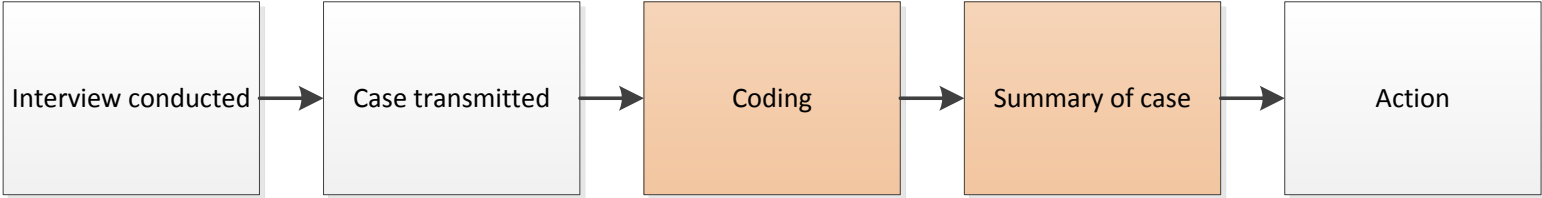# Creating a More User Friendly CARI System

Aaron Maitland
Laura Branden
Susan Genoversa
Naoko Okuma

# Computer Audio Recorded Interviewing (CARI) for Monitoring Field Interviewer Behavior

- CARI offers an opportunity to listen to audio recordings of survey interviews conducted in the field.

- Behavior coding allows for systematic analysis of interviewer behavior across interviews.

- Three primary goals of CARI:
  - Question assessment
  - Interviewer performance
  - Validation

**Westat®**

# CARI process

```
Interview conducted → Case transmitted → Coding → Summary of case → Action
```

**Westat**®

# Research Questions

- How well does a quantitative ranking of interviewers based on coding from the CARI system distinguish between good and poor interviewers?

- Do CARI results differ depending on whether non-project specific or project specific staff conduct the coding?

# National Health and Aging Trends Study (NHATS)

- National sample of Medicare beneficiaries (n~8000)

- Longitudinal study with annual interviews

- Age 65 and older

- Study funded by National Institute on Aging

- Cooperative agreement with Johns Hopkins Bloomberg School of Public Health

**Westat**®

# CARI Recording and Coding Process

| | Number of Interviews | Total Items Recorded | Question Assessment | Interviewer Assessment | Percent Recorded | Percent coded | Coded by |
|---|---|---|---|---|---|---|---|
| **Validation Study** | 304 | 198 | 198 | 0 | 30-100% | 15% | Survey methodologist |
| **Pretest** | 125 | 330 | 330 | 0 | 30-100% | 7% | Survey methodologist |
| **Round 1** | 8,245 | 64 | 52 | 25 | 100% | 24% | Non-project staff |
| **Round 2** | 7,077 | 58 | 36 | 25 | 100% | 24% | Non-project staff |
| **Round 3** | 5,799 | 48 | 25 | 25 | 100% | 22% | Project staff |

Westat®

# CARI Coding Staffing

- Round 2
  - Non-project staff
- Round 3
  - Project specific staff
  - Selection of project staff for coding
    - Field supervisor
    - Quality control skilled interviewer
    - Bilingual interviewer
- CARI Coding Staff trained via WebEx session
  - Review of the coding scheme
  - Review of the functionality of the web-based coding system

**Westat**®

# Sample Question from NHATS

Let's start with your laundry. By laundry we mean cleaning your clothing, sheets, and towels.

Which answer best describes how your laundry got done in the last month? Did you always do this by yourself, always do it together with someone else, did someone else always do it for you, or did it vary?

- Interviewer reads: "Now – chores around the house – laundry – do you do your own laundry?"
- Respondent answers:  "Yes, not often but I do it."
- Interviewer probes: "And you do it by yourself?"

**Westat**®

# Summary of Cases

- Round 2
  - CARI result code calculated for each case based on coding of individual questions and case summary codes.
  - Result codes then tallied across cases for each interviewer.

- Round 3
  - Project staff coded ~5 cases and assigned a score of excellent, very good, good, fair, or poor.

**Westat®**

# Summary Result Codes

- Cleared
- Inaudible
- Skill issues
- Validation risk

**Westat**®

# Computer Ranking

- Calculated the percentage of times that the interviewer made an error across some of their coded cases.

  - total number of errors / total number of questions

- Divided the interviewers into quintiles based on the percentage of the times that the interviewer made an error.

**Westat**®

# Overall Agreement

| Qualitative judgment | Quantitative score | | | | |
|---|---|---|---|---|---|
| | Excellent | Very good | Good | Fair | Poor |
| Excellent | 11 | 2 | 4 | 0 | 0 |
| Very good | 9 | 13 | 5 | 5 | 0 |
| Good | 2 | 2 | 8 | 4 | 1 |
| Fair | 1 | 1 | 2 | 11 | 8 |
| Poor | 0 | 0 | 2 | 2 | 11 |

Weighted kappa = .58

**Westat®**

# Agreement on Dichotomous Outcome

|  | Quantitative Score | |
| --- | :---: | :---: |
| Qualitative judgment | Excellent - Good | Fair or Poor |
| Excellent - Good | 56 | 10 |
| Fair or Poor | 6 | 32 |

Kappa = .68

Hit rate: 32/38 = .84
Miss rate: 6/38 =.16
False alarm rate: 10/66 = .15
$d' = 2$

Westat®

# Does it make a difference who does the coding?

## Round 2: Non-project staff

| Result | Frequency | Percent |
|--------|-----------|---------|
| Cleared | 765 | 58.7% |
| Changed meaning | 417 | 32.0 |
| Probing difficulty | 22 | 1.7 |
| Professionalism concern | 5 | 0.4 |
| Multiple skill issues | 64 | 4.9 |
| Validation risk | <u>31</u> | <u>2.4</u> |
| Total | 1,304 | 100 |

## Round 3: Project specific staff

| Result | Frequency | Percent |
|--------|-----------|---------|
| Cleared | 638 | 58.6% |
| Changed meaning | 276 | 25.4 |
| Probing difficulty | 29 | 1.7 |
| Professionalism concern | 2 | 0.4 |
| Multiple skill issues | 129 | 11.9 |
| Validation risk | <u>14</u> | <u>1.3</u> |
| | 1,088 | 100 |

Westat®

# Classification of Cases by Round

| Result | Round 2:<br>Telephone Coders | Round 3:<br>NHATS staff coders |
|---|---|---|
| Cleared | 58.7% | 58.6% |
| Skill issues | 39.0 | 40.1 |
| Validation risk | 2.4 | 1.3 |
| Total | 100 | 100 |
| (N) | (1,304) | (1,088) |
| p = .33 | | |

Westat®

# Differences in Case Level Coding – Positive Characteristics
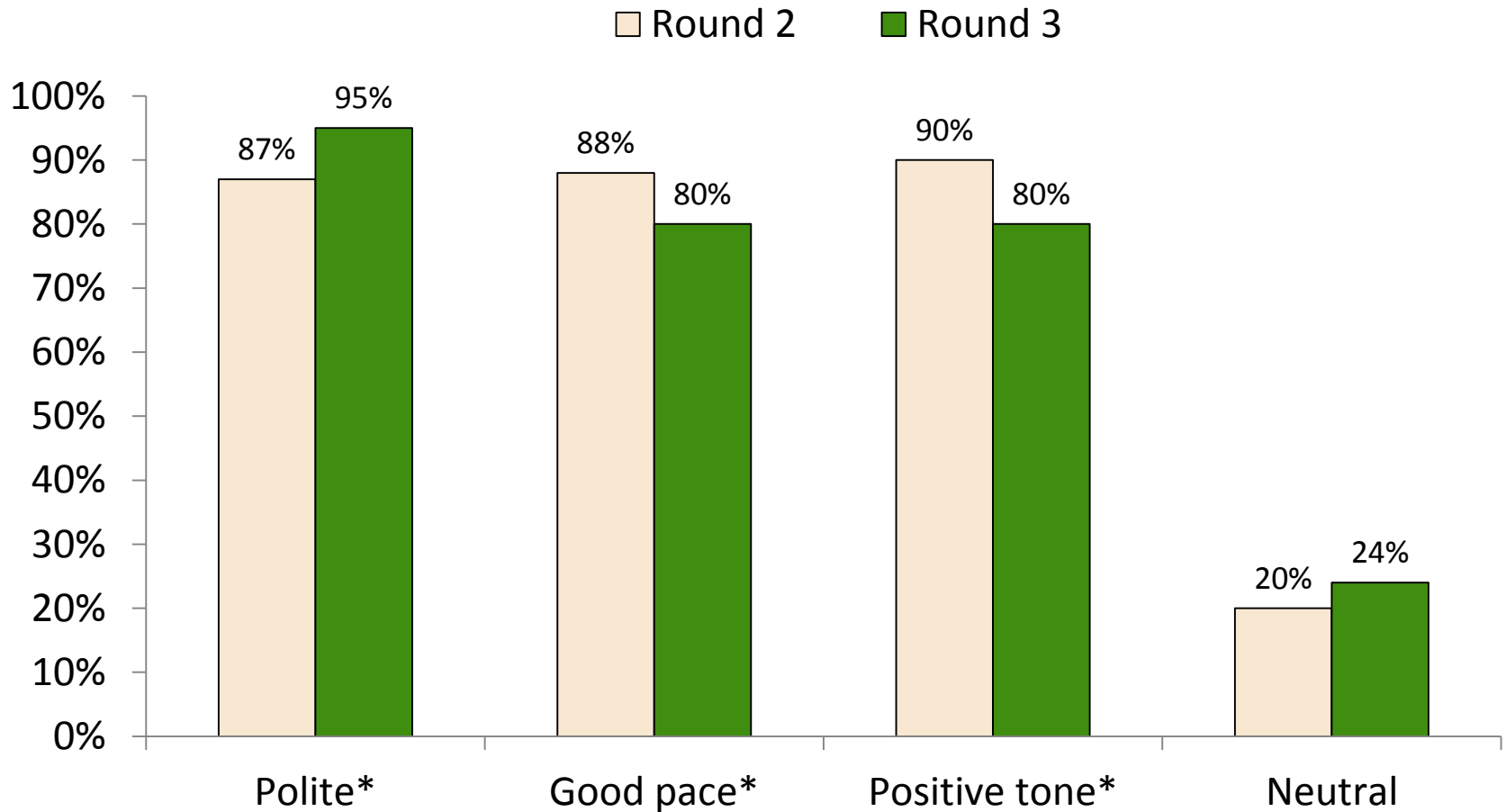


*P < .05

# Differences in Case Level Coding – Negative Characteristics



*P < .05

Westat®
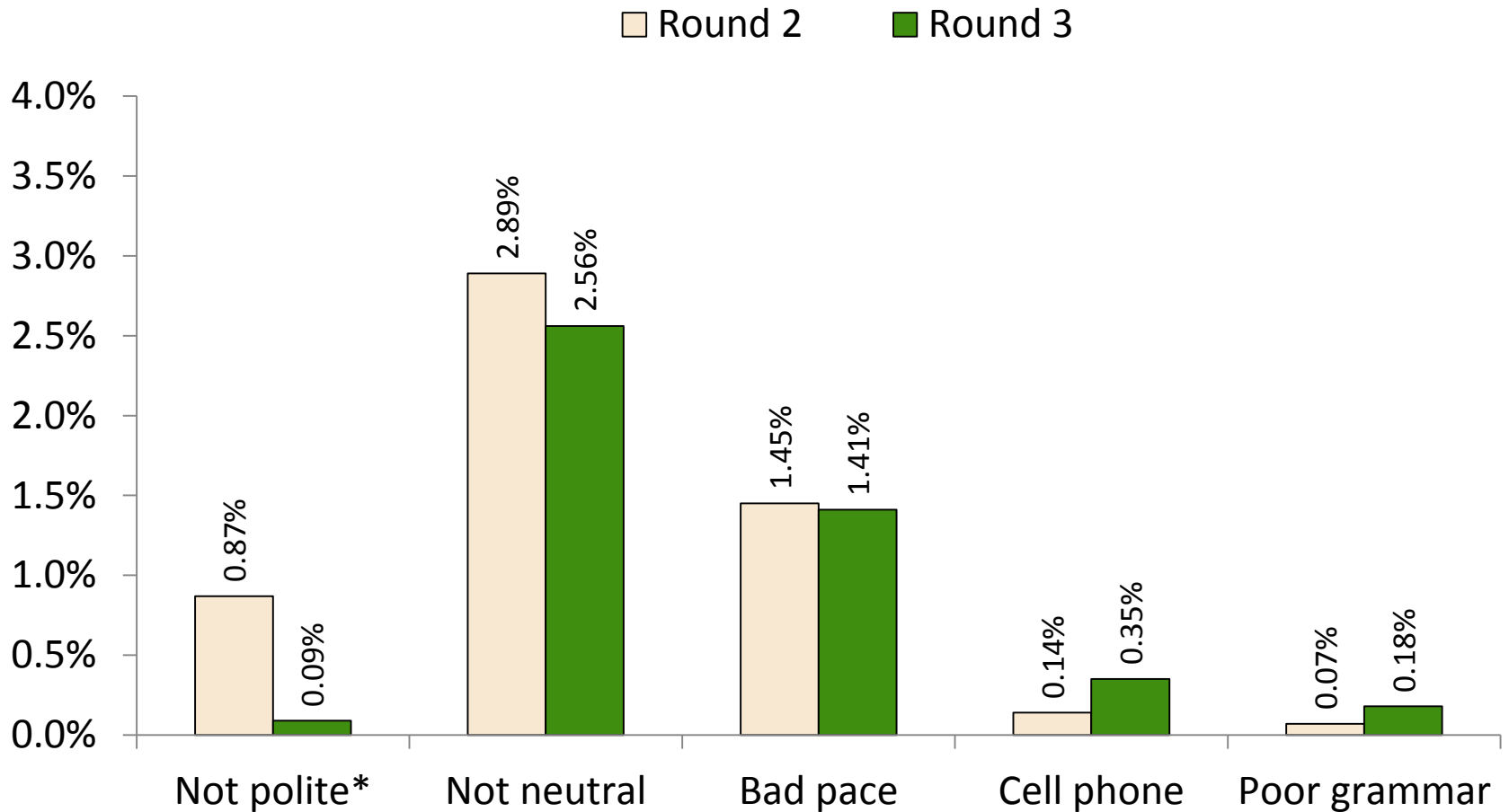
# Discussion

- CARI system provides a useful and meaningful method for providing feedback to interviewers.

- Output from the system needs to be in a format that the supervisors can easily interpret and use to provide feedback to the interviewers and readily identify interviewers that need retraining or feedback.

- It is important to identify both good and poor interviewer behavior.

- The type of coder that was used for this study does not seem to influence the distribution of codes that are assigned to the cases.

- Need to understand more about how the feedback given from the CARI system influences interviewer behavior and data quality.

Westat®