

# Evaluating Survey Questions: An Inventory of Methods

Report of the FCSM Subcommittee on  
Question Evaluation Methodology



*Diane K. Willimack*  
*U.S. Census Bureau*

# Answers to ----

- Why evaluate survey questions?
- What do we mean by question ‘evaluation’?
- Why an inventory of evaluation methods?
- What to do with the Methods Inventory?
- How can you use the Methods Inventory?



# Why evaluate survey questions?

- Do they ask what the researcher intended?
- Are they consistently understood as intended?
- Are respondents willing and able to consistently answer the questions as intended?
- Is it overly burdensome for respondents to provide a full response?

Groves et al., 2009

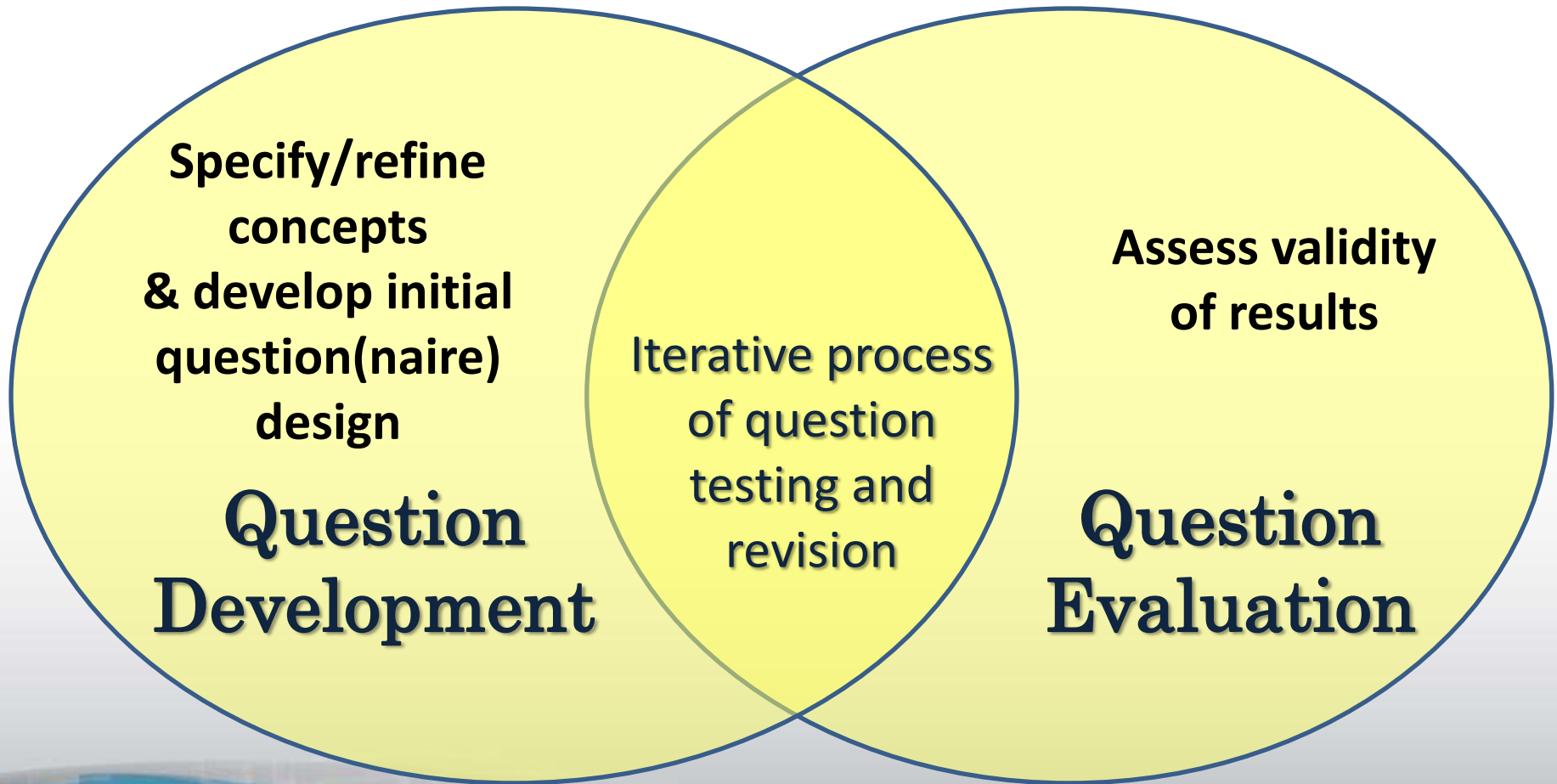
# What do we mean by question 'evaluation'?



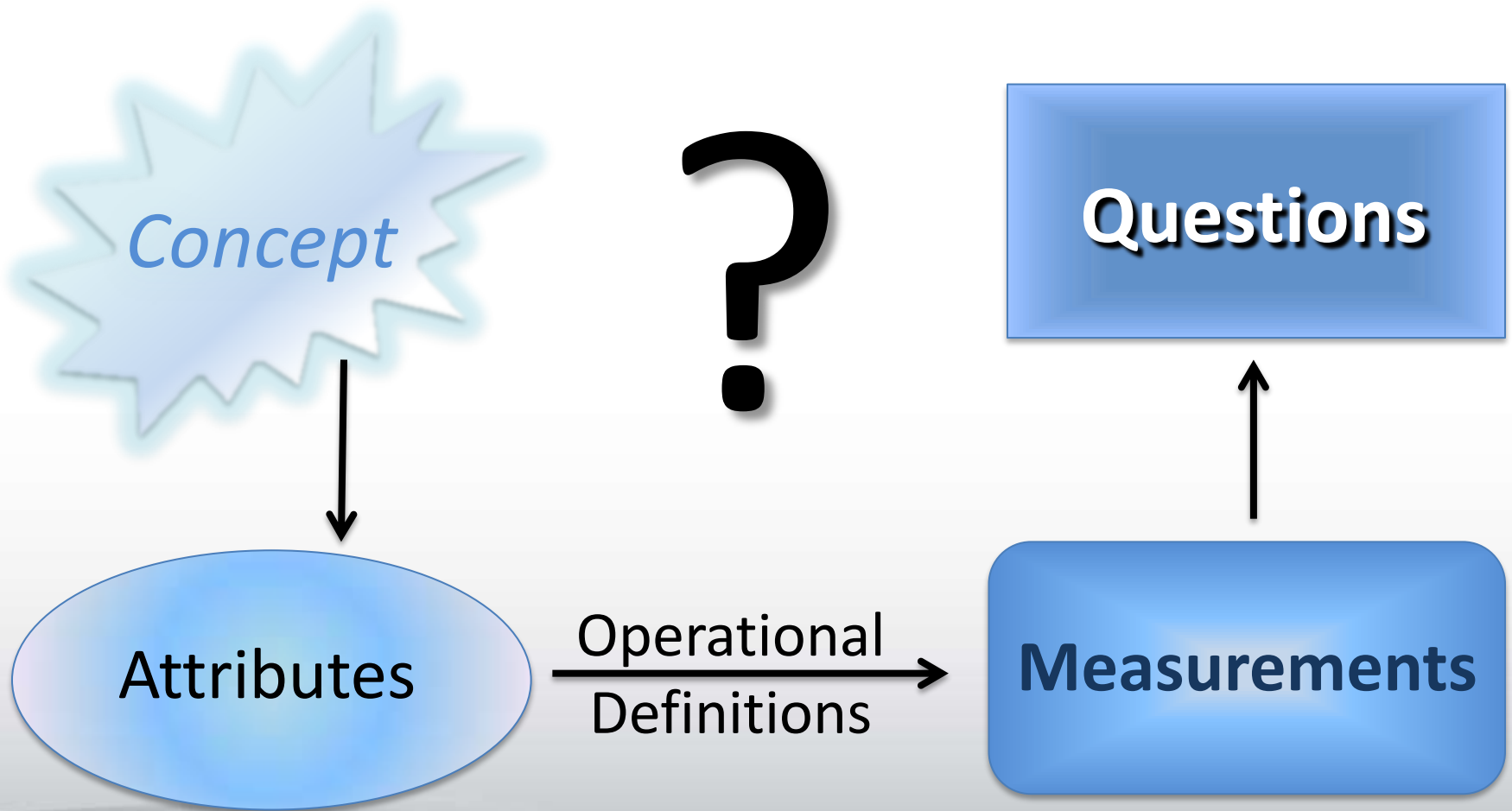
**Question  
Development**

**Question  
Evaluation**

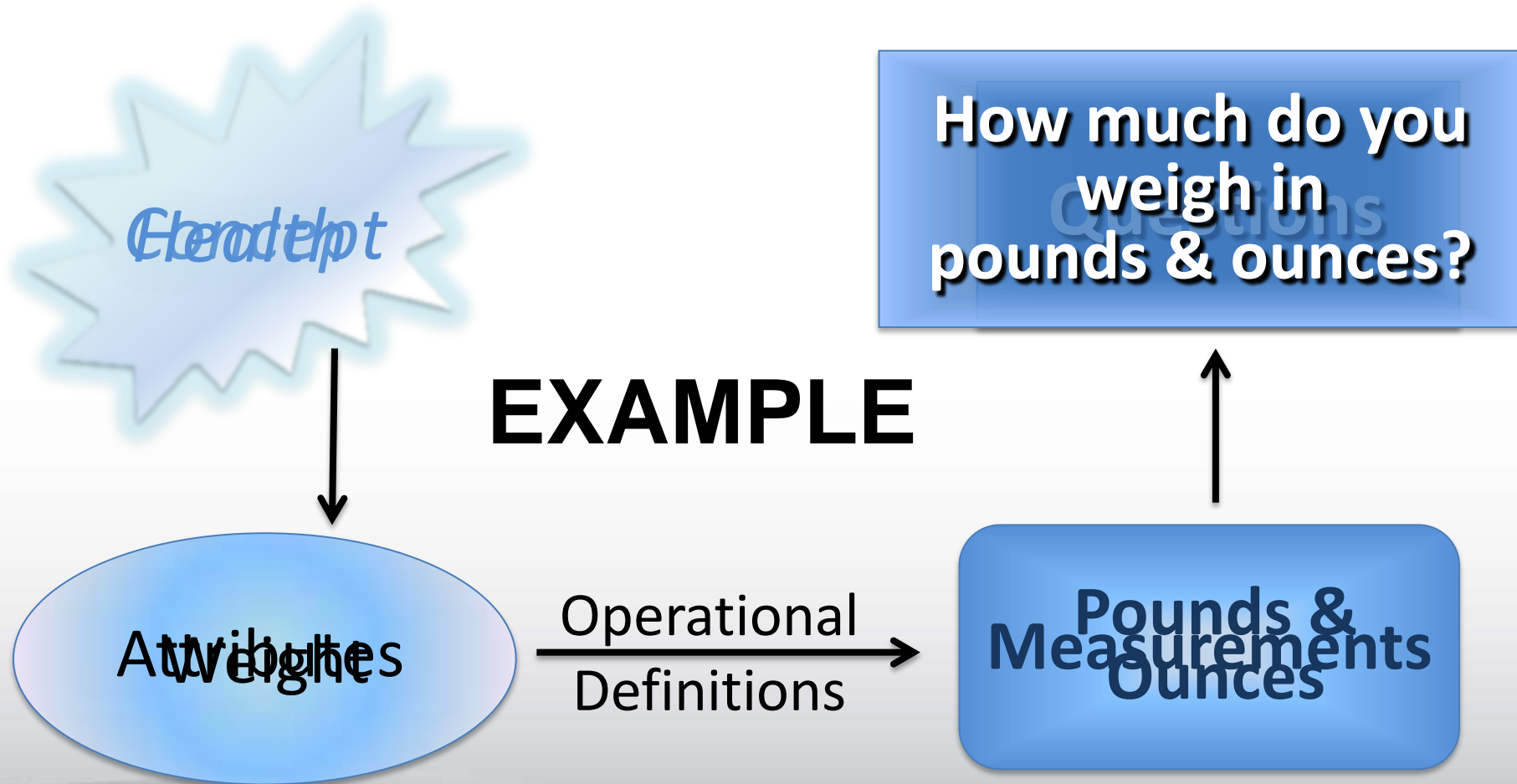
# What do we mean by question 'evaluation'?



# Question Development



# Question Development



# Question Development Methods

- Actively use respondents
  - Ethnography
  - Early Stage Scoping
  - Focus Groups
  - Concept Mapping
  - User Needs Assessment

- Strengths
- Limitations

- Passively use respondents
  - Dimension/Attribute Analysis
  - Mining of Questionnaires

- Strengths
- Limitations



# Question Evaluation Purpose

*Prior to field test  
or production  
data collection*

- **Will respondents' answers meet the intent of the question?**

*During or following  
field test or production  
data collection*

- **Do/did respondents' answers meet the intent of the question?**

# 'Field' Tests

- Differences in sample design
- Terminology
  - Field pilot vs. large-scale
  - Field test vs. large-scale
  - Dress rehearsal
  - ...other
- Experience
- Positive
- Probability



# Why an **inventory** of evaluation methods?

- No single method satisfies all research needs
  - Hammers, screwdrivers, wrenches, and computers
- All methods have strengths and limitations
- Many methods can be used for different purposes
  - One method → multiple applications
- Using more than one method is advised
  - Strengths of one method offset limitations of another
  - Complementary findings provide fuller picture
  - Corroboration strengthens validity of results



# The Inventory of Methods

- **14 methods**
  - Definition
  - Typical uses with examples
  - Strengths
  - Limitations
  - References

# The Methods

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# The Methods

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Strengths

- Respondent-centered
  - In-depth, descriptive information
  - Rich, contextual insight
- Conducive to iteration
- Less resource-intensive

## Limitations

- Lack statistical inference
  - Not representative of target population
  - Results not generalizable
- Not “real-life”

# The Methods

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# The Methods

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

## Strengths

- Representative sample
- Statistical inferences and comparisons
- Diagnostic
- Reflect field conditions

## Limitations

- Resource intensive
- Unable to identify –
  - Nature of problems
  - Possible solutions



# Other Features

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# Other Features

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# Other Features

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- **Vignettes & Fictional Scenarios**
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# Other Features

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# Other Features

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# Other Features

## Qualitative

- Methodological Expert Review
- Cognitive Interviews
- Vignettes & Fictional Scenarios
- Usability Testing
- Feedback from Survey Personnel
- Respondent Debriefing

## Quantitative

- Response Analysis Survey
- Randomized Experiments
- Validation Studies
- Analysis of Paradata
- Re-interview / Content Evaluation
- Response Quality Indicators
- Item Response Theory
- Latent Class Analysis

# Multiple Methods: An Example

## 2012 Census of Agriculture

*National Agricultural Statistics Service*

1. Expert Review
  - Methodological
  - Subject matter
2. Response Quality Indicators from 2007 COA
3. Cognitive Testing
4. Multiple Field Tests that featured ---
  - Randomized Experiments
  - Response Quality Indicators
  - Respondent Debriefings

# What to do with the Methods Inventory?

- What the inventory is not ---
  - Not a cookbook
  - Not an exhaustive list
  - Not a list of methods that must be done
  - Not a checklist
  - Not a “how-to” manual for each method



# What to do with the Methods Inventory?

- What the inventory is ---
  - A compendium
  - A toolkit
  - An informational resource
  - A starting point

# What to do with the Methods Inventory?

- Users of the inventory include ---
  - OMB desk officers
  - Statistical agencies
  - Other government agencies
  - Survey research contractors
  - Researchers
  - Data users

# What to do with the Methods Inventory?

- Introduce a variety of methods for evaluating your survey questions
- Provide a common language for survey researchers, practitioners, and sponsors
- Engage in dialogue with survey designers, e.g. contractors, regarding question evaluation
- Ensure due diligence for including question evaluation strategies in survey proposals/plans
- Refer to cited literature to inform and guide proper implementation

# How can **you** use the Methods Inventory?

- Identify goals of your question evaluation studies
- Consider alternative evaluation method(s)
- Match evaluation method(s) to your goals
- Consider costs and timing
- Consider the quality requirements of the data collected using your survey questions
- Select method(s) that satisfy your goals, while recognizing trade-offs due to constraints

# A **Word** of Cautionary Advice

- Document
- DOCUMENT
- **DOCUMENT!!!**

# Conclusion

- A variety of methods are available for evaluating survey questions
  - Tried & true, well documented in scientific literature
  - Different methods → different purposes
  - Multiple methods → complementary
- Why evaluate survey questions?
  - Satisfy OMB Statistical Quality Standards
  - **Contribute to data quality for official statistics**