

Thematic Analysis of Cognitive Interview Data

A qualitative Framework for Evaluating Large-Scale Federal Survey Questions

Halelujha Ketema, Survey Statistician

Social, Economic, and Housing Statistics Division

U.S. Census Bureau

AAPOR | May 15, 2026

Disclaimer

This presentation is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed are those of the author and not those of the U.S. Census Bureau.

The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7534374, Disclosure Review Board (DRB) approval number: CBDRB-FY26-0182.

Current Population Annual Social and Economic Supplement Survey (CPS ASEC)

- **ASEC:** Health insurance, income, and poverty since 1947.
- Official source for national health insurance coverage rate.

Data Collection Mode Timeline



Modernization Timeline

- **FY2024** ← **this study**
Exploratory interviews + Round 1.
- **FY2025**
Round 2 cognitive interviews.
- **FY2026**
Usability testing (current year).
- **FY2027**
Field test.
- **FY2028**
Production phase-in.

Study Objectives

- Identify where respondents struggle with health insurance questions.
- Transform observer notes and other supplemental resources into analyzable qualitative data.
- Build a standardized, repeatable qualitative framework.
- Provide support to subject-matter experts (SMEs) to analyze content-area questions systematically using supplemental qualitative data sources.

Study design and data source

- **RTI:** Cognitive interviews.
- Capture details via **observer notes** especially useful to content area subject matter experts (SMEs).
- Support question-by-question review to supplement broader summaries.
- Help identify recurring issues in areas targeted for modernization and ISR transition.
- Provide additional evidence to supplement broader interview findings.

Observer notes analysis focus

Verbal Information	Direct respondent statements noted by observers.
	How respondents described their coverage or interpreted a question (probing).
	How respondents rephrased questions in their own words.
Behavioral Information	Hesitation and pauses.
	Signals of confusion or frustration.
	Requests for clarification.
	Answer changes after probing.
	Question-specific concerns noted by SMEs.

Why Health Insurance?

- **Coverage types overlap.**
 - Private, employer, union, Medicaid, Medicare, CHIP, marketplace.
- **Policyholder relationships vary.**
 - Young adults on parental plans, split households, roommates.
- **Recall can be challenging.**
 - Coverage changes mid-year.

Qualitative Framework

CDC qualitative analysis spreadsheet:

- Free, accessible, and reproducible alternative to proprietary qualitative software.
- Structured spreadsheet framework for excerpt entry, code assignment, and theme development.
- Supports an auditable, transparent coding process for observer notes and interview narratives.



QAS-99 FRAMEWORK ADAPTED

QAS-99 STEP	Check	Adapted Codebook Category
1. Reading	Is the question easy to read aloud?	Structure & functionality (S)
2. Instructions	Are directions clear to the respondent?	Structure & functionality (S)
3. Clarity	Is the wording ambiguous or vague?	Comprehension (C)
4. Assumptions	Does the question assume too much?	Interpretation (I)
5. Knowledge/memory	Can respondents reasonably recall the answer?	Recall and calculation (R)
6. Sensitivity	Could answers be socially biased?	<i>Not a primary focus</i>
7. Response categories	Do the answer options work for respondents?	Response options (O)
8. Other	Anything not captured by steps 1–7	Structure & functionality (S)

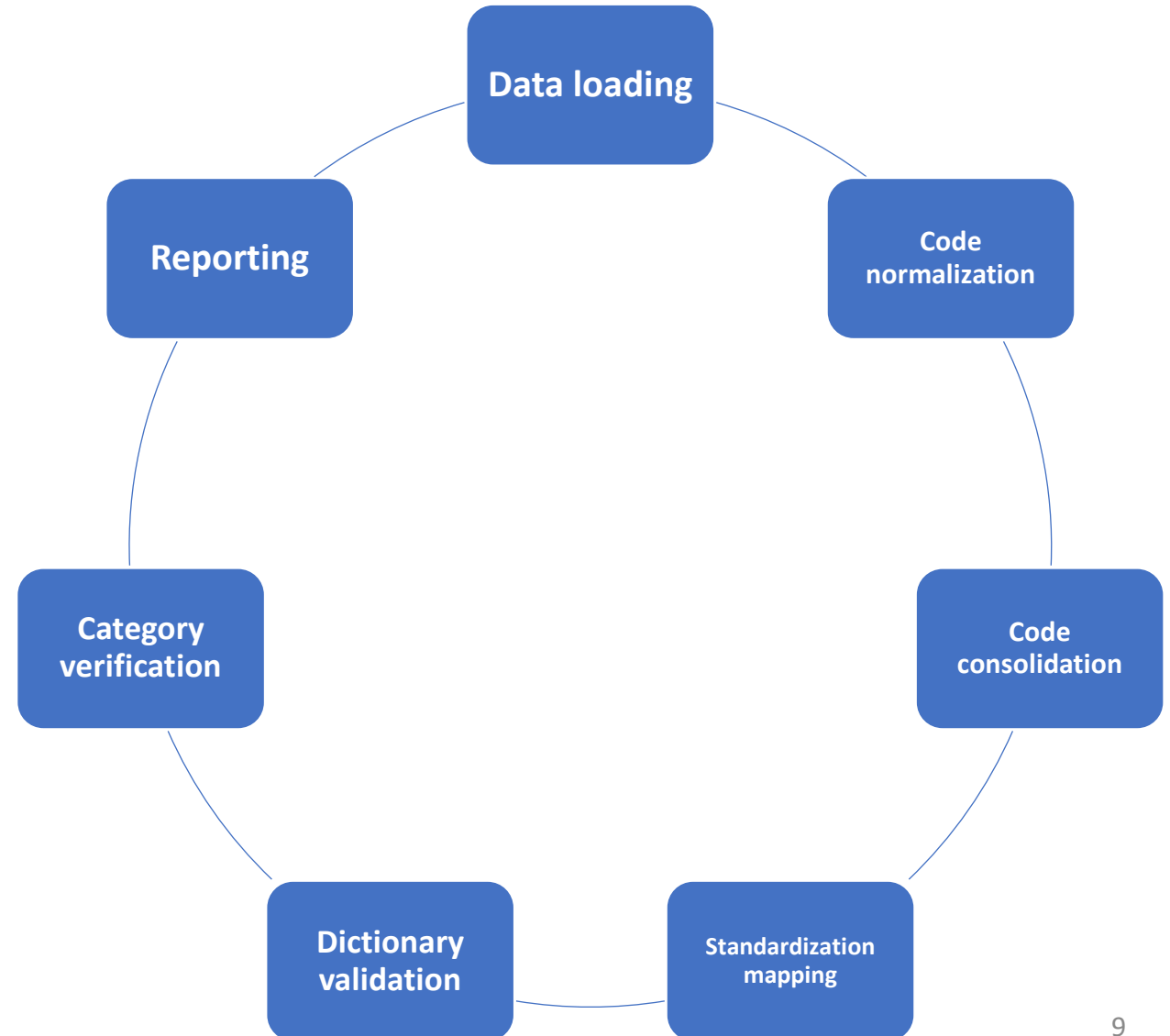
Validation and automation: Python scripting

Automates cleanup of hand-entered codes for consistency.

Renames and consolidates codes into a standardized framework.

Validates all codes against an approved data dictionary.

Logs every change for transparency and replicability.



Frequency codes

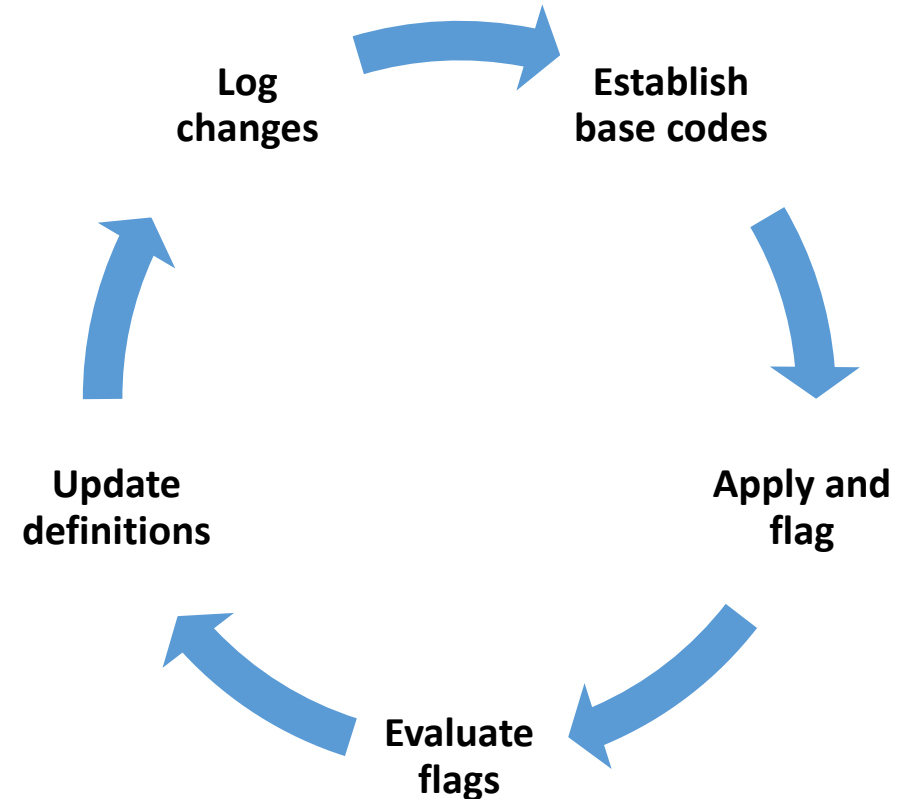
- **Recurring themes and patterns:** Identifies recurring themes and patterns emerging from qualitative data.
- **Focus on data points:** Rather than indicating the number of participants, analysis captures recurring themes within data points for each question.
- **Coding methods for cognitive review analysis:**
 - **Descriptive coding:** Captures key quotes and direct expressions from participants.

Codebook development Process

First cycle: Open coding captured respondent behaviors, comprehension challenges, interpretation patterns, navigation issues, and recall constraints.

- Code set expanded from initial 28 codes to a peak of 90 subcodes before consolidating. **Total of 4 rounds of coding.**
 - Final count of five broad codes and 47 subcodes for health insurance.
- Subsequent cycles iteratively broadened the taxonomy with domain-specific subcodes.

Final codebook: Standardized codes across five categories.



Findings

Finalized coding and analysis of observer notes:

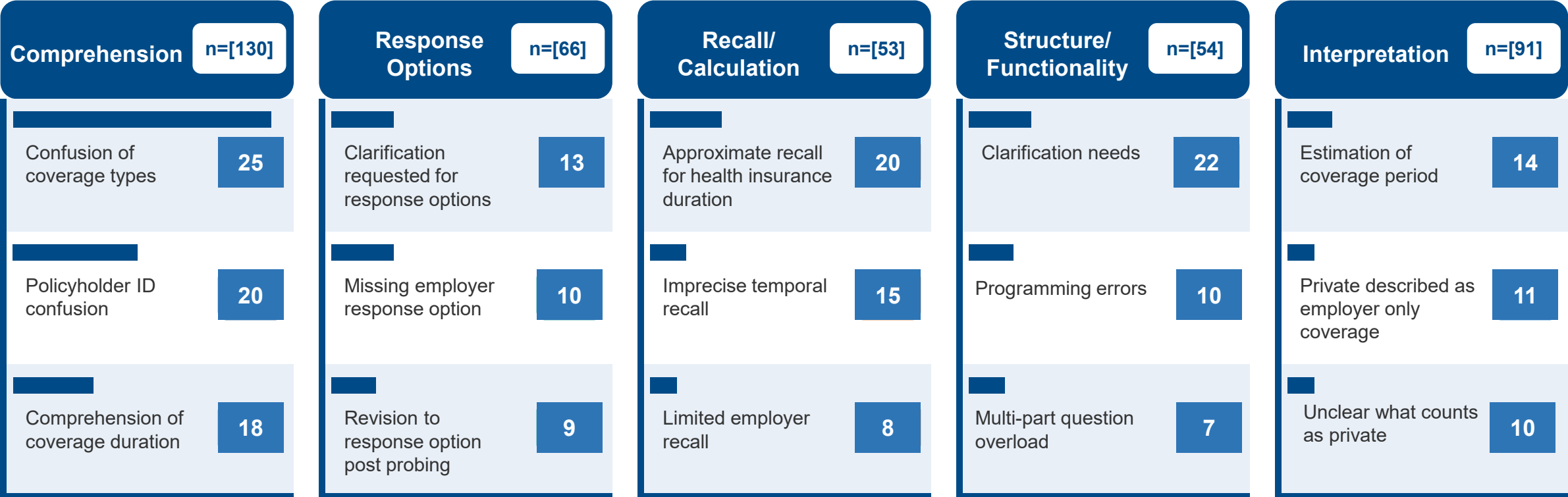
- 18 participants total.
- 60-minute interviews conducted by trained RTI interviewers.
- Semi-structured interviews with probing.
- SMEs captured content-specific observation notes for this research.

Note: This analysis is based on qualitative observer notes and supplemental interview data collected from a limited, nonrandom set of observed interviews. Because the data were not designed to be statistically representative or suitable for inferential testing, counts shown are descriptive only and should not be interpreted as population estimates. Findings identify recurring question-level issues and potential areas for questionnaire refinement but cannot be generalized to all CPS ASEC respondents.

Broad codes and their definitions

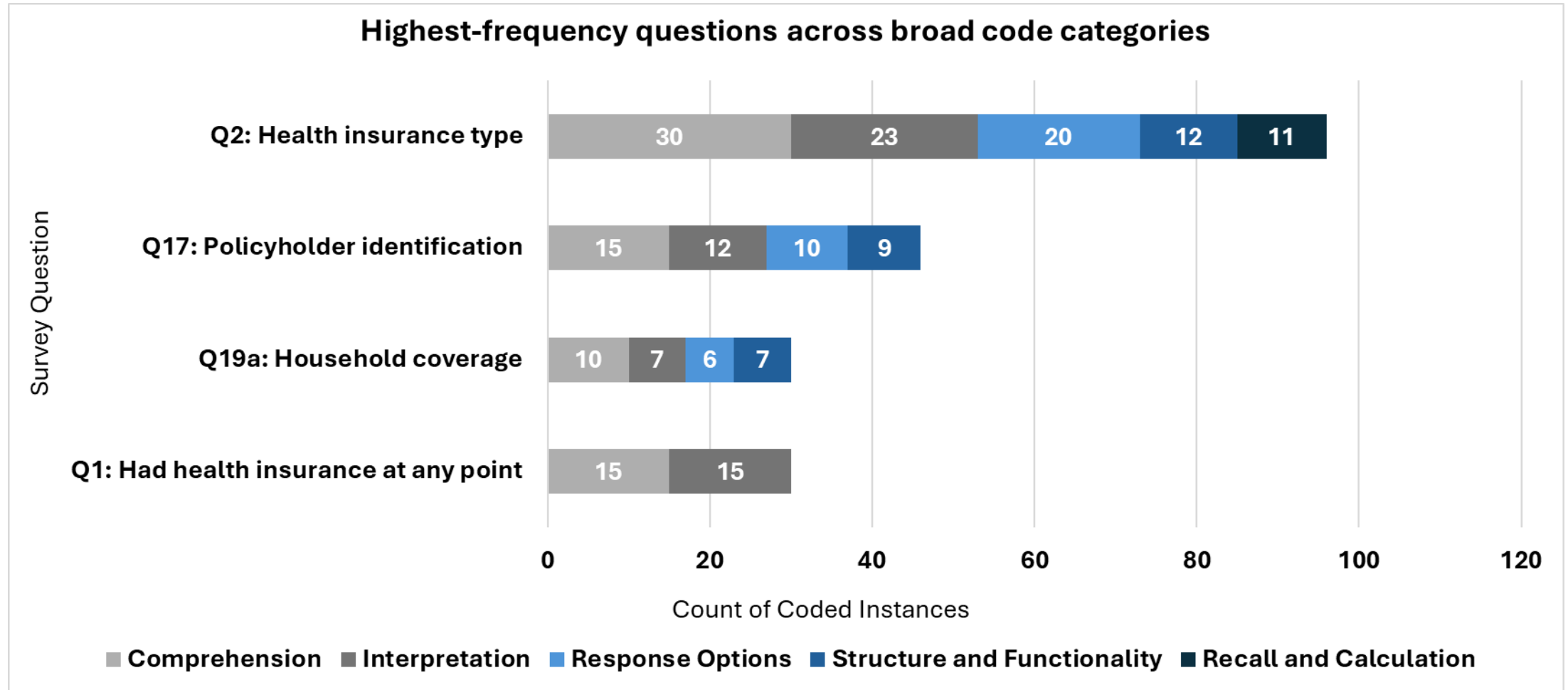
Category	Focus
Comprehension	Terminological and definitional confusion (e.g., coverage types, policyholder identity).
Interpretation	How respondents map their situation onto survey concepts.
Response Options	Adequacy and fit of available response choices.
Structure and Functionality	Survey flow, programming issues, formatting, and repetition.
Recall and Calculation	Memory limitations, imprecise temporal recall, and knowledge gaps.

Count by broad code and subcode



Source: U.S. Census Bureau, CPS ASEC Cognitive Interview Study, Round 1 (2024), conducted by RTI International. Based on a non-probability sample; findings do not support statistical inference and should not be generalized to the U.S. population. Mini bars show relative frequency within each category. Numbers indicate instance count. Bars are not comparable across categories. Top 3 subcodes are shown per category; based on 18 participants and 394 total coded instances (unweighted).

Questions by Broad Code

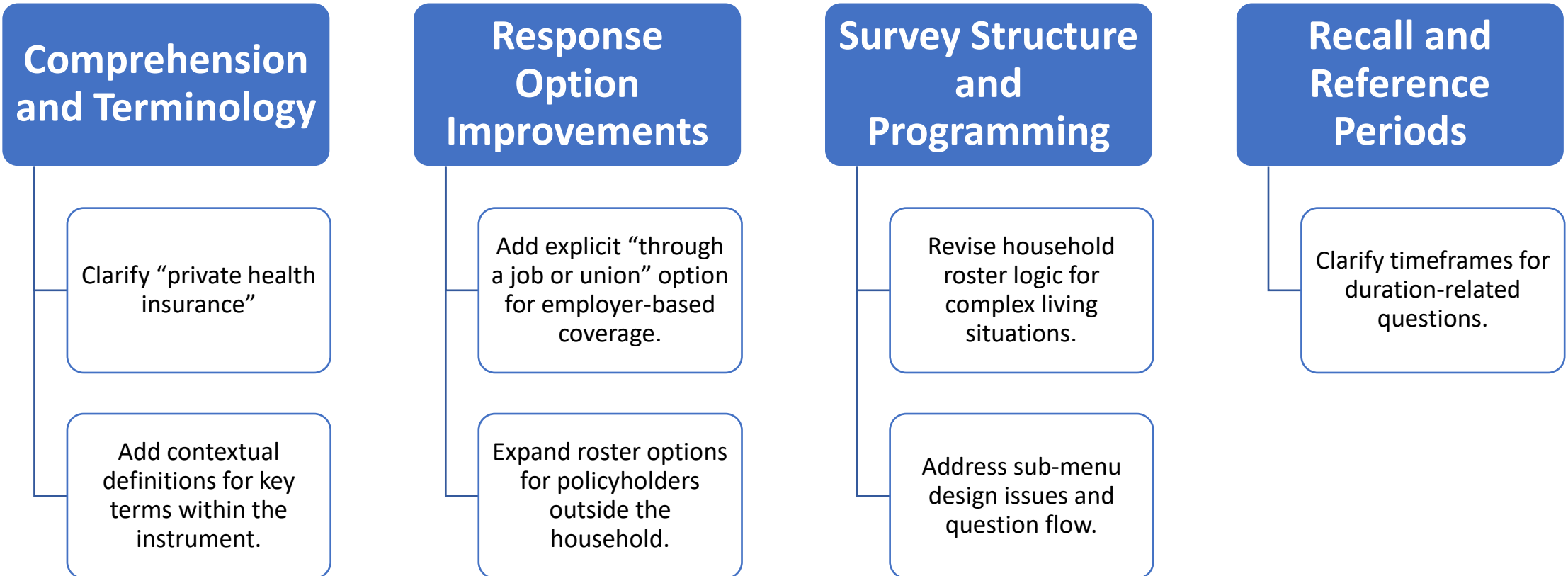


The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7534374, Disclosure Review Board (DRB) approval number: CBDRB-FY26-0182.

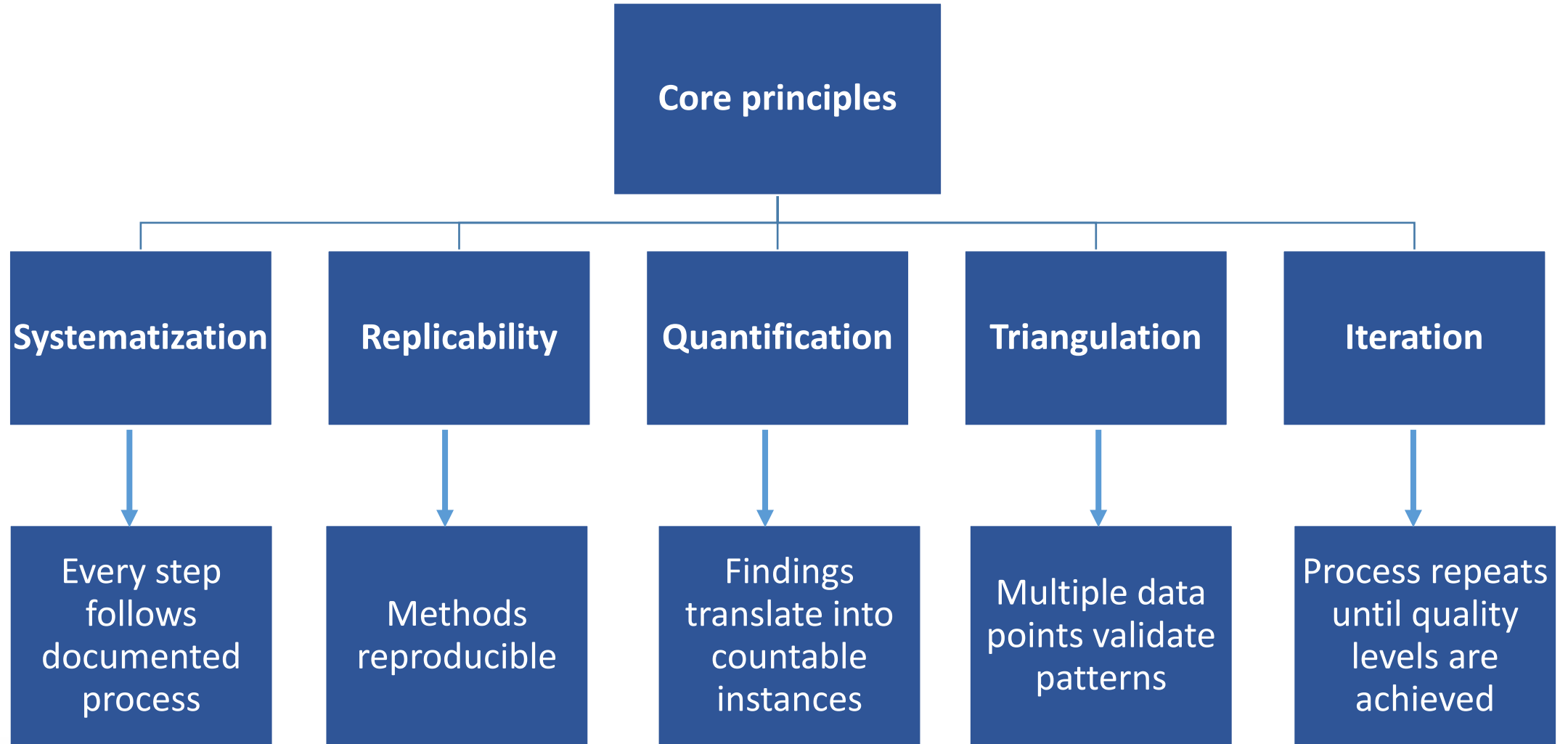
Source: U.S. Census Bureau, CPS ASEC Cognitive Interview Study, Round 1 (2024), conducted by RTI International. Based on a non-probability sample; findings do not support statistical inference and should not be generalized to the U.S. population.

Note: Numbers indicate instance count. Bars are not comparable across categories. Based on 18 participants and 394 total coded instances (unweighted).

Discussion and recommendations



A replicable framework



Cross-survey applications

Survey of Income and Program Participation (SIPP):

- Temporal recall over annual reference periods.
- Income source classification across multiple forms.

American Community Survey (ACS):

- Relationship and marital status measurement.

ASEC Modernization resources and contact INFO

- ASEC modernization website: www.census.gov/programs-surveys/cps/about/modernization/asecmodernization.html.
- Sign up for our [GovDelivery List](#).
- Email us at demo.asec.modernization@census.gov.

Thank you

HALELUJHA.Y.KETEMA@CENSUS.GOV