# Appearance before the Comparative Systems Subcommittee of the Response Systems to Adult Sexual Assault Crimes Panel (RSP)

William J. Sabol, Acting Director
Allen Beck, Senior Statistical Advisory
Bureau of Justice Statistics
April 11, 2014

# Issues of interest to the Comparative Systems Subcommittee

- How accurate are the Workplace and Gender Relations Survey (WGRS) data?
- Does the instrument measure what it purports to measure?
- Is the context of the survey as a “workplace” the right context?
- What survey design considerations are important to understand about the WGRS to help the Committee assess its reliability?
- What are the key design differences between the NVCS and other major surveys of rape and sexual assault, such as NISVS and WGRS?

# Overview

- Brief background on the NCVS
- Key survey design issues and their implications for measuring rape and sexual assault
- Comparison of NCVS, NISVS (military) and Workplace and Gender Relations Survey on key design elements
- BJS's efforts to improve its measurement of rape and sexual assault in the NCVS.

# Key survey design elements and their implications for measuring rape and sexual assault

- Cost and burden
- Purpose
- Periodicity
- Context
- Sample design/sample size/response rates
- Mode
- Respondent selection
- Recall/reference period
- Bounding and telescoping
- Screening for events
- Cuing
- Counting events (incidence) vs. people (prevalence)

# Key survey design elements and their implications for measuring rape and sexual assault

- Purpose: The breadth or narrowness of a survey's scope can impact the measurement of sexual assault. NCVS's purposes are broader than NISVS and WGRS in measuring several types of crime, describing victims' experiences, describing offenders, and meeting legislative mandates.
- Periodicity: Continuous fielding facilitates maintenance of a trained field staff; intermittent fielding requires new training and potential loss of expertise.
- Context: Respondents react to the context in which a survey is presented; a crime survey focuses recall on events respondents consider to be crimes; a health survey on other issues. It is not clear that one focus or another, by itself, will lead to higher or lower reporting of events.

# Key survey design elements and their implications for measuring rape and sexual assault

- Sample design: Random digit dialing and the WGRS Web survey have lower response rates than (the NCVS's) address-based sample.
  - Low response rates raise questions of bias in the estimates; the direction of bias is not known *a priori* and depends upon the nature of the response patterns
    - OMB requires a bias-analysis plan for surveys with expected response rates below 80%.
    - WGRS provides confidentiality but not anonymity; response patterns by groups can be assessed.
  - Avidity bias: Persons interested in the survey or more engaged in the topic of the survey may be more likely to respond than those with less interest in it. This may affect RDD and Web surveys, where followup with sampled respondents is more difficult to do.

# Key survey design elements and their implications for measuring rape and sexual assault

- Sample size/response rates: Associated with precision of estimates
- Mode: (Costs vary)
  - In-person vs. telephone: NCVS uses both; no mode effects detected
  - Self-administered instruments (e.g., Web-based instruments)
    - No interviewer effects associated with responding to sensitive questions
    - Potential difficulties in negotiating a two-step screener/incident approach and distinguishing separate incidents from behaviors occurring within an incident.
  - Computer-assisted modes (CAPI, CATI, ACASI) provide greater control over the interview process.
  - ACASI considered to be the gold standard for rape and sexual assault.
  - In-person modes are more expensive than phone/web/mail.

# Key survey design elements and their implications for measuring rape and sexual assault

- Respondent selection: One person per HH (e.g., head of HH or random selection) vs. multiple persons per HH
  - Multiple persons per HH is less expensive, provides for greater statistical precision of estimates (less variance), and provides for more complete reporting of HH crime. (Recall NCVS's purpose as an omnibus survey is to measure HH crimes in addition to crimes against persons.)
  - Presence of others may increase inhibitions to report some crimes, especially if the perpetrator is present
  - Presence of others may cause fear of retribution.

# Key design elements and their implications for measuring rape and sexual assault

- Reference period/recall:
  - Recall is more complete with a shorter reference period
  - Longer reference periods: Respondents may be at risk of an event occurring for a longer period of time and therefore may report more events covering the reference period.
- Bounding/telescoping: Telescoping is the temporal displacement of events that leads to over-reporting of events; for surveys that ask about events occurring during a period of time until "today" respondents may "forward telescope" and bring events that occurred before the start of the reference period into the reference period.

# Key survey design elements and their implications for measuring rape and sexual assault

- <u>Bounding/telescoping (cont'd)</u>:
  - Not bounding the reference period is likely to lead to inclusion of events that are not eligible
  - NCVS's first interview bounds subsequent interviews; gives the respondent a mental anchor (e.g., "Since our last interview, did…occur?"). Effect of bounding is to reduce estimates for the first interview by 30-40%.
  - NCVS now uses statistical adjustments for bounding; cost saving measure to include the first interview but to adjust for bounding based upon observed relationships between time in sample and reports of crimes.
  - CNSTAT report on RSA recommended bounding adjustments for first interviews to avoid overestimating counts of events.

# Key survey design elements and their implications for measuring rape and sexual assault

- Screening for events:
  - One-step strategies both screen for (identify) events and classify them at the same time
    - All events taken as given; if respondent reports the same event (incident) in different ways, it gets counted twice.
    - E.g., in a single incident of attempted rape a respondent may also experience unwanted sexual touching and would respond "yes" to separate questions about these types of unwanted sexual contact; both responses could be counted.
  - Two-step (NCVS) strategy:
    - Screened events may not be classified as crimes or as the crimes for which they are screened; e.g., some events screened as property crimes end up getting classified as violent crimes based on the information reported in the incident report.
    - Ineligible events are identified by the incident form and not counted.
  - Lower rates observed with a two-step process due to unfounding of incidents; generally more accurate estimates.

# Key survey design elements and their implications for measuring rape and sexual assault

- Cuing: Use of more or less explicit terms for sexual acts and body parts to capture information about incidents; volume and number of cues
  - More cues lead to more reporting
  - The more explicit the cues, the more reporting of incidents
  - Explicit cuing may affect response rates, especially with younger persons (recall the omnibus nature of the NCVS).

# Key survey design elements and their implications for measuring rape and sexual assault

- <u>Incidence and prevalence</u>:
  - Incidence: Measure the number of (criminal) events
  - Prevalence: Measure the number of people in a population who experienced at least one event of interest.
  - Generally, incident rates will be higher because each event is counted whereas in a prevalence, each person is counted
  - BJS's report "*Measuring the Prevalence of Crime with the National Crime Victimization Survey*" (Janet Lauritsen and Maribeth L. Rezey, September 5, 2013):

    (Females only; age 12 and older):
    - RSA victimization rate: 1.9 per 1,000 (or 0.2%)
    - RSA prevalence: 1 per 1,000 (or 0.1%)
    - Mean number of RSA events per victim: 2
    - Percent of victims with one event: 85%.

# One-year prevalence estimates from NISVS and WGRS

NISVS (military)

2010 estimates for "contact sexual violence":

Women, 18-59, general population:
5.2% (4.4, 6.1)

Active duty women:
5.6% (4.2, 6.9)

Wives of active duty men:
3.6% (2.6, 4.6)

WGRS

"Unwanted sexual contact" estimates:

Active duty women:

2006: 6.8%

2010: 4.4%

2012: 6.1%

# Key features of NCVS, NISVS and WGRS

| Design feature | NCVS | NISVS (military) | WGRS |
|---|---|---|---|
| Periodicity | Annual; continuous data collection | 2010 | 2006, 2010, 2012 |
| Purpose | Omnibus crime survey, major categories of violent and property crimes; “dark figure;” describe victims and offenders; supplements: SCS, ID theft, stalking; PPCS; mandates. | Intimate partner violence, sexual violence, and stalking | Sexual assault (unwanted sexual contact); sexual harassment; workplace relations. |
| Context | Crime survey | Health, safety, crime | Workplace relations, crime |

# Key features of NCVS, NISVS and WGRS

| Design feature | NCVS | NISVS (military) | WGRS |
|---|---|---|---|
| Sample design | Address frame: All residential housing units and non-institutional group quarters. Excludes military ships, military barracks, other non-institutional GQs such as shelters for abused women, soup kitchens, etc. Members of military living in housing units included. | Random digit dialing (RDD) for general population; Active Duty Master file (frame for active duty women); Defense Enrollment Eligibility Reporting System (frame for wives of active duty men) | List frame: Active duty population |

# Key features of NCVS, NISVS and WGRS

| Design feature | NCVS | NISVS (Military) | WGRS |
|---|---|---|---|
| Respondent selection | All persons 12 and older in sampled HH | 1 person 18 or older per HH | Active duty service personnel (individual) |
| Sample size | ~180,000 persons in 90,000 HH | ~5,400 for active duty women; 1,408 interviews;<br><br>~4,900 for wives of active duty men; 1,428 interviews. | Sampled 108,000 active duty persons with oversampling for groups; oversampling rates not available; 25,900 responses. |
| Response rates | ~90% for HH; ~88% for individuals | 27.5% for GP; 26.5% active duty; 29.2% wives | 24% in 2012; 32% in 2010 |

# Key features of NCVS, NISVS and WGRS

| Design feature | NCVS | NISVS (Military) | WGRS |
|---|---|---|---|
| Recall/reference period | 6 months | 1-year; 3-year; lifetime | 1-year |
| Bounding | First interview; statistical adjustment | None | None |
| Screening | 2-step: Screener followed by incident form; events “founded” or “unfounded”; each incident reported separately up to 5 incidents; series victimization | Screener/incident | Screener/event that “had the greatest effect on you” |

# Key features of NCVS, NISVS and WGRS

| Design feature | NCVS | NISVS (Military) | WGRS |
|---|---|---|---|
| Cuing | Non-explicit and non-extensive; use of terms “rape” and “sexual assault” whose meaning may differ among respondents; confirm with incident information about presence, type of attack (in-person threat/attack), injury items, and narrative. | Sexual violence cues: Completed/attempt rape; completed alcohol/drug-facilitated penetration; sexual coercion; being made to penetrate someone else; unwanted sexual contact experiences; unwanted non-contact sexual experience. | Unwanted sexual contact cues: Sexually touching, attempted and completed intercourse, attempted and completed oral/anal sex or penetration with object |
| Mode | In-person (first); CAPI phone | CATI | Web |

# Some WGRS design considerations (from "Survey Methodology," Slide 53 of the June 27, 2013 presentation)

Sample design (2012): Oversampling of certain groups (gender, race, rank, Service, deployment status).

What were the sample sizes of each group? What was the sampling fraction for each subgroup? What was the coverage of the frame?

Learned from past administrations of the survey to identify demographic groups that traditionally had low response rates and oversampled these groups

What specific groups had low response rates? How were the sample sizes of these groups adjusted for 2012?

Response rates:
- 24% in 2012
- 32% in 2010
- Not shown for 2006

What were the response rates for each subsampled group? How much did each contribute to the overall response rates? Were there differential response rates for subgroups?
How did the response rates for subgroups change over time, and how did the changes in subgroup response rates affect the change in the overall response rate?

# Some WGRS design considerations (Slide 53 of the June 27, 2013 presentation)

| | |
|---|---|
| Bias analysis | Given that anonymity was not assured and the frame can identify subgroups, how well did the sample perform? (Weighted only by the inverse of the probability of selection?) |
| Weighting adjustments | Describe the various weighting adjustments, e.g., for the probability of selection, nonresponse. Did nonresponse adjustments assume that nonrespondents arose at random? What evidence is there for this assumption? (Analyze the characteristics of the sample in relation to known attributes of the frame.) |
| Statistical controls to ensure reliability | What statistical controls were used to ensure reliability? Modeling of relationships? Weighting adjustments? Others? |

# Some WGRS design considerations (Slide 53 of the June 27, 2013 presentation)

| | |
|---|---|
| Analysis of incidents (item 33): | How many incidents were reported and what was the distribution of all incidents (how many persons reported 1, 2, … 9 or more)? How do the incident counts compare to other surveys? |
| Workplace environment and reporting (policy implications for DoD) | ITEMS 11-22: Trust of supervisor; confidence in the chain of command; morale; satisfaction with military life, etc. Relationship between these responses and reporting unwanted sexual contact to authorities. |

# BJS efforts to improve measurement of rape and sexual assault in the NCVS

- Challenge/problem: Competing national estimates of the level of rape and sexual assault in the U.S.
  - Official, NCVS estimates typically lower than estimates from other federal agencies (e.g., CDC) and private groups
  - National Violence Against Women: Estimates are about 4 times higher than comparable NCVS estimates
  - National Survey of Intimate Partner Violence (NISVS): Estimates are about 10 times higher than NCVS.
- Differences in estimates raises questions about the best methods for collecting self-report data on RSA
- No consensus in the field about the optimum set of procedures for doing this.

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: Differing approaches

- NCVS's "criminal justice" approach: Incident-based, emphasizing forcible rape and criminal events (a more exclusive definition); use of a two-step screening/incident process to identify and classify events into crime categories; bounds the first interview to address telescoping.
- "Public health" approach: Utilizes behaviorally-specific (explicit references to actions and body parts) to cue respondents; presented as health surveys; classify events based on the initial questions that reference behaviors with limited to no followup on the number of incidents; RDD with relatively low response rates; no bounding.

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: BJS National Survey on Health and Safety (NSHS):

- Methodological study to identify, develop, and test alternative methods for collecting self-report data on RSA; aim is to inform the methods to be used in the NCVS for RSA.
- Focus of NSHS is on RSA among adult females
  - Ultimately, BJS will collect RSA from males and females
  - Prevalence of RSA among males is low and neither NCVS or NISVS produce reliable, one-year estimates for males.
- Three distinct components of NSHS:
  - Develop and pilot test an ACASI design
  - Develop and pilot test an RDD design
  - Conduct detailed comparisons of the two designs against each other and the existing NCVS

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: NSHS (cont'd)

- Will employ two methods: (1) in-person interviews and (2) telephone interviews
- Universe for the study is restricted to five large metropolitan areas (Dallas-Ft. Worth-Arlington; Los Angeles-Long Beach-Santa Ana; Miami-Ft. Lauderdale-Pompano Beach; New York-Northern NJ-Long Island; Phoenix-Mesa-Glendale)
- Study comprises four samples:
  - Probability sample using personal visit; ABS to enable personal visit of females aged 18-49;
  - Probability sample using telephone; RDD landline frame and RDD for cell phones for females 18 and older
  - Purposive sample of women aged 18-39: Supplement the analysis and comparisons of designs 1 and 2 by focusing on a group at higher risk of RSA
  - Purposive sample of known victims: Assess instrumention to ensure that it works

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: Design features in the NSHS study:

- Sample design
- Screening strategy
- Reference period
- Bounding
- Cuing strategy
- Context
- Respondent selection
- Modes of interviewing (telephone, in-person interviews, ACASI)

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: Structure of interviews in NSHS:

- Short series of demographic and personal items;
- Series of victimization screening questions to determine if the respondent has experiences 12 types of RSA in their lifetime or in the past 12 months
  - Use of explicit behavioral cues, with references to behaviors that make up the legal definition of RSA
  - Telephone protocol screens first for lifetime, then past 12 months
  - ACASI respondents asked first about past 12 months, then lifetime
- Respondents with one or more incidents in the past 12 months go to a detailed incident form
  - Expect about 5% of respondents will report an incident in the past 12 months;
  - Number of incidents capped at three (burden issue).

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: Structure of interviews (cont'd)

- Following incident reports or screener if no incidents reported, respondents presented with two vignettes that characterize different levels of coercion or alcohol use and are asked to answer questions about the vignettes
- Study will interview about 18,100 respondents; a subset of 1,000 will be contacted about two-weeks after the first interview for a followup interview, which will be identical to the first interview (verification of reports).

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: Schedule for NSHS:

- Feasibility study: May-June 2014: Test all survey protocols
- Pilot study: February-May 2015: Full implementation in five metropolitan areas; 18,100 interviews
- File preparation (cleaning, weighting, etc.) and analysis:
- Final report: February/March 2016

# BJS efforts to improve measurement of rape and sexual assault in the NCVS: Analysis questions of NSHS:

- Comparative response rates and correlates of response rates
- Difference in incidents by methodology
- How respondents differ with respect to defining RSA
- Method that produces the most consistent reports (reinterview)
- Differences in reported incident and measures of survey quality (e.g., missing data, level of effort to complete)
- Comparative costs of the two major designs