

Screening for Specific Experiences: Fine-Tuning Questions in Multi-phase Testing

Mandi Martinez and Mary Davis

Center for Survey Measurement, U. S. Census Bureau

Disclaimer: This talk is to inform interested parties of research and to encourage discussion. The views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

Multi-Phase Pre-Testing of the School Crime Supplement (SCS)

- ▶ A supplement to the National Crime Victimization Survey
 - ▶ Various school related topics, but pretesting focused on section of revised bullying questions
 - ▶ SCS administered every 2 years during January through June
- ▶ Conducted pre-testing prior to finalizing the 2015, 2017, and 2019 supplements
 - ▶ 2015 Redesign prompted by new uniform definition of bullying developed by the Centers for Disease Control and Prevention
 - ▶ Bullying is a complex social processes, interpreted subjectively, and difficult to measure
 - ▶ Different limitations during each phase of testing (i.e., adhering to new definition, time constraints, preserving the trend)

Sources of data that informed approach

- ▶ Cognitive interviews
 - ▶ Allows us to test question performance and revise question wording multiple times before fielding
 - ▶ 2015: 2 rounds, 40 interviews
 - ▶ 2017: 3 rounds, 30 interviews
 - ▶ 2019: 4 rounds, 36 interviews (in progress, 29 interviews completed and analyzed)
- ▶ FR Debriefings
 - ▶ Provides feedback on how measures are performing during administration
 - ▶ FRs can identify questions that are difficult to administer or appear to cause respondent confusion
- ▶ Data from previous fielding
 - ▶ Data can be examined to
 - ▶ Results can be compared to previous years

What drove each phase of testing?

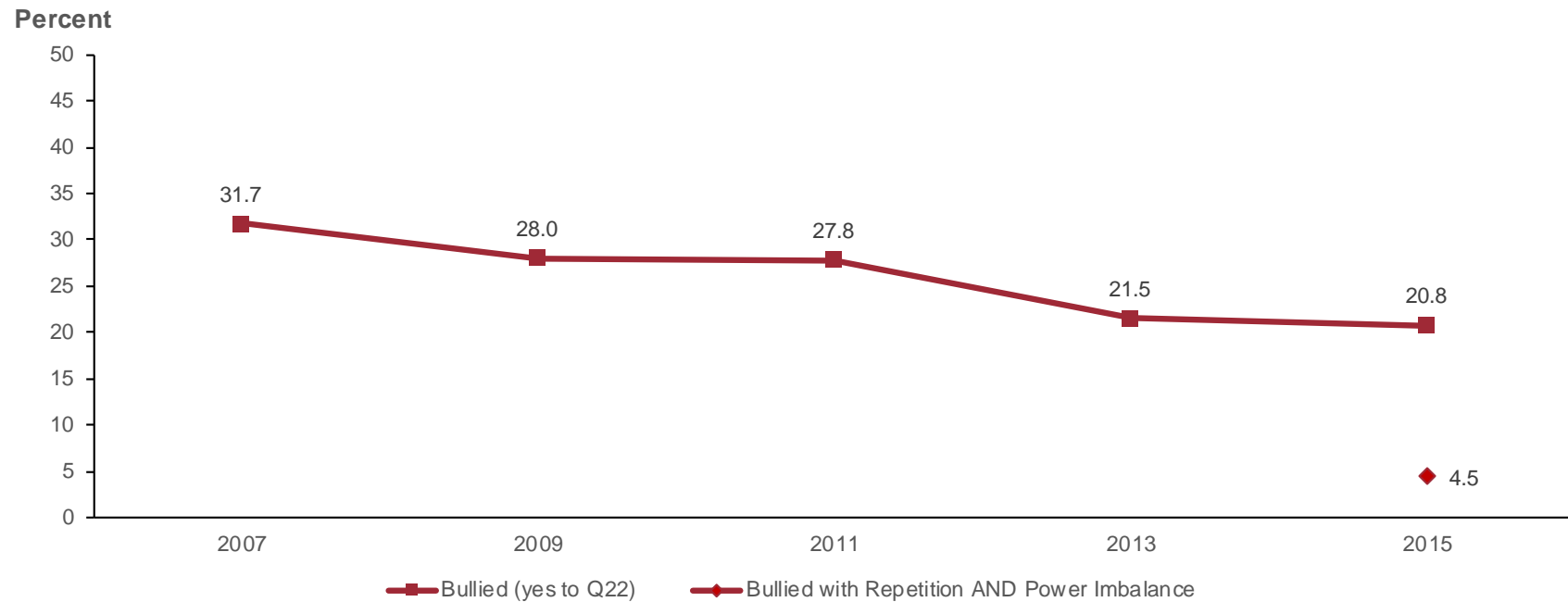
- ▶ Prior to 2015: Make measure of bullying align with official definition developed by the Centers for Disease Control and Prevention
 - ▶ New criteria to measure: repetition and power imbalances
 - ▶ Split-ballot in field
- ▶ Prior to 2017
 - ▶ Improve questions that measure repetition and power imbalances
 - ▶ Examine students' perceptions of what constitutes bullying
- ▶ Prior to 2019 (in progress)
 - ▶ Remove the word "bullying" from the supplement
 - ▶ Possible split-ballot

Evolution of Screener: Phase 1

- ▶ Goal: Align with CDC's new uniform definition of bullying
- ▶ Criteria of repetition and power imbalance
- ▶ Developed 2 versions of questionnaire for split ballot experiment
 - ▶ Version 1:
 - ▶ Kept original SCS measure of bullying (Q22)
 - ▶ Added 2 follow up questions to measure repetition and power imbalance
 - ▶ Version 2:
 - ▶ Replaced Q22 with a new, longer question
 - ▶ Question provides respondent with definition of bullying that specifies the requirement that experiences involve repetition and power imbalance
- ▶ Data from fielding showed that Q22 alone produced a similar rate of bullying
- ▶ Both Version 1 (Q22 + follow-ups) and Version 2 produced significantly lower rates of bullying
- ▶ Rates from version 1 and version 2 significantly differed from each other

Trends in Bullying (2007-2015): Data from the SCS

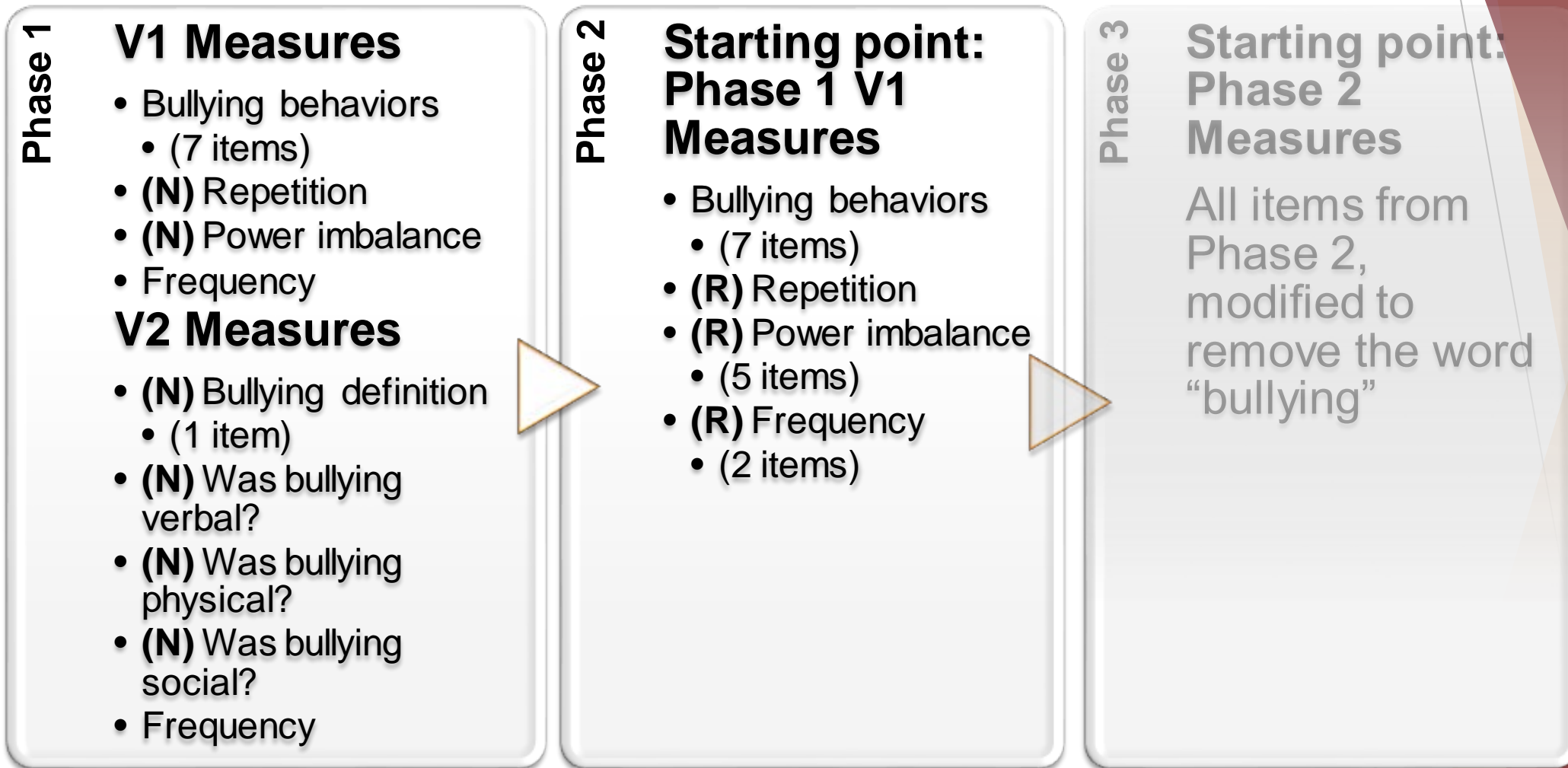
Percentage of students ages 12 through 18 who reported being bullied or called a hate-related word at school, by year: School years 2006–07 through 2014–15



From Data Points released by NCES:

1. **Changes in Bullying Victimization and Hate-Related Words at School Since 2007**, accessed at <https://nces.ed.gov/pubs2018/2018095.pdf>
2. **Repetition and Power Imbalance in Bullying Victimization at School**, accessed at <https://nces.ed.gov/pubs2018/2018093.pdf>

Evolution of Screener Items: Phase 1 to Phase 2

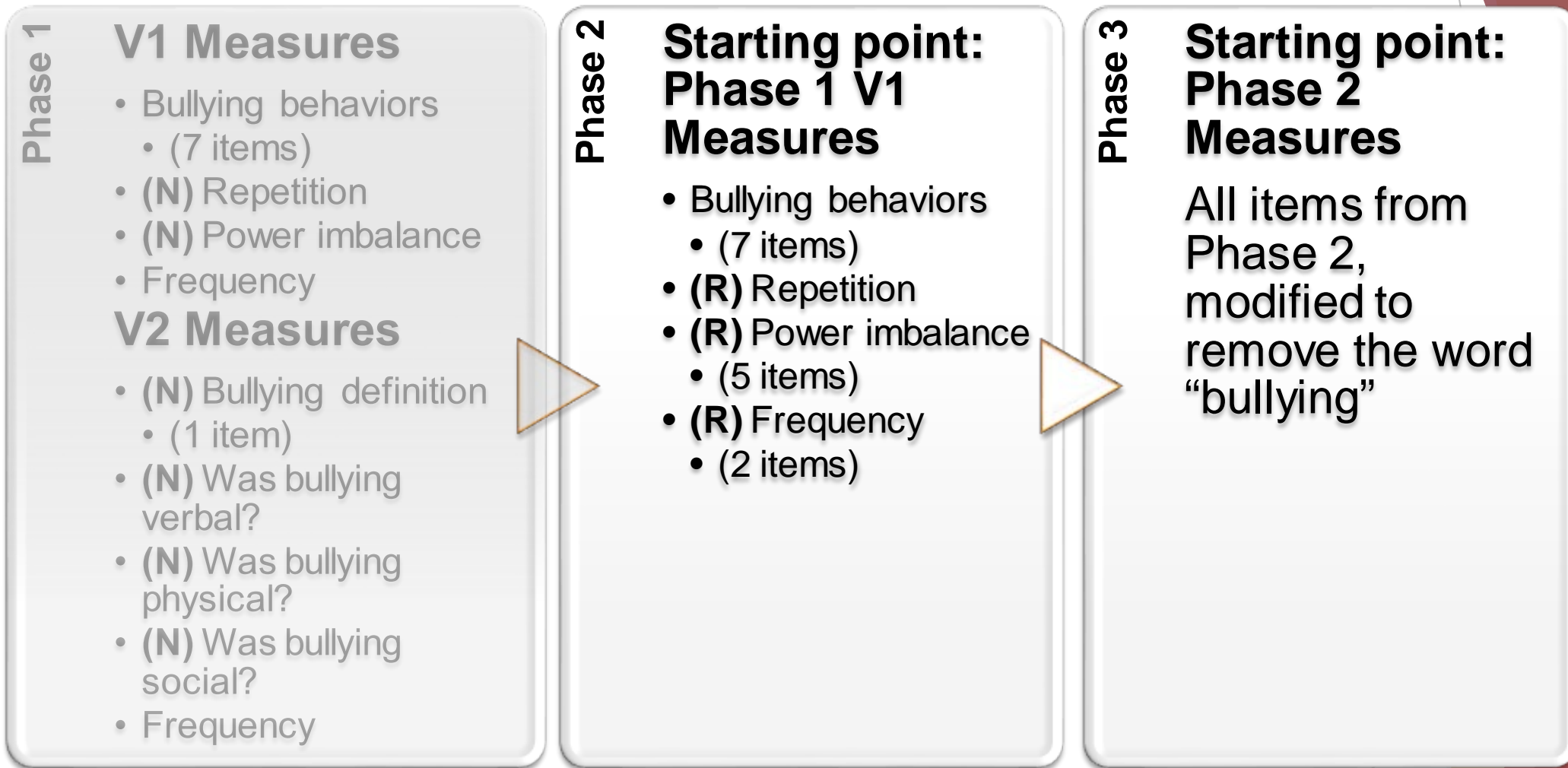


(N) = New
(R) = Revised

Evolution of Screener: Phase 2

- ▶ Goal: Revise 2015 Version 1 follow-up questions to improve measurement
- ▶ Q22 was left untouched to preserve the trend
- ▶ Follow-up questions were revised over 3 iterative rounds of testing
 - ▶ Modified questions measuring frequency of bullying and likelihood of repetition
 - ▶ Changed power imbalance questions from a single yes/no to a series of yes/no questions about different types of power imbalances
- ▶ Used respondents' narratives about experiences to determine if they meet the criteria of bullied, and if their responses to questions aligned with their experiences
- ▶ This allowed us to assess whether new measures produce false positives or false negatives
- ▶ Data from fielding showed that Q22 continued to produce similar rate, and new measures produced a slightly lower rate

Evolution of Screener Items: Phase 2 to Phase 3



(N) = New

(R) = Revised

Evolution of Screener: Phase 3

- ▶ Goal: Revise 2017 questions to remove the word “bullying”
- ▶ Used phrasing from original Q22 to describe bullying behaviors,
“Now I have some questions about what students do at school that make you feel bad or are hurtful to you. ~~We often refer to this as being bullied. (...)~~”
- ▶ After administering questions and follow-up probes for bullying section, asked students if they considered their experiences to be bullying.
- ▶ Of the 14 students whose narratives led us to qualify them as bullied, three did not consider themselves victims of bullying
- ▶ Student feedback on our removal of the word bullying is overwhelmingly positive (some neutral, none negative)

Conclusions

- ▶ Multi-phase testing allows for data from the field to inform approach for subsequent rounds
- ▶ The process allows questionnaire designers to address different issues in different phases, helping navigate limitations during a single phase
- ▶ Cognitive interviews can be used to collect additional data on issues that are not able to be addressed until subsequent rounds

Questions?

mandi.martinez@census.gov