Survey Development
Measuring self-reported attitudes, perceived environments, behaviors ...

- **Self-report:**
  - Interview
  - **Questionnaire/Survey**
    - Measures typical behaviors
    - Respondent answers directly
    - Close-ended questions
  - Diaries
    - Collects data about events, emotions, beliefs
    - Semi-structured format is common
Questionnaires and Surveys

• Most common measurement tool used in health behavior research

• Why?
  – Low cost
    • Relatively inexpensive to distribute and analyze
  – Often most feasible way to measure a concept that is non-existent in the physical sense:
    • Beliefs, attitudes, perceptions...
• **Questionnaire design is a multi-stage process**
  – Begins with definition of the aspects to be examined
  – Ends with interpretation of the results
  – May be inexpensive to administer well-established survey
  – But expensive in terms of design time and validation
Develop (or select) a questionnaire that...

- is valid.
  - Measures what it is intended to measure.

  - Validity is **specific** to the...
    - ... intended purpose
    - ... intended population
    - ... intended conditions

**Example:** A survey validated to measure frequency of physical activity in suburban dwelling adults.
- Valid for adults living in rural setting?
- Valid for children?
VALIDITY: Does it measure what it intends to measure?

Types of Validity

- **Content Validity**: extent to which a measure represents the concept of interest
- **Criterion Validity**: how well scores (or subset of scores) on instrument predicts an outcome
- **Construct Validity**: how well a scale measures or correlates with similar or different constructs
Determining Content Validity

The extent to which a measure represents the concept of interest.

- **Content Validity Assessment:**
  1. Experts rate each item and the overall scale in its relation to the concept of interest.
  2. Compute Content Validity Index\(^A\) (at least .90) or apply N-1 rule\(^B\)
     - A. Number of items all raters judge to be relevant divided by total number of items on the scale (can also be calculated at the item level)
     - B. For an item to be retained, N-1 raters must have judge it to be relevant

- **Face Validity:** cursory expert review – simple reading of items and judgment
  - Lacks the rigor of content validity assessment – no ratings for individual items nor an associated calculation
Determining Criterion Validity

How well scores (or subset of scores) on instrument predicts an outcome.

- **Criterion Validity Assessment:**
  - Correlation between new instrument scores and an outcome measure.
    - “Gold standard” - comparison between the new instrument/survey and an established instrument
    - SAT scores predicting college success (grades)
Determining Construct Validity

How well a scale measures or correlates with similar or different constructs.

- Construct Validity Assessment is applicable to “concepts that are non-existent in the physical sense” without a criterion to predict or a “gold-standard” comparison.
Reliability

Consistency, repeatability, reproducibility

Reliability Assessment

- **Equivalence**
  - Correlation between parallel tests (large # of items, split items into two forms)

- **Stability (over time)**
  - Test-retest method

**Standards of internal consistency:**
- < .70 considered low
- > .70 considered adequate

Note: the number of items and mean correlation should be considered, too.

- **Internal Consistency**
  - Correlations among items within a survey or scale
  - Coefficient alpha
    1. Correlate each item with every other item
    2. Average item-to-item correlations
    3. Adjust for number of items

- Cronbach’s alpha: based on variance-covariance matrix
Reliability

Consistency, repeatability, reproducibility

• Item Analysis

1. Examine mean scores on each item (ceiling effect?)
2. Inspect item correlations for low or high values
3. Inspect correlation between each item and total scale
4. Should any items be considered for removal?

▶ Intrarater Reliability
  ◦ Consistency among scores assigned by one rater

▶ Interrater Reliability
  ◦ Consistency among scores assigned by two or more raters

kappa: calculation of the percentage of agreement between the two raters, taking chance agreements into account

Standards of kappa:
.4 < .75 considered acceptable
>.81 considered strong
Questionnaires – good enough?

- Self-report measures can only measure what people are willing to say, able to remember, and how well the questions are understood.

Entire field of scientific study dedicated the development of questionnaire instruments and related test theory.
Development of Psychosocial Measures for ACT
(Bone Health Study)

Glanz & Steffen, JADA 2008
Purpose of measurement development:

- to define, formulate, pre-test, and pilot test measures of psychosocial constructs likely to be associated with consumption of high-calcium foods in the young adolescent girls participating in the ACT study
- to be used at pre-test (or soon after) and post-test
Goals of the measures:

- pre-intervention assessment
- pre-post assessment
- possible explanatory or mediating variables
Six Steps to Measures

1. Identifying relevant/related measures to generate a pool of items
2. Selection of groups to participate in pre-testing & pilot testing
3. Pre-testing, de-briefing, revising the instrument based on input (Phase 1 field work)
4. Pilot testing & re-testing revised survey with a larger group (Phase 2 field work)
5. Data processing
6. Statistical analysis to determine…
   * psychometric properties
   * best combination of items for the actual survey
   * statistically sound schemes for composites & data reduction
Identifying Items and Constructs

- W191 materials
- Knowledge tests
- Self-Efficacy Scales (T. Wachs)
- CAL-Girls Study (S. French)

*Item Inventory = 130 items*

*Categories, classification: Iterative*
Pre-Test Version = 61 items

- No exercise items (not main intervention focus)
- No specific self-efficacy scale – but related concepts (availability, taste, image, social influence) [per KW]
Sample

- 7 private schools on three islands (Oahu, the Big Island of Hawaii, and Maui)

- **Phase 1: Pre-test & de-briefing**
  2 schools, n=17 girls

- **Phase 2: Pilot test & re-test [1-2 weeks later]**
  5 schools, n=217 girls
  N=199 full sets of surveys (baseline + follow-up)
  95% completion rate
September 19, 2002

Dear [Principal, Dean...]:

We would like to take this opportunity to introduce you and your 5th, 6th, 7th and 8th grade teachers to our exciting new project:

**SHeBA**

*Strong Healthy Bones in Adolescents*

**Purpose:**
The purpose of the project is to explore, understand and create new ways to increase calcium intake during adolescence. This is a critical period of life for building healthy bones.
Bone Health Survey

Directions: This survey asks about you and your views and habits about nutrition. The information you provide will help us understand what girls your age are doing to stay healthy. All information will be kept private and confidential. No information will be tracked back to you individually. Thank you for your help.

Place a check or "X" in the box that matches your answer. If you have more than one answer, please choose the one answer that fits best.

Section A: Your Family, Home & Friends

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Usually</th>
<th>Everyday</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☒</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>☐</td>
<td>☒</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☒</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☒</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>☐</td>
<td>☐</td>
<td>☒</td>
<td>☐</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☒</td>
</tr>
</tbody>
</table>
### Section A: Your Family, Home & Friends Continued

<table>
<thead>
<tr>
<th>Question</th>
<th>Milk</th>
<th>Soda</th>
<th>Juice</th>
<th>Punch</th>
<th>Water</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>22) What do most of your friends drink at lunch?</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>23) What do your parents give you to drink at dinner/supper most nights?</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
</tbody>
</table>

### Section B: Your Opinions

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neither Agree or Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>24) When it is cold outside I like to drink cocoa.</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>25) Now that I am older, I'd rather drink soda-pop or coffee instead of milk.</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>26) I like the taste of soda.</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>27) I am trying to lose weight, so I don't drink milk.</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>28) Soda-pop and other drinks are easier to take with you than milk.</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>
## Section C: Milk

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neither Agree or Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>44) I drink milk so I can have strong bones now.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45) If I drink milk now, my bones will be strong when I am older.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46) Milk is not healthy.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47) I know it is important for people my age to drink milk.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48) After I drink milk, my stomach hurts.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49) I don't drink milk because it is fattening.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50) Only little kids need milk.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51) I drink milk because it is good for me.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For questions 52-61, check which foods are high in calcium (High=Yes) or not high in calcium (Not High=No). If you do not know, you can check (Not sure).

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>52) Broccoli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53) Chicken</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54) Bananas</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Key Constructs

Social and Environmental Factors (18 scaled items + 2 categorical items)
- Availability (9 items)
- Social Influence (11 items)

Attitudes and Preferences (25 items)
- Health Benefits (6 items)
- Preferences (5 items)
- Taste (8 items)
- Weight Concerns (3 items)
- Tolerance (2 items)
- Convenience (1 item)

Knowledge (10 items)
Statistical Analysis

1) Univariate – frequencies, descriptive stats
2) Correlations of items – w/in scales, sub-scales
3) Composites – internal consistency (alpha) for scales & sub-scales
4) Test-retest reliability – Spearman’s Rho, Kappas
5) Association of sub-scales to categorical variables – ANOVAs in ProcGLM
6) Cluster analysis to examine item groupings
Internal Consistency Findings (Alpha)

- **Attitudes & Preferences**: .79 (25 items)
- **Social & Environmental**: .82 (18 items)
- **Knowledge**: .73 (10 items)
Test-Retest Reliability (Spearman Rho/Kappa)

Attitudes & Preferences  .78
Social & Environmental  .76
Knowledge  .41
Friends/Parents Drink/Serve  .73
Categorical Variables & Sub-Scales

? Parents serve @ dinner -- Friends drink at lunch
*Milk *Water or juice *Soda/Punch *Other

“Parents serve” – associated with Social Influences
[F (2,195) = 21.81, p<.0001]

“Friends drink” not significantly related to Social Influences
Results

55-item questionnaire

Assesses key psychosocial variables related to content & hypothesized mechanisms of effect of ACT intervention

Takes about 10 minutes, is easy for girls of the study age to complete
Cluster Analyses

METHODS
- Done on baseline data
- Proc Varclus ➔ oblique principal components analysis
- Uses iterative algorithm
- Assigns item to cluster 2/ highest squared correlation
- Iterative Process
- Stops when each cluster has one eigenvalue > One

FINDINGS
- 3 cluster solution – 44% of variance
- 4 cluster solution – 52% of variance
Key Constructs

Social and Environmental Factors (18 scaled items + 2 categorical items)
- Availability (9 items)
- Social Influence (11 items)

Attitudes and Preferences (25 items)
- Health Benefits (6 items)
- Preferences (5 items)
- Taste (8 items)
- Weight Concerns (3 items)
- Tolerance (2 items)
- Convenience (1 item)

Knowledge (10 items)
STRENGTHS & LIMITATIONS

STRENGTHS

- Brief, acceptable measures
- Good internal consistency & reproducibility
- If asked after baseline FFQ, less potential for bias
- Items can be analyzed as single items, scales, or re-combined based on cluster analyses

LIMITATIONS

- Limited range of calcium-rich foods queried
- No precedent to predict potential for change in vbls
- Brevity led to compromises
- Measurement study mainly Asian, Caucasian girls
ACKNOWLEDGEMENTS

Carol Boushey
Kathleen Welshimer
Jennifer Chee
Diana Evensen
Simone French
Jared Kuroiwa
Elly Mar Aganon
Tomo Ono

... and participating schools & students