side effects
- I don't believe the vaccine is effective in preventing HPV

- Partially Closed-ended

- She may become sexually active
- She may have side effects
- I don't believe the vaccine is effective in preventing HPV
- Other: _____

Open-Ended Responses

Advantages

- Useful in exploratory work
- Researcher does not need to know the universe of possible answers
- Respondent not influenced by specific alternatives suggested
- Can be used to build rapport in an interview
- Get more in-depth information

Open-Ended Responses

Advantages

- Useful in exploratory work
- Researcher does not need to know the universe of possible answers
- Respondent not influenced by specific alternatives suggested
- Can be used to build rapport in an interview
- Get more in-depth information

Disadvantages

- Require lots of effort, esp. if written (respondents may skip)
- Respondents may vary in ability/willingness to articulate (e.g., sensitive topics, socially unacceptable opinions or behaviors)
- Data may be irrelevant, vague, or easy to misinterpret
- Difficult, time-consuming & expensive to code/analyze responses
- Can't ask as many questions because survey will take too long

Closed-Ended Responses

Advantages

- Easier for respondents
- Focuses respondents on the topic
- Limits “don’t know” and non-responses
- Easier to standardize time frame of reference and other aspects
- Less variability in interviewer performance
- Less time to administer and record responses
- Easier to score and analyze resulting data
- Get greater coverage of information

Closed-Ended Responses

Advantages

- Easier for respondents
- Focuses respondents on the topic
- Limits “don’t know” and non-responses
- Easier to standardize time frame of reference and other aspects
- Less variability in interviewer performance
- Less time to administer and record responses
- Easier to score and analyze resulting data
- Get greater coverage of information

Disadvantages

- Need to know appropriate response categories in advance
- Respondent may be forced into an unnatural frame of reference
- Potentially non-accurate information. (e.g. easier to get a response that you *think* is meaningful but is not)
- Respondent may not feel as involved or motivated by questionnaire
- Lacks depth of open-ended questions

Creating new measures: Question wording

- Include only one idea per question—
 - e.g., No “double-barreled” questions such as “How much did you agree or disagree with what your **nurse and doctor** told you?”
- Give an appropriate time frame of reference
 - e.g., “In the last two weeks”
- Use plain language and appropriate reading level
 - 2 grades below avg. reading level of your population
 - 6th grade reading level is typical
 - Use vocabulary appropriate for your audience
 - Assess Reading Level using Microsoft Word

Creating new measures: Question wording

- Don't use leading questions (unless you *want* to bias responses)
 - This clinic was recently ranked number 1 in patient care. How was your experience today?
- Don't use double negatives
 - Do you think it is unsafe not to use condoms when having sex with a stranger?
- Don't provide incomplete or overlapping response categories
 - Are you white, black, or Latino?

Creating new measures: Question wording

“Unload” loaded questions to minimize social desirability bias:

As you may know, there is now a greater discussion about shop lifting in this community. Some people feel it is a serious problem and others feel it is not a serious problem. How about yourself?

1. Do you consider shop lifting to be a serious, moderate, slight, or no problem at all in this community?

Serious Moderate Slight No problem at all

2. During the past few years, do you think the frequency of shoplifting has increased, stayed about the same, or decreased in this community?

Increased Stayed about the same Decreased

Creating new measures: Question wording

Improve recall accuracy:

- Shorten the reference period (but remember that this limits the time period to which you can apply the data)
- Use temporal “landmarks” to aid recall of dates
 - e.g., Timeline follow-back: Ask client to mark on a calendar the significant events in their lives during a time period; then ask them to report drinking behavior on the same calendar.

Creating new measures: Scale Development

- This is a lengthy process involving:
 - Generating an “item pool”
 - Determining the format for measurement
 - Having the item pool reviewed by experts
 - Administering the items to a development sample (focus group, pilot study)
 - Evaluating the items (correlations, variance, factor analysis, internal consistency)
 - Optimizing scale length (dropping “bad” items)
- **Don't try this at home!!**

Building your survey

- It can be helpful to keep track of the constructs you are measuring with a table (this can also be helpful to include in a grant application)

Building your survey

Construct (IV/DV)	Time points	Items	Response options
Viral load (DV)	Baseline, 6m, 12m	-Have you had a viral load test in the past 6 months? -At that time, were you undetectable?	No, Yes, Don't know, Refuse to answer
Incarceration history (IV)	Baseline	-Have you ever been in prison or jail?	No, Yes, Refuse to answer
Stress (IV)	6m, 12m	How much stress have you experienced over the past 6 months from the following: -Started disability -Death of family member -Discrimination -Change in viral load or T cell count -Fired or laid off from a job	No stress, A little stress, Some stress, A lot of stress, Did not happen in the last 6 months, Refuse to answer

Building your survey

- Choose the order of your survey items carefully—optimizing order can reduce missing data and decrease respondent burden
 - Early: Inviting, non-threatening, interesting questions
 - Middle: Most important questions, and those requiring a lot of thought
 - End: Most sensitive questions, easy questions that can be answered with minimal effort (by now they may be fatigued!)

Building your survey

- Once you have compiled and ordered measures for all constructs, create your survey document, including:
 - Introduction/framing: *Why am I being asked to do this survey? How long will it take?*
 - Instructions: *Which questions am I supposed to answer? What should I be thinking about when I answer?*
 - Transitions: *What am I being asked about now? Do I need to shift my focus?*

How do you develop a survey?

- Questions to consider:
 - What constructs do you want to measure?
 - How will you measure those constructs?
 - **How will you administer the survey?**
 - How will you test the survey in your target population?

How to administer the survey?

- Decide which mode of administration is best for your survey (and budget)
 - Face-to-face, computer-assisted or unassisted
 - Telephone
 - Mail
 - Internet
 - Audio Computer Assisted Self-Interview (ACASI)
 - Computer-Assisted Personal Interview (CAPI)
 - Computer-Assisted Telephone Interview (CATI)
 - Mobile device

How to administer the survey?

- Access resources to develop your survey
 - UNC CFAR (QDS, Qualtrics)
 - UNC Survey Research Unit
 - UNC Odum Institute
 - Qualtrics
 - RedCap
 - MediaLab

How do you develop a survey?

- Questions to consider:
 - What constructs do you want to measure?
 - How will you measure those constructs?
 - How will you administer the survey?
 - **How will you test the survey in your target population?**

Why test your survey?

- It is STAGGERING how many surveys go into the field with:
 - Typos
 - Vague language
 - Incorrect instructions or skip patterns
 - Items that are not asking what the researcher is intending to measure
- Such mistakes lead to (at best) respondent annoyance / lack of confidence in the survey, and (at worst) to missing or misleading data

Testing your survey

- Review by knowledgeable colleagues and analysts
- Cognitive interviews with a sample similar to your target population
- Small pilot study, if possible
- Final check by people inside and outside the study

Initial Reviews

- Carefully read the entire survey multiple times **yourself**, in the mindset of a respondent
- Let **key players** on your study read the entire survey, in the mindset of a respondent
 - Study leadership can provide insight into your research questions and constructs at this point
 - Data collectors may be able to anticipate problems respondents may have in understanding the questions, or ambiguities in the instructions
- If the survey is programmed, identify several people to test the survey logic to ensure skip patterns are working correctly (may want to develop matrices or scenarios to guide testing)

Cognitive interviewing

Cognitive interviewing can be used to:

- improve the ability of a survey to collect meaningful data
- ensure that participants are interpreting questions as intended
- explore the cognitive processes that respondents use to answer survey questions

Cognitive interviewing

Cognitive interviewing helps to answer:

- How do respondents understand the survey question?
- Do different respondents understand the survey question differently?
- Does the question mean the same in all of the cultures that it is asked?
- In processing a question, do all respondents recall information and form an answer the same way?
- How much effort is required of respondents to answer each survey question?

Cognitive interviewing

- Think-aloud
 - Subjects are explicitly instructed to "think aloud" as they answer the survey questions
 - The interviewer records and/or otherwise notes the processes that subject uses in arriving at an answer to each question
 - The interviewer interjects little else, except to say "tell me what you're thinking" when the subject pauses
- Verbal probing
 - After the interviewer asks the survey question, and the subject answers, the interviewer then asks for other, specific information relevant to the question, or to the specific answer given
 - In general, the interviewer "probes" further into the basis for the response
 - Verbal probing can be done either as the survey goes along or at the end of the survey

Cognitive Interviewing Example

- Survey question: “What percentage of North Carolinians are infected with HIV?”
- In CI, participants are asked “How confident are you of your answer?”
 - All respond “I totally guessed.”
- PI decides to remove the question from the survey—data would be misleading and not useful

Cognitive Interviewing Example

- Survey question: “How often did you see a doctor in the past month?”
- In CI, participants are asked “Which doctors did you consider in coming up with your answer?”
 - Several respond that they only included their primary care doctor in their answer
- PI decides to add to the question to clarify the types of providers: “Please include any visits with doctors including your primary care provider, specialists, or other medical doctors.”

In Conclusion...

- Survey quality = data quality
- The more carefully you plan your survey:
 - The more useful and user-friendly your data
 - The more pleasant the experience for your participants
 - The more rigorous and meaningful your research

Contact CFAR for Assistance

We love to help with your surveys!

- Visit <http://cfar.med.unc.edu/content/social-and-behavioral-science-core> and click “Request services”
- Email Catherine at grodensk@med.unc.edu
- Visit the SABI at www.cfarsbcdirectory.unc.edu/sabi/index.asp

Acknowledgements

- Many thanks to the following individuals who provided content and feedback to this presentation:
 - Noel Brewer
 - Robert DeVellis
 - Becky White
 - David Rosen
 - UNC CFAR Biostatistics Core