

# Creating Lesbian, Gay and Bisexual Visibility Through Advocating for the Collection of Sexual Orientation Data

Randall Sell

Alternate title:

A few of the things I've learned while trying to get the federal government to collect sexual orientation data

# What is demography?

- Demography is the scientific study of human population, including its size, distribution, composition, and factors that determine changes in its size, distribution, and composition.

Demography therefore focuses on four aspects of human population:

1. Size
2. Distribution
3. Composition
4. Change in Size, Distribution, or Composition

# Population Composition

## (Demographic Variables)

- Age
- Year of Birth
- Nativity
- Sex
- Race
- Ethnicity
- Religion
- Citizenship
- Marital Status
- Educational Level
- Labor Force Status
- Income/Wealth
- Sexual Orientation

# Why do people interested in health collect demographic data?

- Civil Rights Enforcement
- Program Administration/Evaluation
- Research
- Public Health Surveillance

# Do we really need to collect demographic data in research studies?

1. Is the variable relevant to the study?
2. How much will it cost to collect?
3. Will subjects refuse to answer questions or break-off interviews?
4. Could findings stigmatize populations?
5. Will there be enough power to analyze data?
6. Is the variable a proxy for more relevant variables?
7. Are there political ramifications of collecting this data?
8. Are there valid and reliable measures of the variable?

# 1. Is the variable relevant to the study?

- Catch 22: must prove relevancy before the variable can be added.
- Opposite of Catch 22: HHS requires the collection of race and ethnicity data.
- Do you consider the relevancy of every demographic variable in your research studies, or are some reflexively included?



## 2. How much will it cost to collect?

- Current NHIS supplement paid for by NCI cost \$1.8 million and consists of 7 sections.
- In 1993 you could add a question to NHIS for \$15,000.
- NHIS is estimating it will cost \$2.6 million to add a sexual orientation variable (including testing).
- Do you always consider the (opportunity) costs of including variables in your research?

### 3. Will subjects refuse to answer questions or break-off interviews?

- In 2002, 16.5% of NHIS respondents broke off interviews. Of these, 17.9% did so because the questions were too personal. (About 3% overall).
- Respondents are more likely to refuse or break off on income questions than virtually any other question (31% refused family income on 2002 NHIS).
- NESARC has had 0 break offs on the sexual orientation questions in over 30,000 interviews (with only 1.7% skipping these questions).
- Do you consider item refusal or break off before including a variable?

## 4. Could findings stigmatize populations?

- “for two centuries there have been scientists obsessed with proving that minorities, poor people, foreigners, and women are innately inferior to upper-class white males of northern European extraction.” (*The Science and Politics of Racial Research*)
- The Current Population Survey debated the addition of nativity to its dataset because of concerns that the data could be abused.
- Ask Larry Summers about data on sex.

#### 4. Could findings stigmatize populations?

- Health care professionals have used the following “treatments” to cure and/or suppress homosexual feelings and behaviors in the United States: surgical removal of the ovaries, clitoridectomy, castration, vasectomy, pudic nerve section, lobotomy, hypnosis, shock therapy (electric and chemical), aversion therapy, psychoanalysis (individual and group therapy), drug therapy (including hormones, LSD, sexual stimulants, and sexual depressants), primal therapy, vegetotherapy, prolonged abstinence and sitz baths.

## 4. Could findings stigmatize populations?

- Often data is released for public use and can be used by anyone to make their case. “Lies, damn lies, and statistics.” And the current director of NIH is advocating policies to make data more accessible to the public.
- Do you consider how data you collect could be abused?

## 5. Will there be enough power to analyze data?

- Asian and Pacific Islanders.
- Gays, lesbians and bisexuals (NESARC: identity 1.7%, attraction 5.1%).
- A solution is to oversample.
- Do you calculate power for each demographic variable before deciding to include it?

## 6. Is the variable a proxy for more relevant variables?

- Race: socioeconomic status, discrimination, cultural factors, unspecified biological differences, etc...
- If Race is a proxy for these other variables, then shouldn't we measure these instead?
- Sex vs. Gender
- Religion vs. Religiosity

## 7. Are there political ramifications of collecting this data?

- Race: Racial Privacy Initiative (California, Proposition 54, 2002)
- Sexual Orientation: Defunding of The National Health and Social Life Survey
- Have you ever been concerned with how collecting or reporting certain demographic data could impact your ability to conduct future research?



## 8. Are there valid and reliable measures of the variable?

- EX: **Sex** is not even asked in most telephone interviews, but assigned by the interviewer according to voice. Rarely are more than 2 response categories provided.
- EX: **Marital Status** was asked differently by each of 18 different programs in a HRSA SPNS multi-site evaluation.
- EX: **Income** was found to be underreported by 12% on the Current Population Survey (and as presented earlier, is often refused).

## 8. Are there valid and reliable measures of the variable?

- EX: **Race** data is collected multiple ways in Add Health; youth are asked their race at home and in school; are allowed to identify with multiple racial groups, but then asked to select a “best” monoracial group; are assigned a race by interviewers; can have their race surmised from the races reported by their biological parents; and are asked these questions three times over 7 years.

## 8. Are there valid and reliable measures of the variable?

- EX (continued): Only 88 percent of Add Health respondents give the same response to the school and home race questions. More than one racial group is selected by seven percent of adolescents at school, four percent at home, and eight percent in at least one of the two contexts. Only 67 percent of youth who identify as white and black, and then select white as the race that best describes them, are also identified as white by interviewers. By contrast, 95 percent of the white-black youth that most identify with blacks are identified as black by interviewers.

## 8. Are there valid and reliable measures of the variable?

- Before you can measure a variable you must define it. The measure should consequently reflect the definition.
- Among demographic variables, race, ethnicity, sex, and sexual orientation, are particularly hard to define and consequently hard to measure.
- See [GayData.org](http://GayData.org) for examples of sexual orientation variables.
- Have you ever seriously questioned the validity of any demographic data collected in your research?

## Ask yourself:

1. Is the variable relevant to the study?
2. How much will it cost to collect?
3. Will subjects refuse to answer questions or break-off interviews?
4. Could findings stigmatize populations?
5. Will there be enough power to analyze data?
6. Is the variable a proxy for more relevant variables?
7. Are there political ramifications of collecting this data?
8. Are there valid and reliable measures of the variable?