The Role of Question Characteristics in Designing and Evaluating Survey Questions

Jennifer Dykema
University of Wisconsin Survey Center
University of Wisconsin-Madison

Nora Cate Schaeffer
University of Wisconsin Survey Center
Department of Sociology
University of Wisconsin-Madison

Dana Garrabrant
Department of Sociology
Loyola University Chicago

AAPOR Webinar
January 21, 2016

Question writers focus on question characteristics
... question length

Do you feel you have ever been treated unfairly because of your race or ethnicity?

1> YES
2> NO

In general, would you say that your health is excellent, very good, good, fair, or poor?

1> EXCELLENT
2> VERY GOOD
3> GOOD
4> FAIR
5> POOR

During the past week, on about how many days did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage, or liquor?

___ NUMBER OF DAYS (0-30)

Word Count = 15
Word Count = 16
Word Count = 29

Question writers focus on question characteristics
... difficulty level

Do you feel you have ever been treated unfairly because of your race or ethnicity?

1> YES
2> NO

In general, would you say that your health is excellent, very good, good, fair, or poor?

1> EXCELLENT
2> VERY GOOD
3> GOOD
4> FAIR
5> POOR

During the past week, on about how many days did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage, or liquor?

___ NUMBER OF DAYS (0-30)

Grade Level = 8.4
Grade Level = 6.0
Grade Level = 12.2
Question writers focus on question characteristics
… response format

- Do you feel you have ever been treated unfairly because of your race or ethnicity?
  <1> YES  <2> NO
  Yes/No Response Format

- In general, would you say that your health is excellent, very good, good, fair, or poor?
  <1> EXCELLENT  <2> VERY GOOD  <3> GOOD  <4> FAIR  <5> POOR
  Selection Response Format

- During the past week, on about how many days did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage, or liquor?
  _____ NUMBER OF DAYS (0-30)
  Discrete Value Response Format

Where we are now … making recommendations

- Recommendations for writing questions are
  • formulated around question characteristics
  • based on research about impact on outcomes

Where we are now … knowledge has cumulated

- Know a lot about effects of some characteristics on data quality
Where we are now ... still under development

- Developing a comprehensive typology in which
  - characteristics are cataloged and organized
  - effects on respondents' and interviewers' cognitive processing are understood
  - effects on data quality are documented

Outline and objectives

- Overview of approaches
- Taxonomy of features
- Case studies
- Summary and future directions

Notes to help frame our goals for today

- Not presenting a list of guidelines
- Broader goals
- Self-documenting slides
Outline and objectives

Overview of approaches

Taxonomy of features

Case studies

Summary and future directions

General approaches to identify and evaluate question characteristics

<table>
<thead>
<tr>
<th></th>
<th>Experimental</th>
<th>Observational</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scope</strong></td>
<td>Limited</td>
<td>Wide-ranging</td>
</tr>
<tr>
<td><strong>Design</strong></td>
<td>Write alternative</td>
<td>Code questions</td>
</tr>
<tr>
<td></td>
<td>question forms that</td>
<td>along the</td>
</tr>
<tr>
<td></td>
<td>incorporate the</td>
<td>dimensions of all</td>
</tr>
<tr>
<td></td>
<td>characteristic</td>
<td>characteristics</td>
</tr>
<tr>
<td></td>
<td>Other characteristics held constant</td>
<td>Other characteristics vary across questions</td>
</tr>
<tr>
<td><strong>Administration</strong></td>
<td>Rs randomly exposed to a subset of characteristics</td>
<td>Rs exposed to all questions and characteristics</td>
</tr>
<tr>
<td><strong>Evaluation</strong></td>
<td>Assess effect on predetermined criterion</td>
<td>Assess impact on outcome available for all questions</td>
</tr>
</tbody>
</table>

Two kinds of observational approaches

- Ad-hoc
- System-based
Ad-hoc observational approaches: Kinds of characteristics

- Examine individual question characteristics
  - General characteristics
    - Question type
    - Question length
    - Response format
    - Instructions to respondents
  - Specific characteristics
    - Ambiguous terms
    - End-point only labeling
    - Mismatch between question text and response categories

Ad-hoc observational approaches: Outcomes

- Outcomes
  - Response times, item-missing responses, respondent and interviewer behaviors

- Some representative studies

System-based observational approaches

- Code questions using established scheme or system
- Identify problems through different sources
  - Observation in cognitive interviews
  - Speculation based on response process model
- Past analysis of effect of characteristics on data quality
- Representative studies: Forsyth et al. 2004 (CCS); Saris & Gallohofer 2007 (SQP); Willis 2005 (QAS)
Outline and objectives

Overview of approaches

Taxonomy of features

Case studies

Summary and future directions

---

### Taxonomy of features

<table>
<thead>
<tr>
<th>Classes of Characteristics</th>
<th>Examples of individual features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question topic</td>
<td>Health, politics</td>
</tr>
<tr>
<td>Question type</td>
<td>Event or behavior, evaluation or judgment, classification</td>
</tr>
<tr>
<td>Response dimension</td>
<td>Occurrence, frequency, intensity, valence</td>
</tr>
<tr>
<td>Conceptualization and operationalization of the target object</td>
<td>Labels for target object and response dimension</td>
</tr>
<tr>
<td>Question structure</td>
<td>Filter and follow-up question, battery</td>
</tr>
<tr>
<td>Response format or question form</td>
<td>Yes/no, selection, discrete value, field-coded open, record-verbatim open</td>
</tr>
<tr>
<td>Response categories</td>
<td>Type, number, and labeling</td>
</tr>
<tr>
<td>Question wording</td>
<td>Length, readability</td>
</tr>
<tr>
<td>Question implementation</td>
<td>Mode, orientation of scale on screen, instructions to interviewers</td>
</tr>
</tbody>
</table>
Question characteristics and decisions: Two examples

- Decisions for questions about events and behaviors
- Decisions for questions about evaluations and judgments
- Two illustrative example decisions
  - Filter questions
  - Response category labels

Question about events and behaviors – decisions (Schaeffer & Dykema 2011)

Question asking for evaluations and judgments – decisions (Schaeffer & Dykema 2011)
### Decision: Filter questions

<table>
<thead>
<tr>
<th>Events</th>
<th>Evaluations &amp; Judgements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure 1 – With Filter</strong></td>
<td></td>
</tr>
<tr>
<td>During the last twelve months, did you go to a museum?</td>
<td>Are you concerned about pesticides in your food?</td>
</tr>
<tr>
<td><strong>IF YES:</strong> During the last twelve months, how many times did you go to a museum?</td>
<td><strong>IF YES:</strong> How concerned are you about pesticides in your food: a little bit, somewhat, very, or extremely?</td>
</tr>
<tr>
<td><strong>Structure 2: Without filter</strong></td>
<td></td>
</tr>
<tr>
<td>During the last twelve months, how times, if any, did you go to a museum?</td>
<td>How concerned are you about pesticides in your food: not at all, a little bit, somewhat, very, or extremely?</td>
</tr>
</tbody>
</table>

### Comparison: Event question and evaluation question

- Differences
  - Target objects
  - Response dimensions
  - Cognitive processes
- Implications
  - Meaning differs
  - Impact on quality may differ

### Decisions and analytic challenges

- The consequences of item characteristics may vary by question type or response format
- There may be statistical interactions
- Estimating interactions with statistical power requires designing appropriate combinations of question characteristics in sample
### Decision: Category labels

<table>
<thead>
<tr>
<th>Events</th>
<th>Evaluations &amp; Judgements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form: Discrete value</td>
<td>Form: Selection with Ordered Categories</td>
</tr>
</tbody>
</table>

**Version 1 – Unlabeled**

In the last 12 months, how many times, if any, did you visit a museum?

____(NUMBER OF TIMES)

**Version 1 – End-points labeled**

**Version 2: Labeled**

**Decisions and analytic challenges**

- The consequences of item characteristics may depend on question type or response format
- There may be statistical interactions
- Estimating interactions with statistical power requires designing appropriate combinations of question characteristics in sample
- Some question characteristics are relevant only for some question types or response formats
- There are structural dependencies or empty cells

**Implications**

- Consider possible interactions and empty cells when designing analyses
- Consider possible interactions and empty cells when generalizing from or applying results from studies

### Outline and objectives

- Overview of approaches
- Taxonomy of features
- Case studies
- Summary and future directions
Case studies: Overview of design and methods

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Mode</th>
<th>Dates</th>
<th>Question topics</th>
<th>Question types</th>
<th>Outcomes</th>
<th>Unit of analysis</th>
<th>Analytic method</th>
<th>Further details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wisconsin Longitudinal Study (WLS)</td>
<td>1/3 random sample of WI high school class of 1957</td>
<td>Telephone</td>
<td>2003-2005</td>
<td>Health</td>
<td>Events and behaviors</td>
<td>Q-reading accuracy; Any &quot;problem&quot; behaviors</td>
<td>Question-answer sequence (n = 8,150)</td>
<td>Cross-classified random effects logistic models</td>
<td>Dykema et al. 2013</td>
</tr>
</tbody>
</table>

Question characteristics and coding systems

- General question characteristics
  - Response format
  - Question wording
    - Question length
    - Readability
    - Comprehension difficulty (QUAID)

- Coding systems
  - Problem Classification Coding Scheme (CCS)
  - Question Appraisal System (QAS)
  - Survey Quality Predictor (SQP)

General characteristics: Response format

- Format projected by question for responding
  - Yes-No
    - Have you been able to see at all?
  - Selection
    - In general, would you say your health is excellent, very good, good, fair, or poor?

- Predictions
  - Interviewer question-asking accuracy
    - Selection question → reading error more likely
  - Respondent problem behaviors
    - Selection question → problem behavior more likely
General characteristics: Question wording measures

- Question length
  - Count of total number of words

- Question readability
  - Flesch-Kincaid Grade Level score

- Question comprehension difficulty
  - Question Understanding Aid (QUAID) (Graesser et al. 2006)

General characteristics: Question wording predictions

- Interviewer less likely to read question exactly with
  - Longer
  - Harder-to-read
  - Higher comprehension difficulty

- Respondent more likely to display problem behavior with
  - Longer
  - Harder-to-read
  - Higher comprehension difficulty

Overview of coding systems

- Problem Classification Coding Scheme (CCS) (Forsyth et al. 2004)

- Question Appraisal System (QAS) (Willis 2005)

- Survey Quality Predictor (SQP) (Saris & Gallhofer 2007)
Predictions for coding schemes

- Interviewer less likely to read question exactly with
  - Higher CCS scores
  - Higher QAS scores
  - Lower SQP scores

- Respondent more likely to display problem behavior with
  - Higher CCS scores
  - Higher QAS scores
  - Lower SQP scores

Values for measures for two questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Length</th>
<th>Readability</th>
<th>Comprehension Difficulty</th>
<th>CCS Problems</th>
<th>QAS Problems</th>
<th>SQP Data Quality Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Long</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Q2</td>
<td>Short</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Q1. Because of any impairment or health problem, do you need the help of other persons in handling your routine needs, such as everyday household chores, doing necessary business, shopping or getting around for other purposes?

Q2. Have you ever been diagnosed with a mental illness?

Results from bivariate cross-classified random effects logistic models of interviewer and respondent outcomes on question characteristics

<table>
<thead>
<tr>
<th>Question Characteristics &amp; Systems</th>
<th>Interviewer Exact Reading (Odds)</th>
<th>Respondent Problems (Any) (Odds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response format [Yes/No] Selection</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Question length</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Question readability</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Comprehension difficulty (QUAID)</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>Problem Classification Coding System (CCS)</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Question Appraisal System (QAS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey Quality Predictor (SQP)</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001
Case studies: Overview of design and methods

<table>
<thead>
<tr>
<th>Sample</th>
<th>Wisconsin Longitudinal Study (WLS)</th>
<th>General Social Survey (GSS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>Telephone</td>
<td>Face-to-face</td>
</tr>
<tr>
<td>Question topics</td>
<td>Health</td>
<td>Varied, public opinion</td>
</tr>
<tr>
<td>Question types</td>
<td>Events and behaviors</td>
<td>Varied, 2/3 are evaluations</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Q-reading accuracy; Any “problem” behaviors</td>
<td>Individual question reliabilities</td>
</tr>
<tr>
<td>Unit of analysis</td>
<td>Question-answer sequence (n = 8,150)</td>
<td>Questions (n = 196)</td>
</tr>
<tr>
<td>Analytic method</td>
<td>Cross-classified random effects logistic models</td>
<td>Regression</td>
</tr>
<tr>
<td>Further details</td>
<td>Dykema et al. 2013</td>
<td>Schaeffer et al. 2015</td>
</tr>
</tbody>
</table>

Histogram of distribution of reliability, n=196 GSS core items, 2006-2010

General Social Survey – an item with high reliability

IF CURRENTLY MARRIED OR WIDOWED:
A. Have you ever been divorced or legally separated?
   Yes 1
   No 2
**General Social Survey – an item with low reliability**

Do people in these groups tend to be unintelligent or intelligent?

<table>
<thead>
<tr>
<th>Hand Card A15</th>
<th>Unintelligent</th>
<th>Intelligent</th>
<th>Don't Know</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7</td>
<td></td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

A. Where would you rate whites in general on this scale?  
B. Blacks?

---

**Case Study: General Social Survey**

- Operationalizing complexity of question wording
- Decisions about batteries

---

**Battery**

- Two or more items
  - Presented together
  - With the same response categories
  - The first target question is usually preceded by an introduction.
**General Social Survey: The nat* battery**

1. First I would like to talk with you about some things people think about today. We are faced with many problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A)... are we spending too much, too little or about the right amount on (ITEM)? READ EACH ITEM; CODE ONE FOR EACH.

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Too much</th>
<th>Too little</th>
<th>About right</th>
<th>DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. The space exploration program</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>B. Improving and protecting the environment</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>C. Improving and protecting the nation’s health</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

**General Social Survey: The nat* battery – item A**

1. First I would like to talk with you about some things people think about today. We are faced with many problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A)... are we spending too much, too little or about the right amount on (ITEM)? READ EACH ITEM; CODE ONE FOR EACH.

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Too much</th>
<th>Too little</th>
<th>About right</th>
<th>DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. The space exploration program</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>B. Improving and protecting the environment</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>C. Improving and protecting the nation’s health</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

**General Social Survey: The nat* battery – item B**

1. First I would like to talk with you about some things people think about today. We are faced with many problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A)... are we spending too much, too little or about the right amount on (ITEM)? READ EACH ITEM; CODE ONE FOR EACH.

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Too much</th>
<th>Too little</th>
<th>About right</th>
<th>DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. The space exploration program</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>B. Improving and protecting the environment</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>C. Improving and protecting the nation’s health</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>
First I would like to talk with you about some things people think about today. We are faced with many problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A)… are we spending too much, too little or about the right amount on (ITEM)? READ EACH ITEM; CODE ONE FOR EACH.

### A. The space exploration program
- Too much: 3
- Too little: 1
- About right: 2
- DON'T KNOW: 8

### B. Improving and protecting the environment
- Too much: 3
- Too little: 1
- About right: 2
- DON'T KNOW: 8

### C. Improving and protecting the nation's health
- Too much: 3
- Too little: 1
- About right: 2
- DON'T KNOW: 8

**Question wording: Measures of the complexity of language**

- Number of words
  - All words needed to answer the questions
    - In batteries, we included the preamble in calculating the word count for second and later items
    - Words actually read to the respondent – according to the script
    - Difference between “all words” and “read words”
  - QUAID measures (Graesser et al. 2006)@
    - Number of left-embedded words - words before main verb
    - Sum of number of problem words – technical or vague – and other problems
  - Reach measures

**General Social Survey: The nat* battery – item B**

1. First I would like to talk with you about some things people think about today. We are faced with many problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A)… are we spending too much, too little or about the right amount on (ITEM)? READ EACH ITEM; CODE ONE FOR EACH.

### A. The space exploration program
- Too much: 3
- Too little: 1
- About right: 2
- DON'T KNOW: 8

### B. Improving and protecting the environment
- Too much: 3
- Too little: 1
- About right: 2
- DON'T KNOW: 8

### C. Improving and protecting the nation's health
- Too much: 3
- Too little: 1
- About right: 2
- DON'T KNOW: 8
General Social Survey: The nat* battery – item B

1. First I would like to talk with you about some things people think about today. We are faced with many problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A)… are we spending too much, too little or about the right amount on (ITEM)?

READ EACH ITEM; CODE ONE FOR EACH.

<table>
<thead>
<tr>
<th></th>
<th>Too much</th>
<th>Too little</th>
<th>About right</th>
<th>DON’T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. The space exploration program</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>B. Improving and protecting the environment</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>C. Improving and protecting the nation’s health</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

Question wording: Measures of the complexity of language

- Number of words
  - All words needed to answer the questions
  - In batteries, we included the preamble in calculating the word count for second and later items
  - Words actually read to the respondent – according to the script
  - Difference between “all words” and “read words”
- QUAI measures (Graesser et al. 2006@)
  - Number of left-embedded words - words before main verb
  - Sum of number of problem words – technical or vague – and other problems
  - Flesch measures

Reliability by the number of words read to the respondent, items not in a battery, N=115
Reliability by the number of words read to the respondent, items first in a battery, \( N = 15 \)

Reliability by the number of words read to the respondent, items second or later in a battery, \( N = 66 \)

Regression of item reliability on measures of question wording and position in battery, \( N = 196 \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left-embedded words (QUAID)</td>
<td>***</td>
</tr>
<tr>
<td>Problem words &amp; language problems (QUAID)</td>
<td></td>
</tr>
<tr>
<td>Number of words read</td>
<td></td>
</tr>
<tr>
<td>First item [Not in battery omitted]</td>
<td></td>
</tr>
<tr>
<td>First item*Number of words</td>
<td></td>
</tr>
<tr>
<td>Second or later item</td>
<td></td>
</tr>
<tr>
<td>Second or later item*Number of words</td>
<td></td>
</tr>
</tbody>
</table>

\( *=p<0.10, **=p<0.05, ***=p<0.001 \)
Regression of item reliability on measures of question wording and position in battery, N=196

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left-embedded words (QUAID)</td>
<td>***</td>
</tr>
<tr>
<td>Problem words &amp; language problems (QUAID)</td>
<td>*</td>
</tr>
<tr>
<td>Number of words read</td>
<td>***</td>
</tr>
<tr>
<td>First item [Not in battery omitted]</td>
<td>***</td>
</tr>
<tr>
<td>First item*Number of words</td>
<td>*</td>
</tr>
<tr>
<td>Second or later item</td>
<td>***</td>
</tr>
<tr>
<td>Second or later item*Number of words</td>
<td>***</td>
</tr>
</tbody>
</table>

*=p<0.10, **=p<0.05, ***=p<0.001

Reliability estimates by the number of words read to the respondent, by placement in a battery

<table>
<thead>
<tr>
<th>N=15</th>
<th>N=15</th>
<th>N=66</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not in a Battery</td>
<td>First in Battery</td>
<td>Second or later</td>
</tr>
</tbody>
</table>

Question type – Type of target object + response dimension

- Asking for an evaluation or judgment (“Evaluation questions”) – N=150
- Asking about events or behaviors (“Event questions”) – N=27
- Event-based classifications and calculations (“Event-based classifications”), N=19
Regression of reliability on language measures, position in battery, and question type, N = 196

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question wording</td>
<td></td>
</tr>
<tr>
<td>Left-embedded words (QUAID)</td>
<td>+***</td>
</tr>
<tr>
<td>Problem words &amp; language problems (QUAID)</td>
<td>+</td>
</tr>
<tr>
<td>Number of words read</td>
<td>+***</td>
</tr>
<tr>
<td>First item [Not in battery omitted]</td>
<td>-</td>
</tr>
<tr>
<td>First item*Number of words</td>
<td>+</td>
</tr>
<tr>
<td>Second or later item</td>
<td>+</td>
</tr>
<tr>
<td>Second or later item*Number of words</td>
<td>+**</td>
</tr>
<tr>
<td>Question type [Evaluation omitted]</td>
<td></td>
</tr>
<tr>
<td>Event question</td>
<td>+***</td>
</tr>
<tr>
<td>Event-based classification</td>
<td>+***</td>
</tr>
</tbody>
</table>

*=p<0.10, **=p<0.05, ***=p<0.001

Case study: Comments on the GSS

- Results depend on the sample of items and distribution of item characteristics in the GSS
- Some evaluation questions in the GSS have very good reliabilities
- Evaluation and event questions are not interchangeable in measurement
- Topic is confounded with other characteristics

Outline and objectives

- Overview of approaches
- Taxonomy of features
- Case studies
- Summary and future directions
Summary

• Comments on the case studies
• New coding system
• Importance of a criterion
• Future directions

Thank You!

For questions about this webinar or more information, contact:

Jennifer Dykema
dykema@ssc.wisc.edu
Nora Cate Schaeffer
schaeffe@ssc.wisc.edu

Please visit us at:
www.uwsc.wisc.edu

References

References (continued)


