

DESIGNING & TESTING QUESTIONNAIRES

Tuesday, December 9, 2008

Jack M. Guralnik, M.D., Ph.D.

**Laboratory of Epidemiology, Demography and Biometry
National Institute on Aging**

Outline

- I. Determine Data to be Collected
- II. Questionnaires vs. Interviews
- III. Questionnaire Design
- IV. Wording of Questions
- V. Data Forms
- VI. Data Entry

- I. Determine Data to be Collected
(independent variables and outcomes)
 - A. Factors whose variability in study population may explain variability in dependent variable (outcome)
 - B. Demographics (age, race, sex, education)
 - C. Time in study (date of admission, examination, contact)

D. Other descriptive variables to define study population and allow others to compare to yours (insurance status, urban/rural, etc.)

E. Variables important to testing hypothesis

F. Practical considerations

1. Limited by money, time, participant burden
2. Incompleteness rate increases with length of interview/questionnaire

3. Priority should be given to variables of interest

4. Important variables should have redundant questions to assess validity and consistency

G. Caveat - easy to be overly inclusive, ***BE SURE*** hypotheses and investigators identified so data will be used once collected

H. Consider use or modification of existing instruments

II. Questionnaires vs. Interviews

(after Hulley SB and Cummings SR, *Designing Clinical Research*, Williams and Wilkins, Baltimore, 1988)

A. Advantages of questionnaires

1. Economy: ability of participants to complete questionnaires on their own reduces staff time
2. Convenience: if mailed to home, participants can complete at leisure and obtain necessary corroborating information such as dates and location of hospitalization

3. Standardization: written instructions reduce biases from differences in administration or interactions with interviewer

4. Anonymity: privacy encourages candid and honest responses to sensitive questions

B. Advantages of interviews

1. Ease for participant: avoids difficulty with poor vision or low literacy, may be more enjoyable

2. Clarity: interviewer can clarify questions and responses as needed
3. Richness: interviewer can collect more complex answers and make observations about respondent's appearance and behavior
4. Completeness: interviewer can minimize missing and inappropriate responses

C. Method of administration

1. Questionnaires can be mailed or given to subjects in person
 - a. Mailed questionnaires cheaper, offer wider population, but much lower response rate
 - b. Distributed questionnaires allow researcher to clarify questions and review completeness
 - c. Questionnaires to reduce burden of interview – select parts of assessment that are easiest to do by written questionnaire

2. Interviews can be in person or by telephone
 - a. Telephone much less expensive, but subjects must have phones and be able to hear and converse clearly for extended periods
 - b. In-person interviews allow direct observation, examination

III. Questionnaire Design

A. Sequence of questionnaire

1. Identifier page
2. Instruction page with clear directions
3. Annoying, sensitive questions best at end

- B. Instrument should begin with brief description of purpose of study or questionnaire if not already explained
- C. To ensure standardized responses, all instruments (self- or interviewer-administered) should have instructions on how to complete them

Specs – detailed instructions in Manual of Procedures on all possible contingencies, exceptions

- D. To improve "flow," group questions on major subject area together and introduce with headings or short descriptions

E. Begin with simple questions about age, sex, date of birth to "warm up" participant to process

F. Place questions with similar formats together if possible; when changing formats be sure to indicate clearly how to respond

1. For interviews, for questions with multiple possible responses use cards to show choices

G. Nature of questions

1. Clear, unambiguous
2. Simple language, appropriate to level of education
3. Use several questions to measure some important point
4. Use standardized, validated questionnaires whenever possible

H. Types of questions

1. Close-ended
 - a. Small number of possible answers
 - b. All possibilities covered

- c. If scaling an opinion, use even number of options (agree strongly, agree, disagree, disagree strongly)
- d. Disadvantages: lead respondent in certain directions, may not always include response most appropriate for given respondent
- e. If not all possibilities can be specified, include space for "Other (specify _____)"
- f. Responses should be mutually exclusive unless specify more than one answer and can deal with coding them

g. Avoid instructing respondent to mark "all that apply" as approach does not force respondent to consider each possibility and unmarked item may not apply or simply have been overlooked

Which of the following increases the chance of having a heart attack?

		Yes	No	Don't Know
— Smoking	Smoking	[]	[]	[]
— Being Overweight	Being Overweight	[]	[]	[]
— Stress	Stress	[]	[]	[]

2. Open-ended

- a. Subject responds freely
- b. Very difficult if not impossible to code and analyze
- c. Often useful to give study subject chance to expand or clarify, but rarely can be used in data analysis
- d. Good in pilot studies to help create close-ended questions

I. Questionnaire format

1. Visual design and attractiveness is important - make as easy as possible for respondent to complete all questions in correct sequence
2. If too complex, respondents or interviewers may skip questions, provide incorrect information, or refuse to complete
3. Neat format with plenty of space-avoid crowding and cluttering

4. Large type particularly for elderly
5. Answers to closed ended questions should be lined up vertically with boxes or brackets (not open blanks) for responses

6. Avoid categorizing continuous variables if not necessary

Number of different medicines taken daily
(check one)

- None
- 1-2
- 3-4
- 5-6
- 7 or more

More information can often be obtained in less space with continuous response:

Number of different medications taken daily:

[_ _]

7. Use branching questions if need to follow up some answers with more detailed questions

Have you ever been told that you have high blood pressure?

Yes



If yes, about how old were you when you were first told that you had high blood pressure?

No

Years Old



Go to Next Page

- J. Completion time should be considered carefully and evaluated in pilot
 - 1. Phone - 20 minutes
 - 2. Interview – 30 minutes ideal, up to an hour reasonable
 - 3. Questionnaire – less than 1 hour

IV. Wording of Questions

- A. Clarity: avoid abstract concepts or words requiring judgment
 - 1. "How much exercise do you usually get?" contains "exercise," an abstraction, and "usually," a judgment
 - 2. "How many flights of stairs do you climb during a typical day?"

B. Simplicity

1. Avoid complex words, technical terms and jargon (very difficult for researchers to do)
2. Simple sentences: fewest words and simplest grammatical structure possible
3. Aim for sixth- to eighth-grade reading level for educated populations; fourth-grade level for barely literate populations

4. Avoid "loaded" terms and stereotypes suggesting there is a most desirable answer
 - a. "During the last month, how often did you drink an excessive amount of alcohol?"
 - b. "During the last month, how often did you drink 5 or more drinks in one day?"

5. Set a tone that permits respondent to admit potentially unacceptable behaviors

Introduction: "People sometimes forget to take the medicines prescribed by their doctor.

Broad question: *"Do you sometimes forget to take your medicine?"*

Specific question: *"About how often has this happened in the past month?"*

V. Data Forms

- A. Precoded if at all possible; greatest number of errors occurs in transcription
- B. Use *one study ID number*
- C. Provide space next to item for response, keep responses right-justified
- D. Provide units, codes and criteria for all possible responses (unknown, not recorded, other {specify _____})
- E. Provide boxes or spaces for as many decimal places as maximal value you could encounter (weight in kg to 3 spaces, blood glucose in 4 spaces, etc.)

- F. Present items in logical order, place sensitive items such as income, sexual habits, etc. last
- G. Place study name and location on each page
- H. Place participant ID on all pages

L. *Pretest all forms* (usually need only a few to find glaring problems)

1. Circulate to colleagues, research assistants, family members
2. Absolutely, positively must be pilot-tested, preferably in participants similar to those you'll be studying
3. Evaluate for missing data, consistency, reasonableness of answers
4. Ask pre-test participants for direct feedback
5. Use duplicate administrations to assess reproducibility

VI. Data Entry

A. General rule: The more times the data are handled, the more errors will be introduced

B. Methods

1. Data from precoded forms may be entered directly into data base if codes listed on right margin
2. Non-precoded data may need to be transferred to code sheets before data entry
3. Duplicate data entry and verification best approach for identifying key punch errors

4. Scorable forms improving in reliability; bar codes for participant ID

C. Computer entry at interview

1. CAPI – Computer Assisted Personal Interview
2. CATI – Computer Assisted Telephone Interview

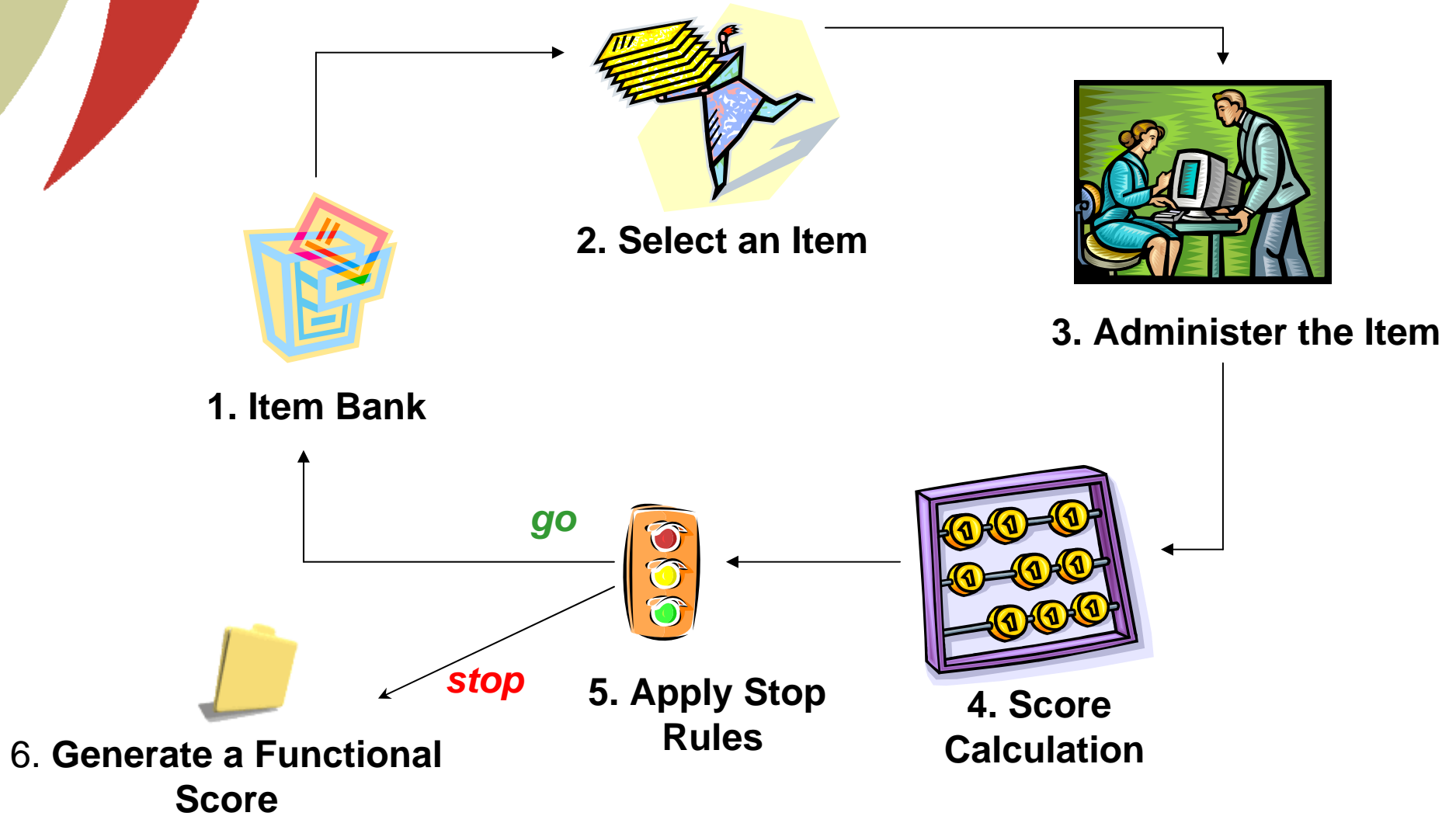
D. Innovative data collection

1. Participant entry into electronic data base – ATM-like machine
2. Audio tape interview with paper and pencil responses
3. Computer Adaptive Testing

Item Response Theory and Adaptive Testing

- Reduces the amount of time required to complete assessments and increases precision by asking only those questions that are relevant to the person's functional level.
- In an adaptive test, as the person responds to items, revised estimates of functional levels are estimated, and then subsequent items for administration are chosen accordingly.
- Computer-adaptive testing (CAT) employs an algorithm that selects items directly tailored to the person, and shortens or lengthens the test to achieve either the desired precision or a preassigned item stopping rule.
- Item response theory (IRT) methods are used to create hierarchically ordered item banks, and then CAT software algorithms select items to match the person's functional level.

CAT process



Source: Alan Jette

**ICC Correlations Between CAT-based and RANDOM-based Scores
with IRT-Criterion Physical Functioning Scores after Administering 5,
10 and 20 items, for three score ranges**

	Low-range <i>N</i> = 131	Mid-range <i>N</i> = 521	High-range <i>N</i> = 150
CAT-5 items	0.87	0.88	0.83
RANDOM-5 items	0.64	0.77	0.53
CAT-10 items	0.92	0.93	0.91
RANDOM-10 items	0.80	0.89	0.76
CAT-20 items	0.97	0.97	0.95
RANDOM-20 items	0.92	0.96	0.93