Contrasting Stylized Questions of Sleep with Diary Measures from the American Time Use Survey

Robin Kaplan, Brandon Kopp, and Polly Phipps
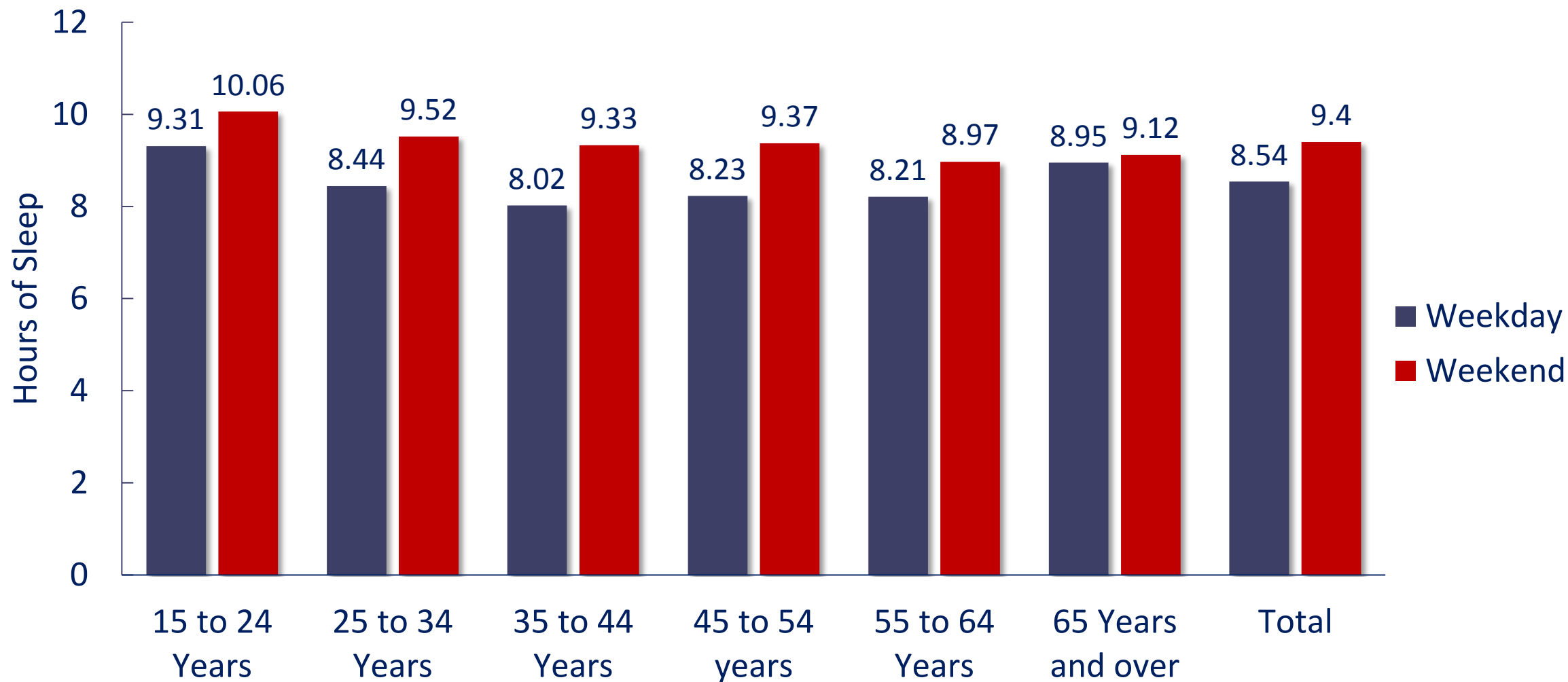
Bureau of Labor Statistics

Office of Survey Methods Research (OSMR)

QDET2

November 12, 2016

*Opinions expressed are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

BLS

Sleep Duration in the ATUS

Hours of Sleep

	Weekday	Weekend
15 to 24 Years	9.31	10.06
25 to 34 Years	8.44	9.52
35 to 44 Years	8.02	9.33
45 to 54 years	8.23	9.37
55 to 64 Years	8.21	8.97
65 Years and over	8.95	9.12
Total	8.54	9.4

■ Weekday
■ Weekend

* Data are annual averages for 2014.

BLS

Diary Collection of Sleep Estimates

Introduction:

"Now I'd like to find out how you spent your time yesterday, [day of week], [date], from 4:00 in the morning until 4:00 a.m. this morning. I'll be asking where you were and who else was with you. If an activity is too personal, there's no need to mention it."

Sample interview excerpt:

- I: "What were you doing at 4 a.m.?"
- R: "I was sleeping."
- I: "What time did you wake up?"
- R: "7:00."
- I: "Okay. And what did you do next?"

Stylized Questions

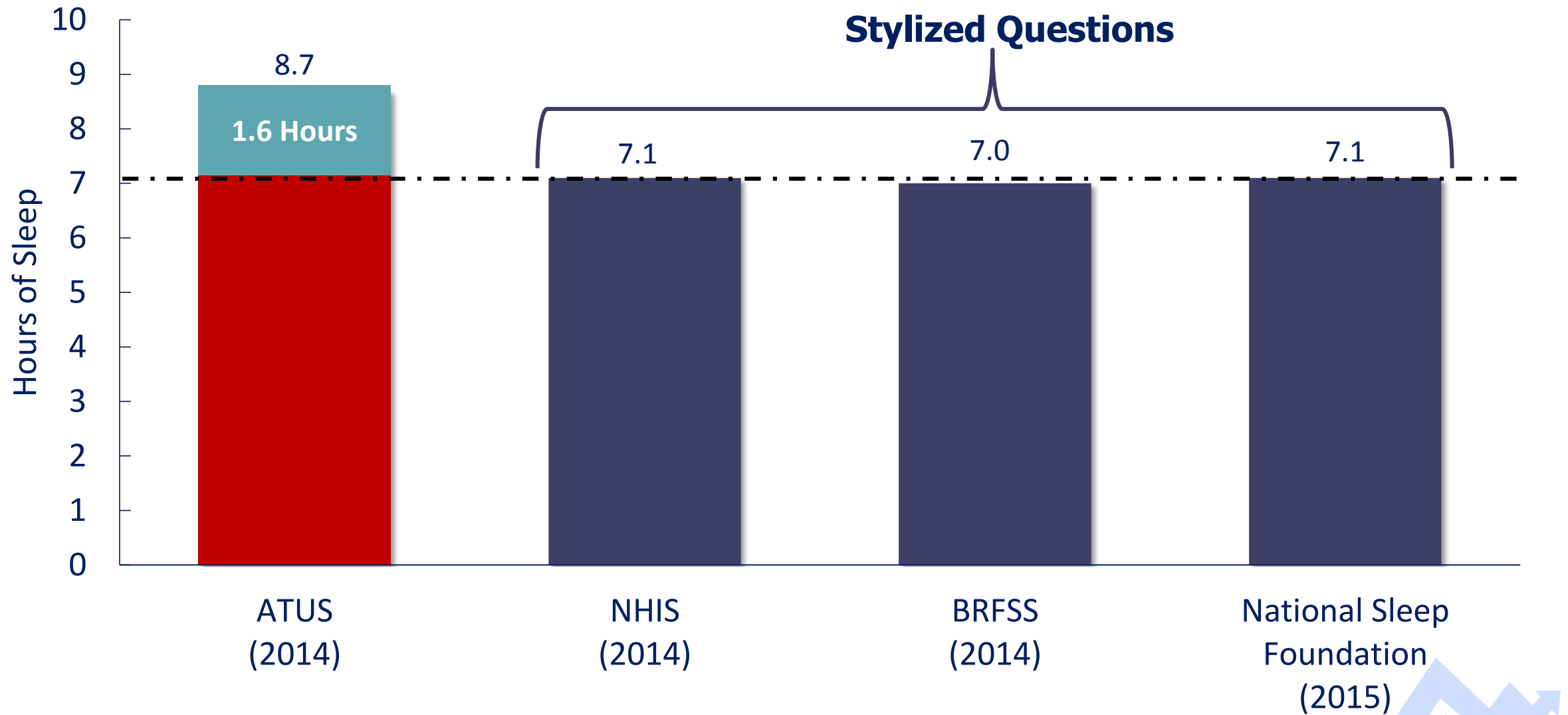| National Health Interview Survey	• "On *average*, how many hours of sleep do you get in a 24-hour period?"
Behavioral Risk Factor Surveillance System	• "On *average*, how many hours of sleep do you get in a 24-hour period?"
National Sleep Foundation	• "…, about how much actual sleep would you estimate you typically get on work nights or weeknights?"

BLS

Sleep Duration in Other Surveys

Stylized Questions

Hours of Sleep

8.7 — ATUS (2014)
1.6 Hours
7.1 — NHIS (2014)
7.0 — BRFSS (2014)
7.1 — National Sleep Foundation (2015)

BLS

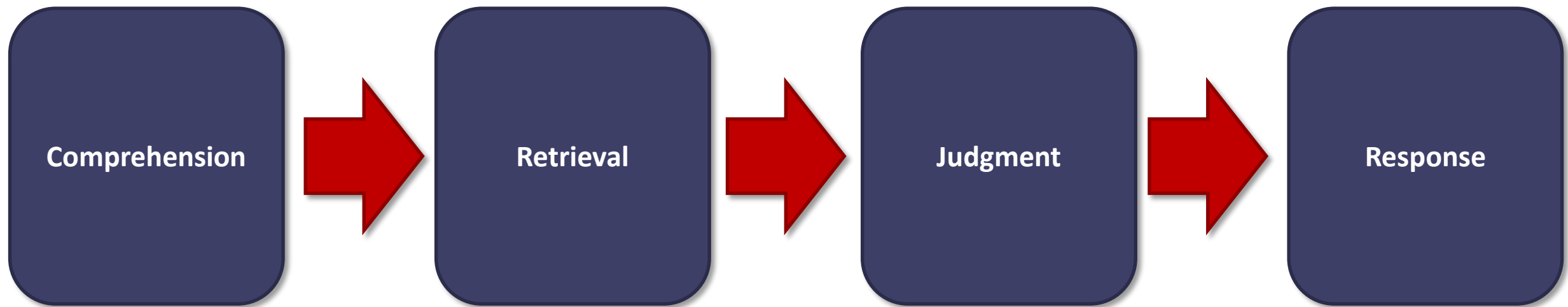# Potential Reasons for the Sleep Gap

Diary

- List of activities

- Broad sleep lexicon

- 30-minute rule

- Naps are captured

Stylized

- Single activity

- Sleep not defined

- Single sleep episodes

- Naps may not be captured

Response Process

```
Comprehension  ➡  Retrieval  ➡  Judgment  ➡  Response
```

* Tourangeau, Rips, & Rasinski (2000)

Questionnaire Evaluation Methods

■ Behavior coding

■ Cognitive interviews

■ Quantitative online study

■ Validation study

Behavior Coding Study

Research Interests

- Opportunity to observe the diary collection of sleep episodes in the production setting

- Identify concepts and tasks that respondents and interviewers may struggle with in reporting and recording sleep

- Detect practices that could be associated with sleep measurement issues

BLS

Study Methods

- N=104 transcripts, 2008 ATUS respondents

  - ~ 3 interviews per interviewer, for the 36 ATUS interviewers on staff

  - Non-random

  - Study sample is demographically representative of full 2008 sample
    - 8.53 hours of sleep on average, closely matching ATUS estimates

- Behavior coding of I and R interactions

  - Main unit of analysis is each conversation around sleep
    - Conversational "turns" related to sleep activity
    - Type of sleep activity; sleep time/duration; respondent recall and qualifications; interviewer probing

  - Reliability coding of 12% of transcripts

  - All results met a minimum kappa 0.4 +, range was from 0.45 to 1.0

BLS

Conversational "turns"

Sample interview excerpt:

- I: "What were you doing at 4 a.m.?"
- R: "I was sleeping."
- I: "What time did you wake up?"
- R: "7:00."
- I: "Okay. And what did you do next?"

Reporting and Recording of Sleep

- Average of 21 turns to record sleep activities in an interview
  - Accounted for 12% of diary interactions

- Sleep mentions are times during an interview when respondent reports a wake time, sleep time or sleep duration
  - ▶ Each mention of waking or going to sleep took 6.7 turns
  - ▶ Reporting and recording most complex for <u>going to sleep</u>
    - Average of 11 turns for going to sleep compared to 3.9 for waking

Interviewer Probes

- **"What time did you wake up?"**
  - ▶ Leading question wording 69% of the time
    - – What time did you get up?
    - – What time did you wake up yesterday morning?

- **"I went to bed"**
  - ▶ Respondents used this phrase in 73% of interviews
  - ▶ Interviewers followed up with a probe 95% of the time
    - – Probes taught in training were used 42% of the time
    - – Most common non-training probe was "did you go to sleep immediately?"

Time Qualifiers

■ Use of "about," "around," "maybe"

  ▶ An indicator of respondent confidence in answers

  ▶ Average of 1.8 qualifiers

  ▶ Respondents were slightly more likely to qualify responses when reporting time of falling asleep (61%) compared to waking up (49%)

Recall Strategies

■ Explicit recall strategy reported in 25% of 331 sleep mentions across all interviews
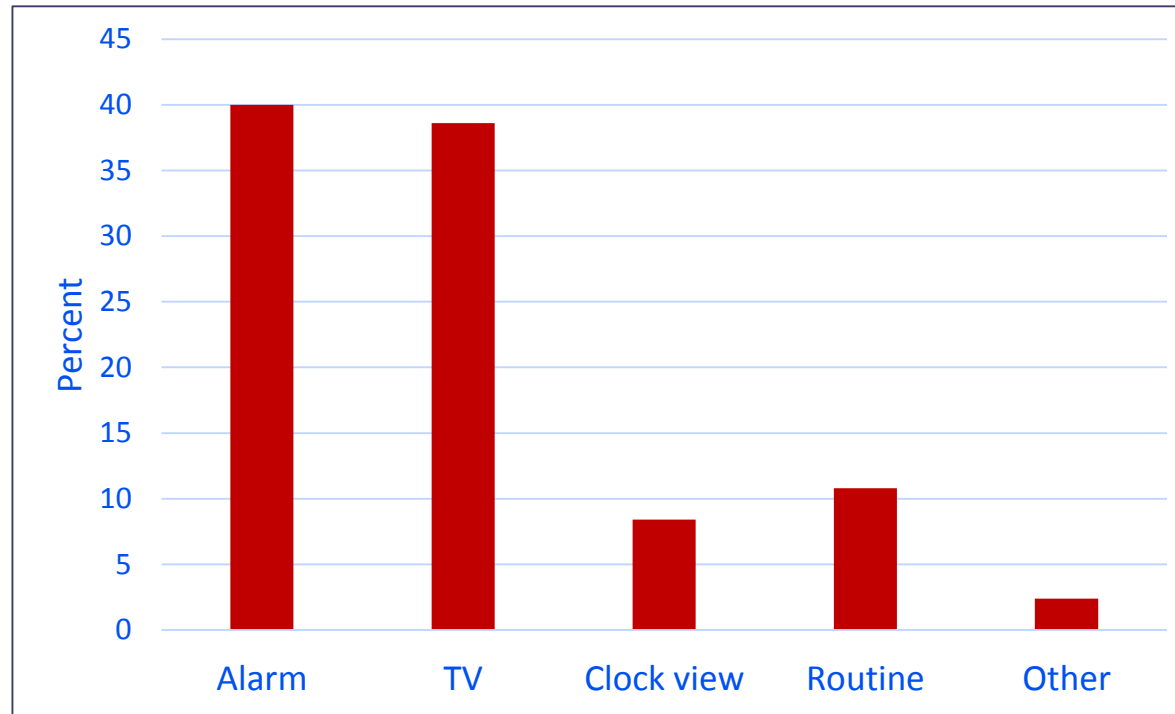
Behavior Coding Summary

■ Interactions are complex and respondents are uncertain, especially for falling asleep time

▶ Respondents may have difficulty recalling or estimating sleep

▶ Recall strategies are used to formulate waking and sleep response

■ Interviewer probing may contribute to underreporting of time spent awake at night and in morning, inflating sleep time

■ But behavior coding study only suggestive

▶ Indirect evidence of measurement error

▶ No follow up questions or stylized questions for comparison

Cognitive Interview Study

Methodology

- Interviewed participants in the lab
- Asked both diary and stylized questions
  - ▶ Counterbalanced order
- Retrospective probes

| Comprehension | → | Retrieval | → | Judgment | → | Response |
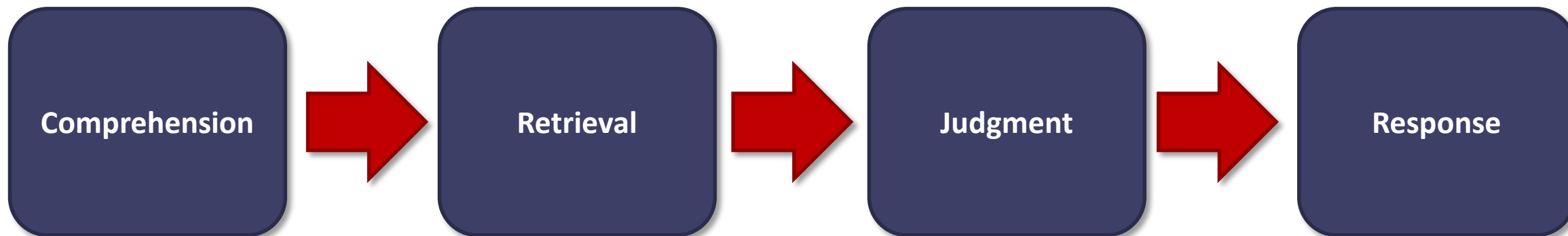
* Tourangeau, Rips, & Rasinski (2000)

Demographics

- Recruited from our participant database

- *N* = 29
  - ▶ 11 male; 18 female
  - ▶ Mean age = 46 (*SD* = 14.05)
  - ▶ Age range of 21 to 69 years old

BLS

Comprehension

- What does "sleep" mean to you?

	Diary mean hours of sleep	Stylized mean hours of sleep
Narrow sleep definition (n=15)	7.25	5.93
Broad sleep definition (n=13)	8.18	6.93

Diary Recall

■ Wake time is easy to recall

▶ Alarm

▶ Look at clock when waking up

▶ Structured schedule

■ Sleep time is difficult to recall

▶ Sleep time varies

▶ Don't check clock

▶ Infer time from TV program

Stylized Estimates

Strategy	Frequency	Mean hours of sleep
Recalled directly	10	6.22
Rate retrieval	9	7.02
Rate and adjustment	4	5.62
Calculation	4	5.50
Estimate/Guess	2	6.50

Reporting

What is the [maximum/minimum] number of hours appropriate to sleep per night?	Diary	Stylized
Maximum (M = 8.87 hrs)	r = 0.39*	r = -0.10, n.s.
Minimum (M = 5.97 hrs)	r = 0.23, n.s.	r = 0.40*

*$p < .05$

Summary

- **Comprehension**
  - ▶ Broader definition = more sleep
- **Recall**
  - ▶ Easy to recall wake times
  - ▶ Hard to recall sleep times
- **Estimation**
  - ▶ Wide range of strategies that may be prone to measurement error
- **Reporting**
  - ▶ Social desirability concerns

25

Quantitative Study

Design

- Participants completed Diary and Stylized sleep measures
  - ▶ Embedded questions about other activities

- 2 Definition Conditions
  - ▶ With definitions; Without definitions
- 2 Order Conditions
  - ▶ Diary first; Stylized first

Participants

- 1233 participants completed the study
  - Recruited through Amazon Mechanical Turk
  - 46% male; 54% female
  - Mean age = 36 (SD = 11 ; range = 19 - 77)

Definition Condition

■ **Sleeping**: By sleep, we mean the number of hours you actually spend sleeping. This may be different from the number of hours you spend in your bed, time you spend preparing to go to sleep, resting with your eyes closed but not actually asleep. Please <u>include</u> any times you were sleeping during the day (or napping)

▶ Also read definitions for other activities (exercise, work)

Sleep Questions

■ Modified version of the ATUS interview

■ Past Week Stylized

▶ "Thinking back to the past week (that is, during the previous 7 days), how many hours of sleep did you get on average each weeknight (excluding weekends)?"
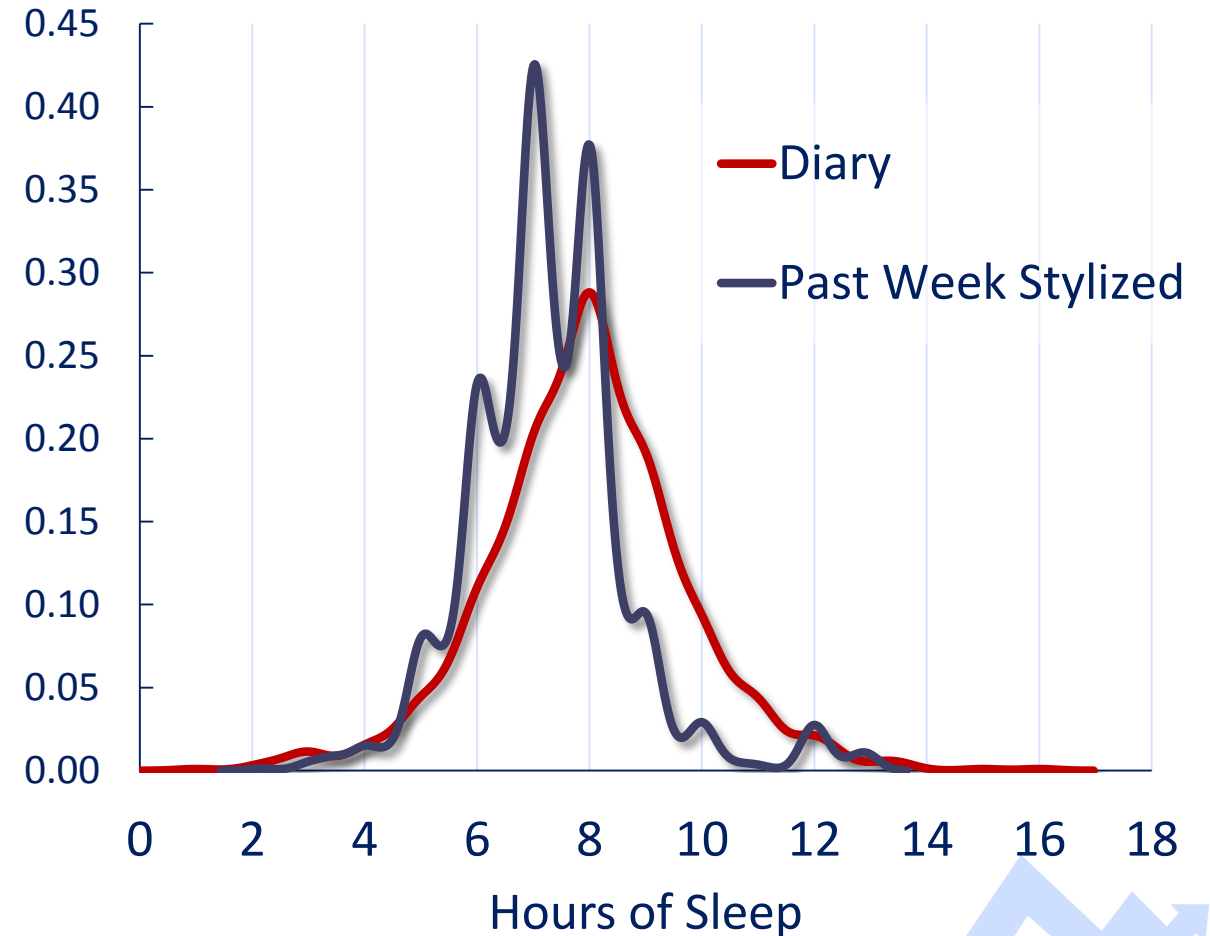
	Activity	Start Time (Hour)	Start Time (Minute)	Start Time (AM or PM)	End Time (Hour)	End Time (Minute)	End Time (AM or PM)
2	Grooming	7	:00	AM	7	:30	AM
3	Eating/Drinking	7	:30	AM	7	:45	AM
4	Travelling & going from place to place	7	:45	AM	8	:15	AM
5	Work & Work-Related activities	8	:15	AM	12	:45	PM

BLS

Results

Means and Standard Deviations
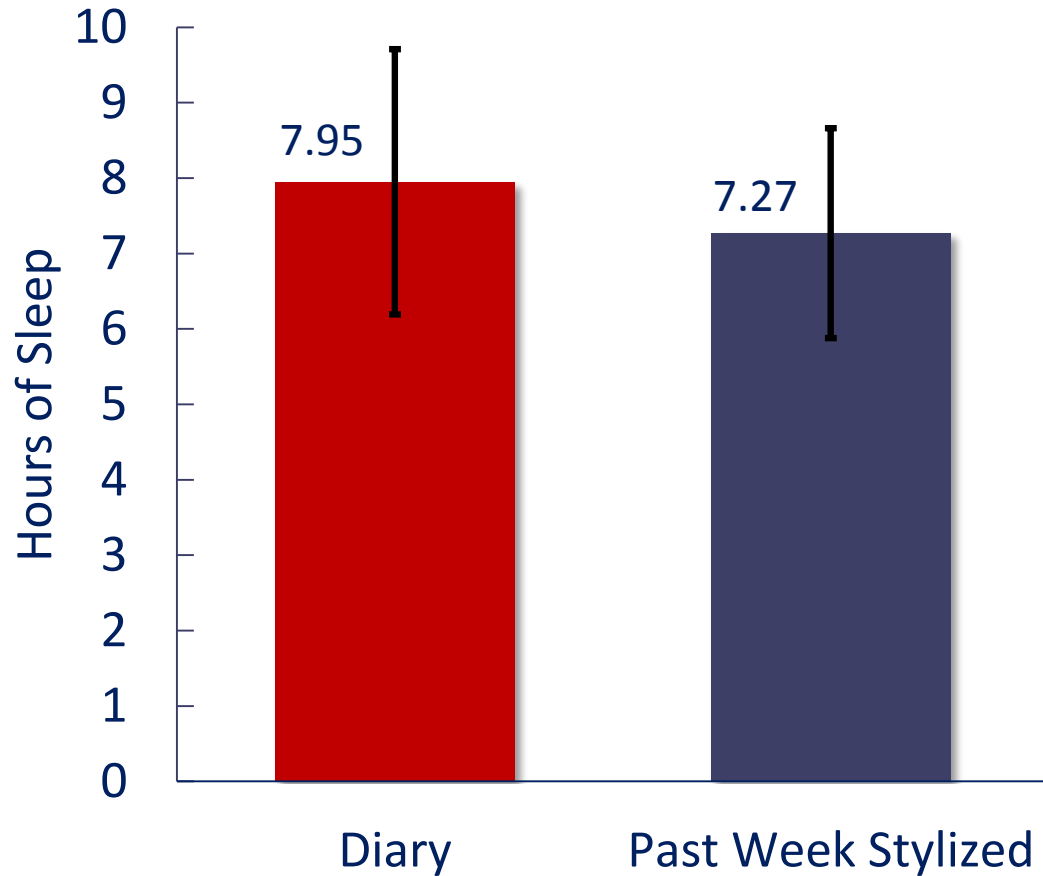
Response Distribution

Results

- 2 (Question Type) X 2 (Definition) X 2 (Order) ANOVA

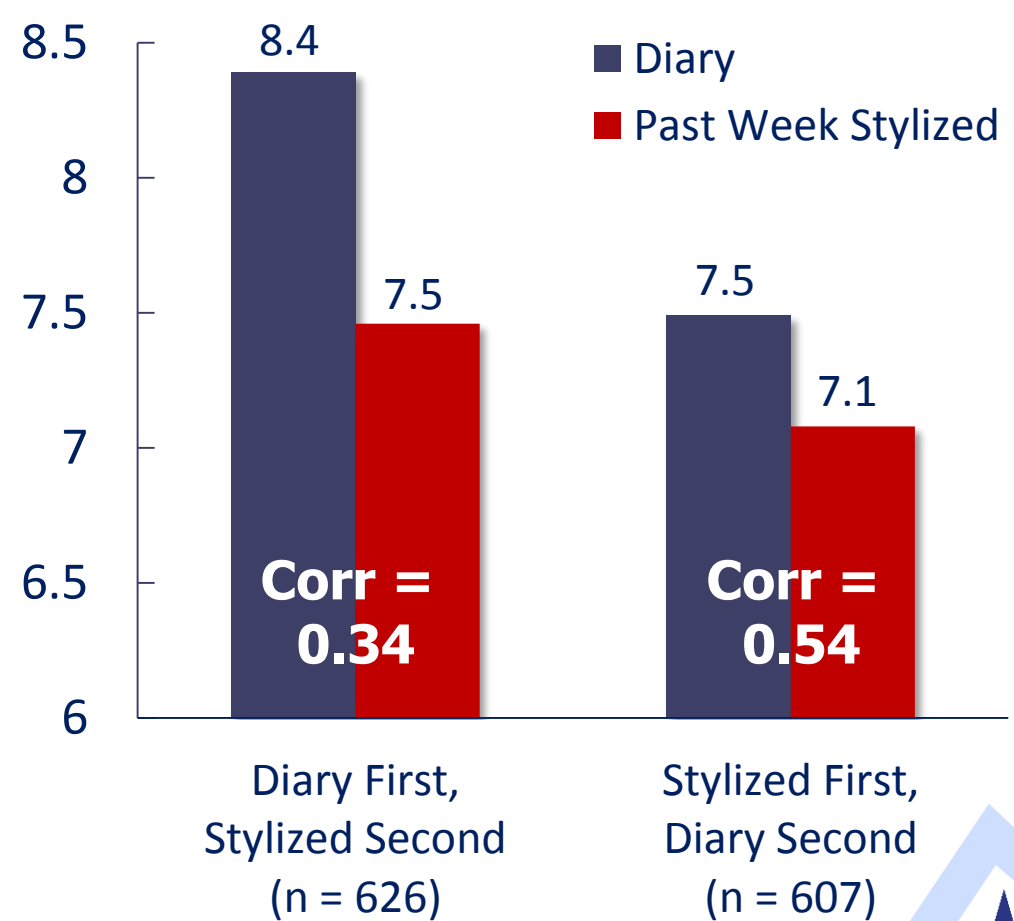|	df	F Value	p-value
Question Type	1	200.49	p < 0.01
Question Type X Definition	1	5.84	p < 0.05
Question Type X Order	1	29.49	p < 0.01
Question Type X Definition X Order	1	0.18	p = 0.67
Residuals	1229		

- Main effect of Question Type

- Question Type X Order Interaction

- Question Type X Definition Interaction

Results

Question Type X Definition

Hours of Sleep

- Diary
- Past Week Stylized

Without Definitions (n = 622): Diary 8.0, Past Week Stylized 7.2
With Definitions (n = 611): Diary 7.9, Past Week Stylized 7.3

Question Type X Order

- Diary
- Past Week Stylized

Diary First, Stylized Second (n = 626): Diary 8.4, Past Week Stylized 7.5, Corr = 0.34
Stylized First, Diary Second (n = 607): Diary 7.5, Past Week Stylized 7.1, Corr = 0.54
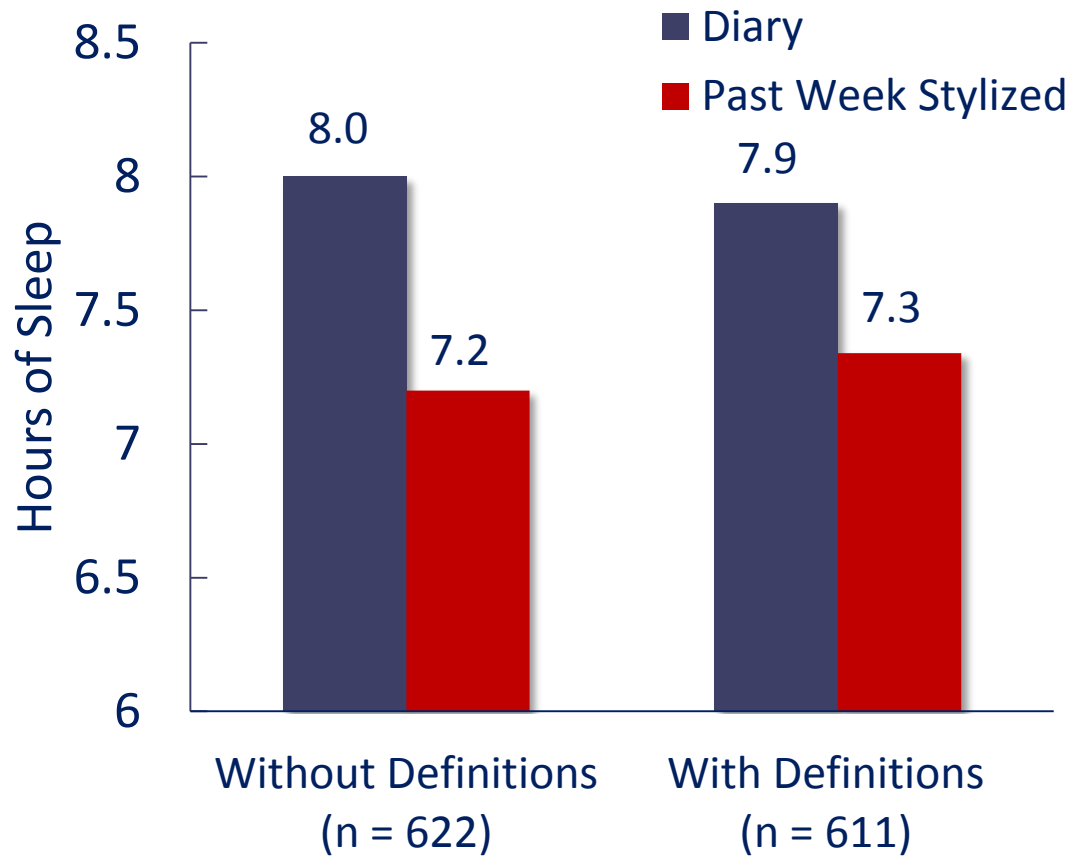
Summary

- Diary > Stylized

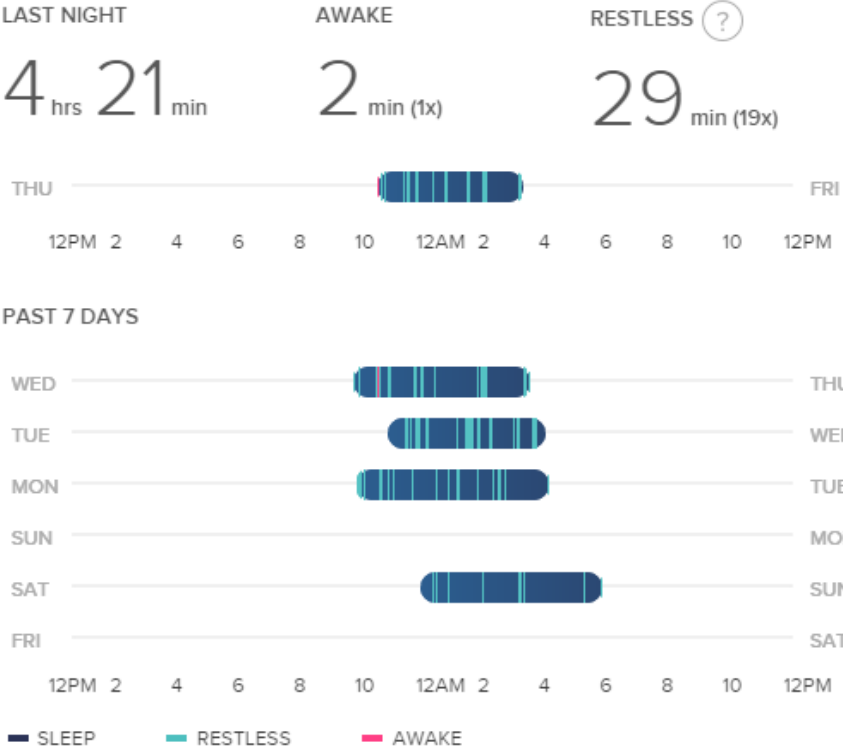- Definitions brought measures closer together

- Order effect

Validation Study

Wearable Activity Devices

Methodology

- 2 visits one week apart

  ▶ Visit 1:
  - Demographic questions
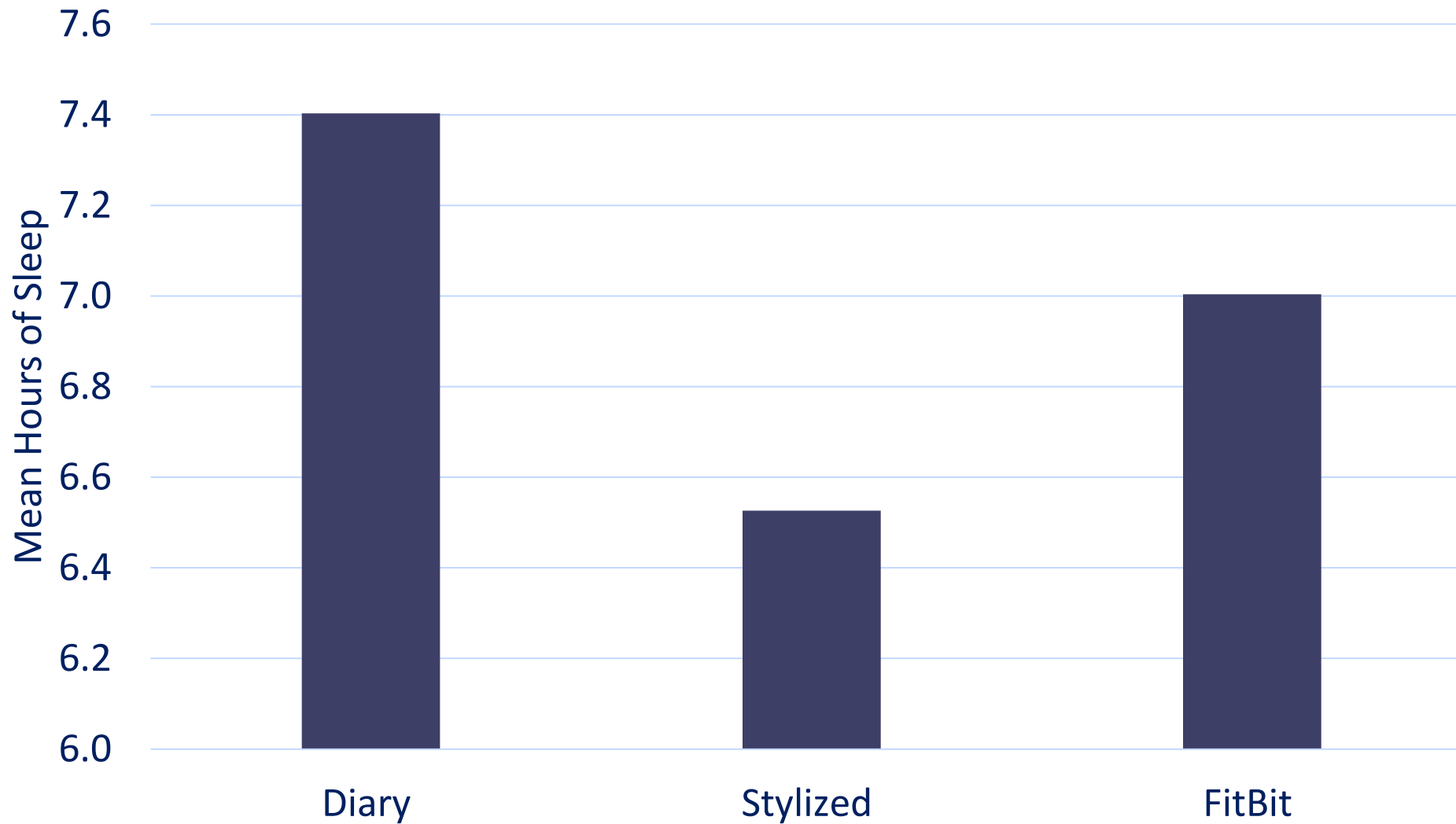 - Fitbit instructions

  ▶ Visit 2:
  - ATUS interview
  - Stylized questions
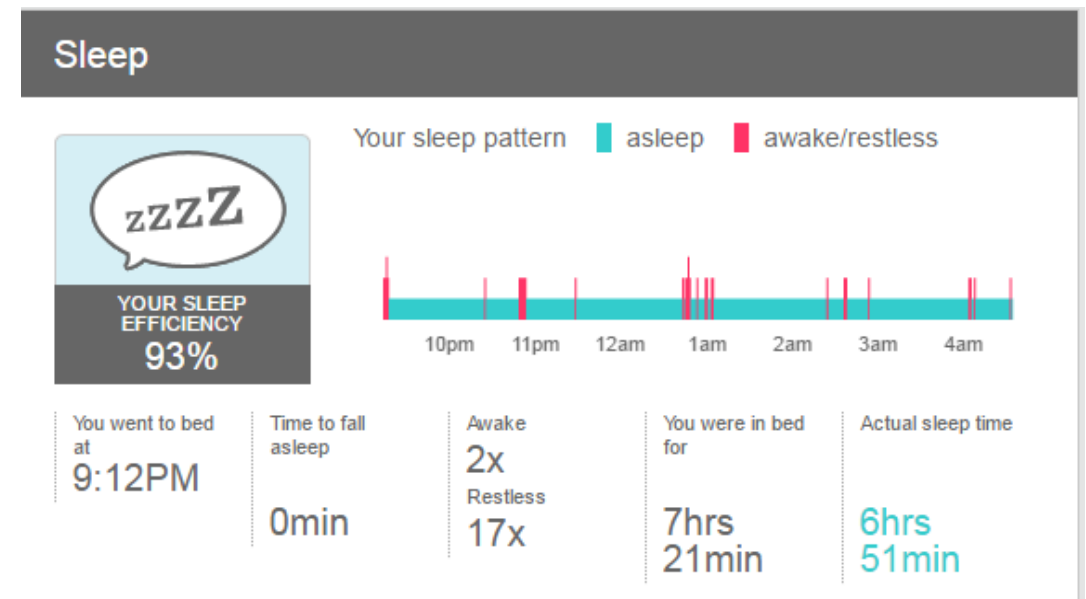 - Debriefed about Fitbit-recorded sleep data

Participants

- ■ 19 participants from the Washington, DC metro area
  - ▶ 13 female, 6 male
  - ▶ Mean age = 44.58 years

Results

Debriefing

- **12/19 thought the Fitbit sleep data was about right**
  - ▶ 5 thought it underestimated
  - ▶ 2 thought it overestimated
- **Recall aid**
  - ▶ "Oh yeah, I hit the snooze button that morning"

Logistics with Validation Research

- **Measurement error**
  - ▶ Records tossing and turning as times awake (underestimates)
  - ▶ Records inactivity as naps (overestimates)
- **User error**
  - ▶ Did not adhere to instructions to wear device each night (n=5)
  - ▶ Only wore device one work night of the week (n=7)
  - ▶ Lost device (n=2)
  - ▶ Device fell off night before Visit 2 (n=1)

Summary

- Preliminary evidence that an objective sleep measure might fall in between diary and stylized reports of sleep
  - But has its own set of measurement and user error
- Device data can be used to aid memory in time use research
- Conduct more interviews in the future

Conclusions

Multi-method approach

■ **Behavior coding**

▶ Pros: Uses actual production survey transcripts, examine questions, probes, answers, unscripted conversation

▶ Cons: Time consuming!, Only indirect evidence of measurement error, no comparison to stylized questions

■ **Cognitive interviews**

▶ Pros: Insight into each stage of the response process, generate hypotheses

▶ Cons: Small sample size, limited geographic region, qualitative data

Multi-method approach

■ **<u>Quantitative study</u>**

▶ Pros: Collect data quickly with large sample, make statistical comparisons

▶ Cons: Convenience sample (voluntary panel), no data on accuracy

■ **<u>Validation study</u>**

▶ Pros: self-report data validated against objective data, relatively inexpensive with new technologies

▶ Cons: logistically complex, user error, and measurement error

Multi-method approach

■ **Recommendations:**

▶ Weigh the pros and cons of each method

▶ Capitalize on the strengths of each type of method

▶ Build results off of one another to gain deeper understanding of an issue

Contact Information

Robin Kaplan, Brandon Kopp, and Polly Phipps
Bureau of Labor Statistics
Office of Survey Methods Research (OSMR)
QDET2
November 12, 2016

E-mail:
Kaplan.Robin@bls.gov
Phipps.Polly@bls.gov

BLS