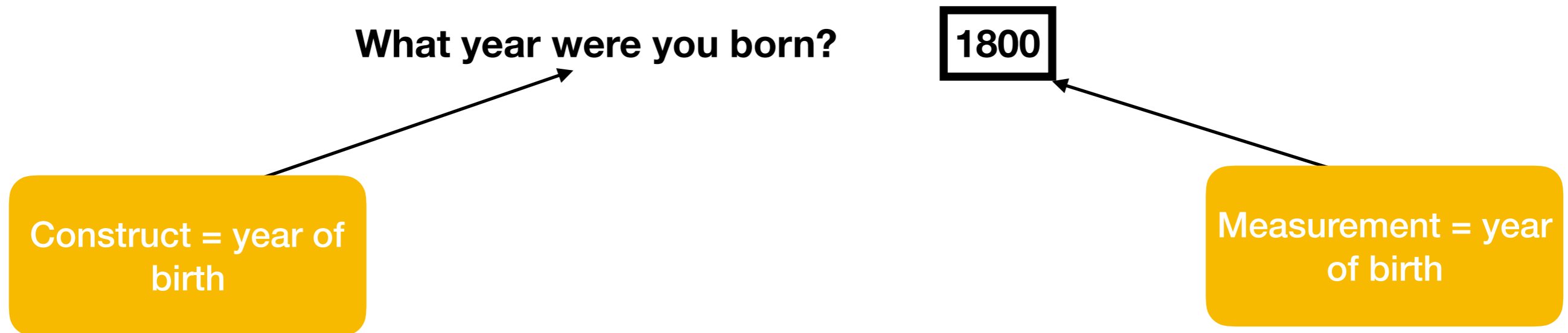


# Using gender and sex in survey statistics

Lauren Kennedy, Katharine Khanna,  
Daniel Simpson and Andrew Gelman

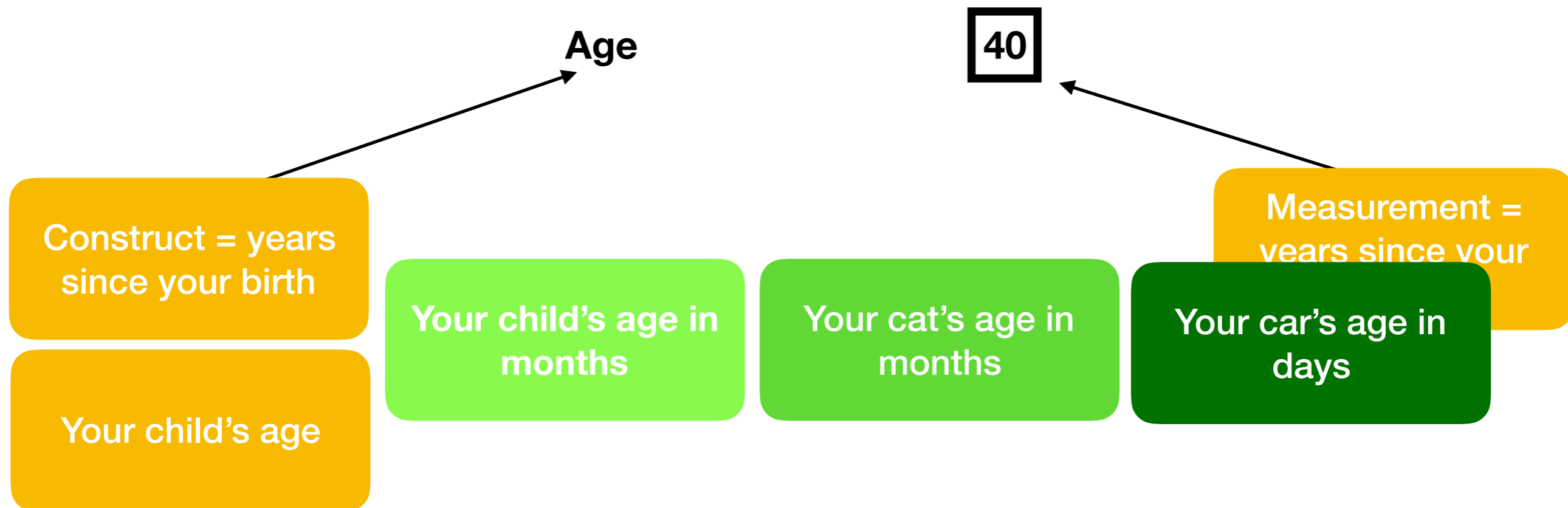
# Constructs in survey adjustment

- Constructs are the “things” we hope that our survey questions measure
- Sometimes the question makes it very clear what the construct is:



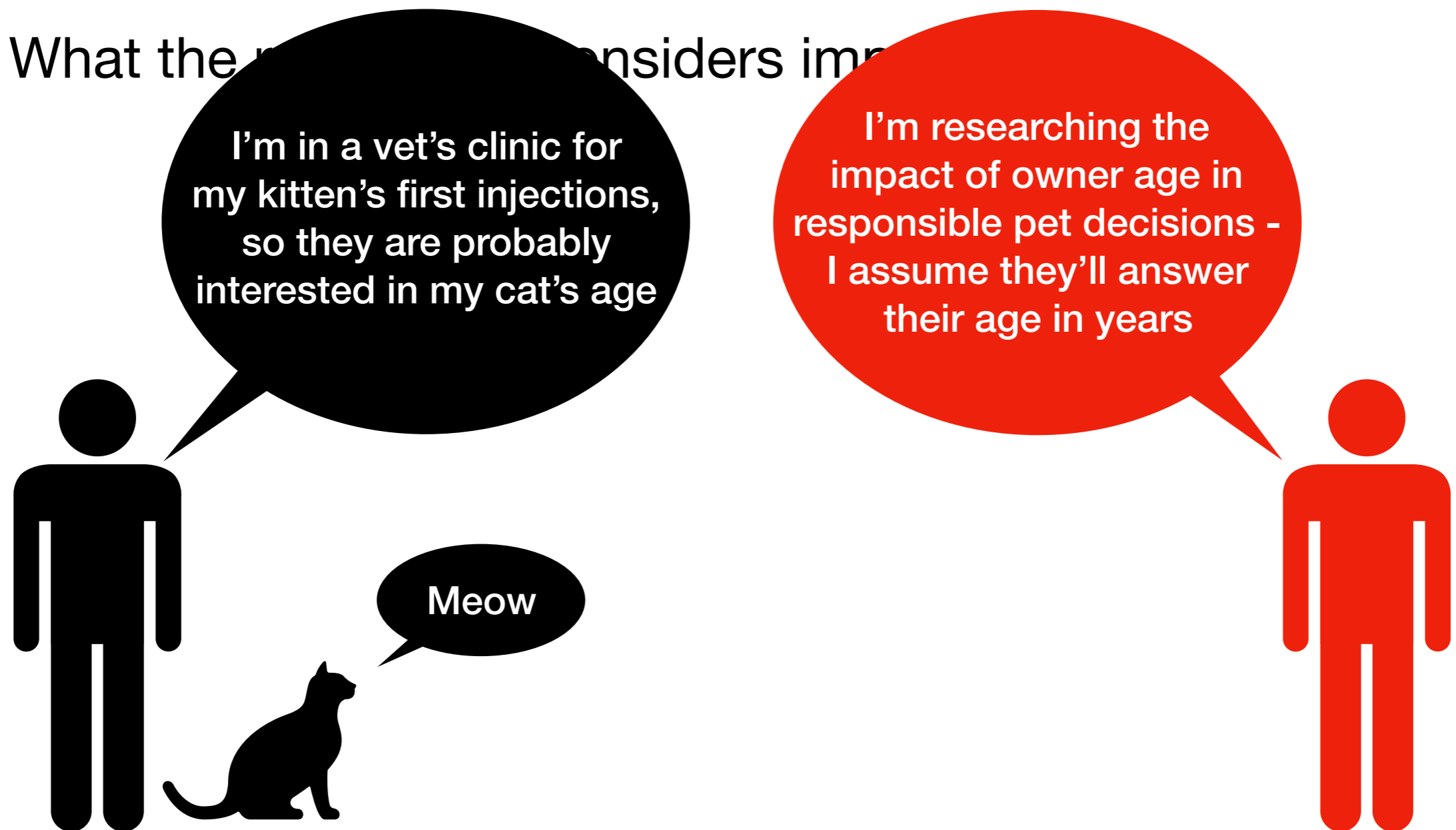
# Constructs in survey adjustment

- Constructs are the “things” we hope that our survey questions measure
- Sometimes the question relies on shared assumed knowledge of what the construct is:



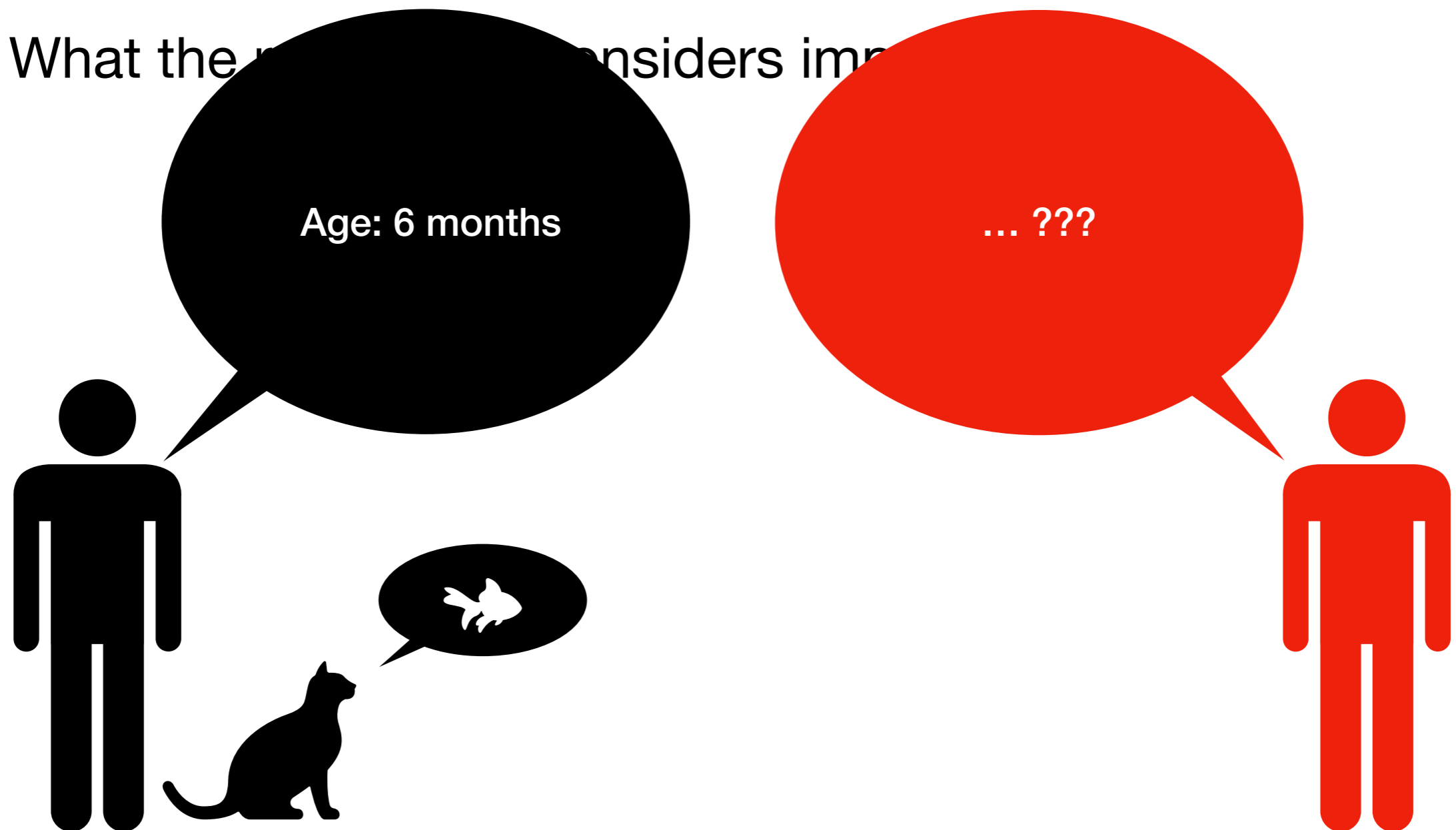
# Shared knowledge

- What the survey respondent considers important
- What the researcher considers important



# Shared knowledge

- What the survey respondent considers important
- What the respondent considers important



# Society

- What we view is important is determined by the society we live in, the values we have and our past experiences.
- All of these things are not static and can change over time and lives
- What might have been previously understood might now be more ambiguous

# What is being measured?

**Male**

**Female**

# What is being measured?

**Male**

**Female**

**Intersex**



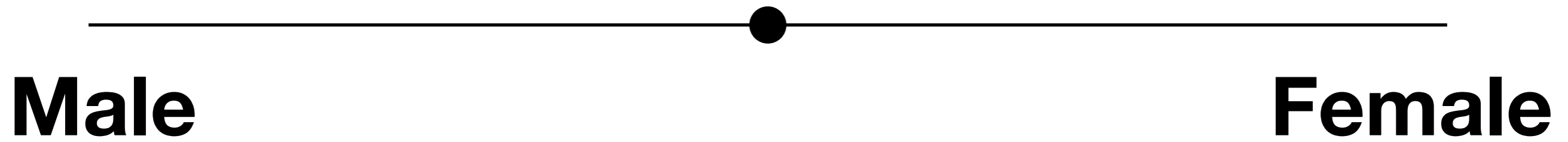
# What is being measured?

**Male**

**Female**

Enter text

# What is being measured?



# What is being measured?

**Male**

**Female**

# Definition

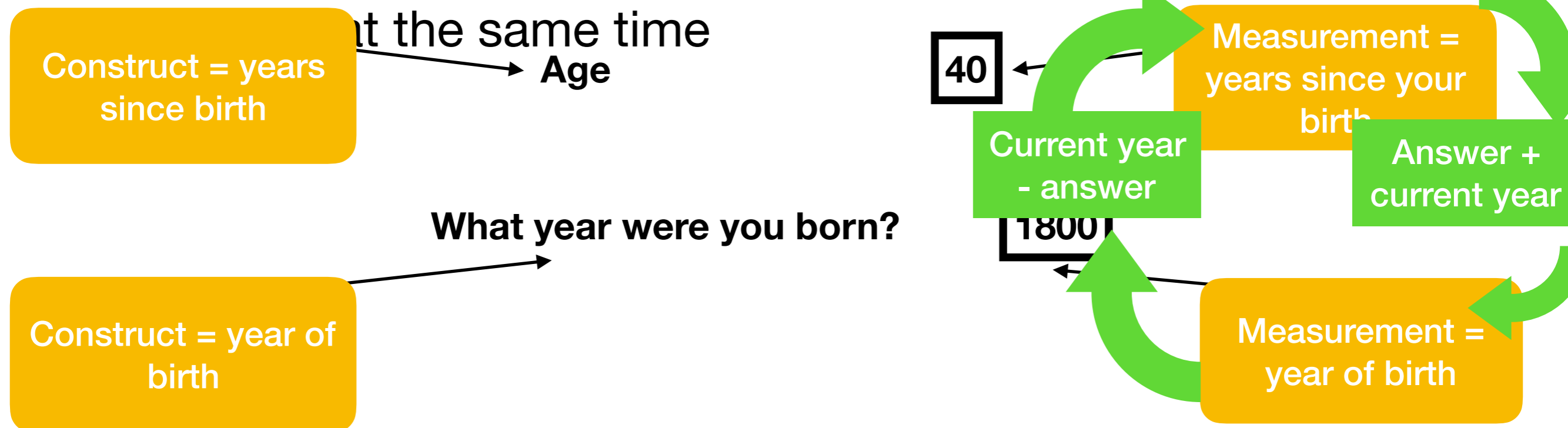
“Sex refers to a set of biological attributes in humans and animals. It is primarily associated with physical and physiological features including chromosomes, gene expression, hormone levels and function, and reproductive/sexual anatomy. . . .

Gender refers to the socially constructed roles, behaviours, expressions and identities of girls, women, boys, men, and gender diverse people. It influences how people perceive themselves and each other, how they act and interact, and the distribution of power and resources in society. . . .”

Canadian Institutes of Health Research

# Mapping constructs

- (surprising no one) Not all surveys measure the same thing
- Many surveys measure a set of common constructs - demographics like age, education level, etc
- Not all surveys do this in the same way, with the same



# Why is this important?

- Post-stratification based methods:
  - Post sampling adjustment of sampling weights to population totals
  - Aim to match key population demographics (i.e., construct to construct)
  - Useful to adjust for demographic differences due to higher than desired non-response
  - Adjust for non-probability samples

# What should we adjust for?

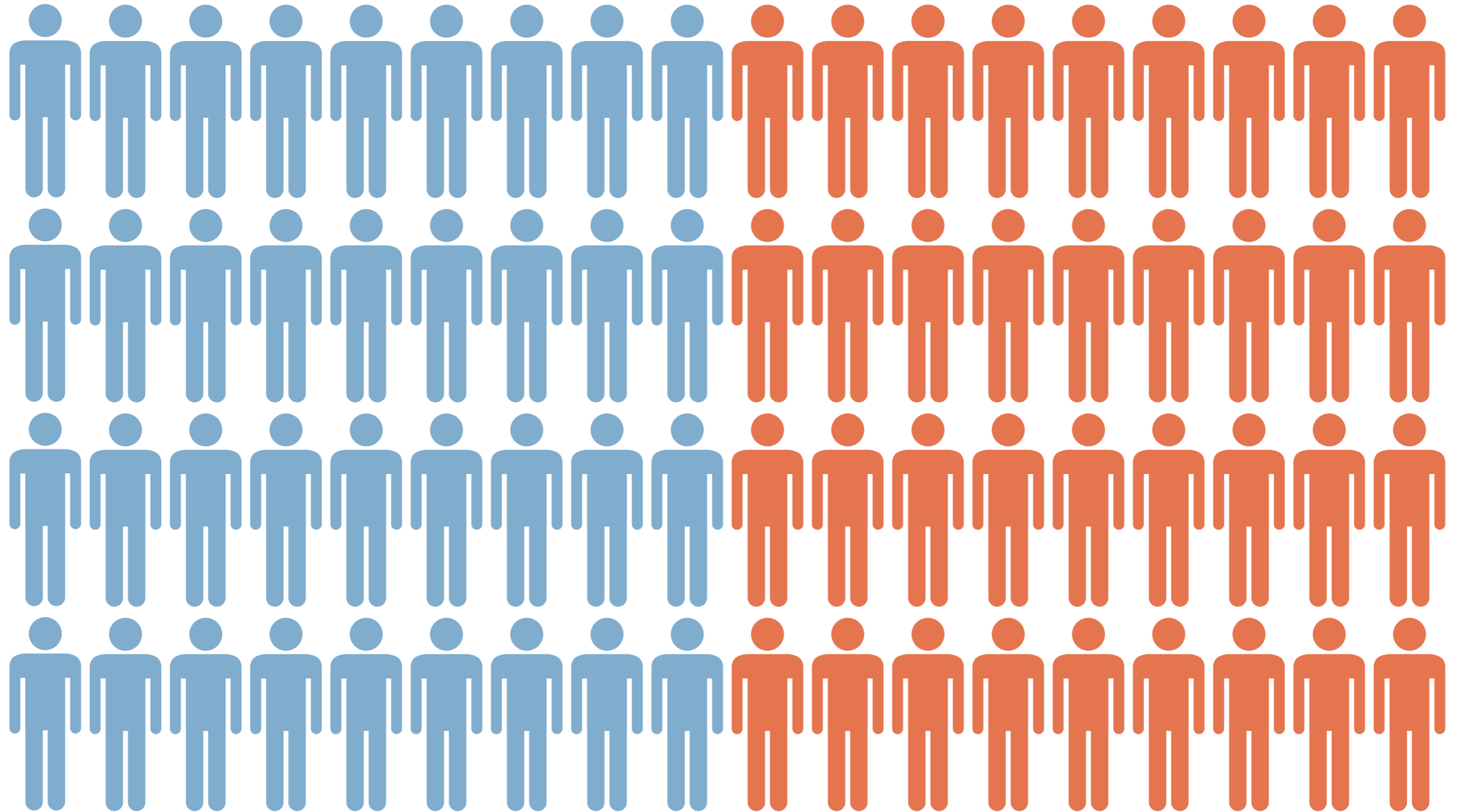
- Historical/Sociological reasons
- Predictive of outcome
- Predictive of inclusion

# Sample adjustment

- Adjust a non-representative sample to known population demographics.
- ALL post stratification methods (weighting or MRP) rely on mapping sample characteristics to population characteristics.
- (relatively) easy when we mirror the demographic questions in the census in our surveys

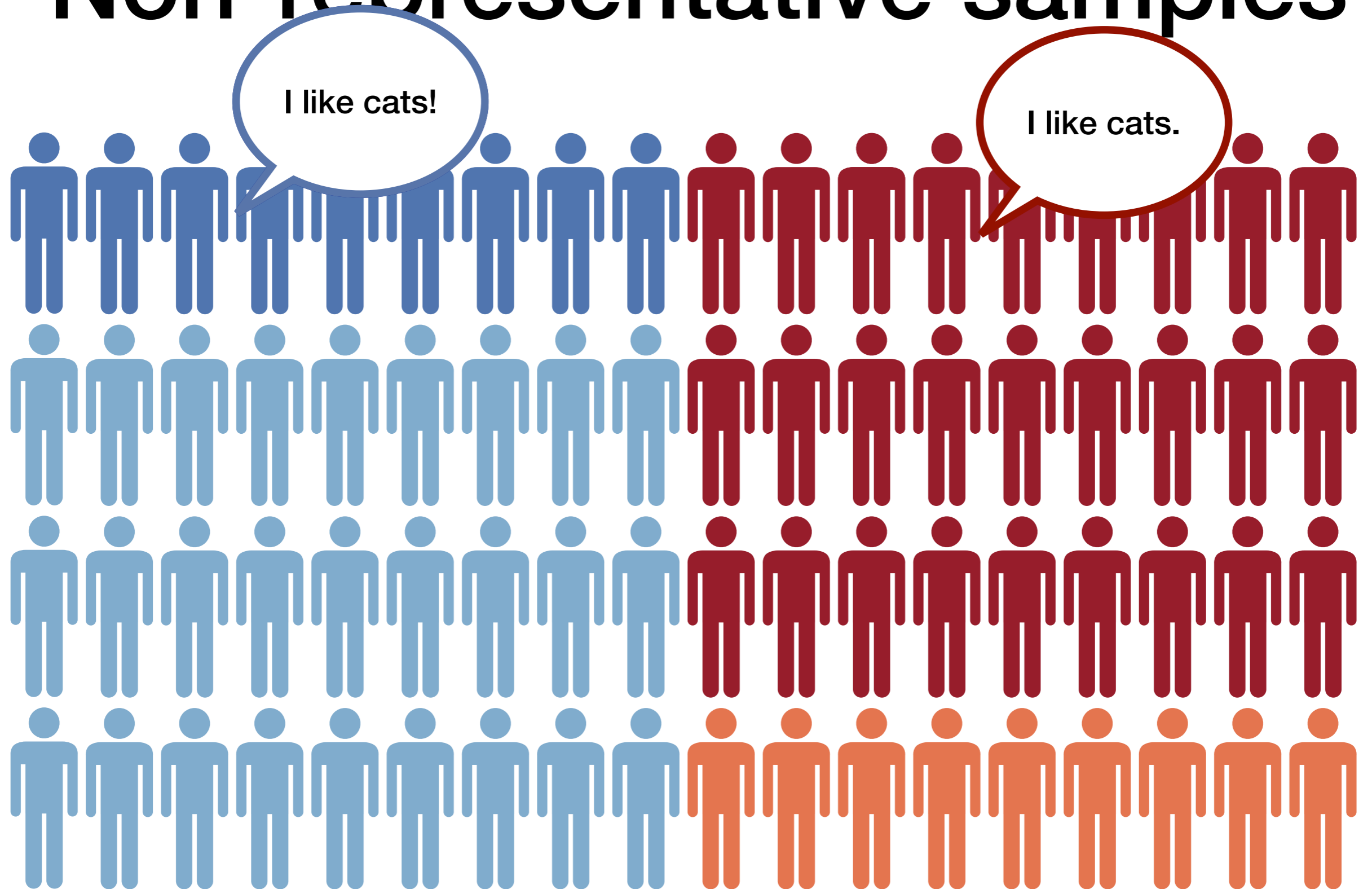


# Non-representative samples

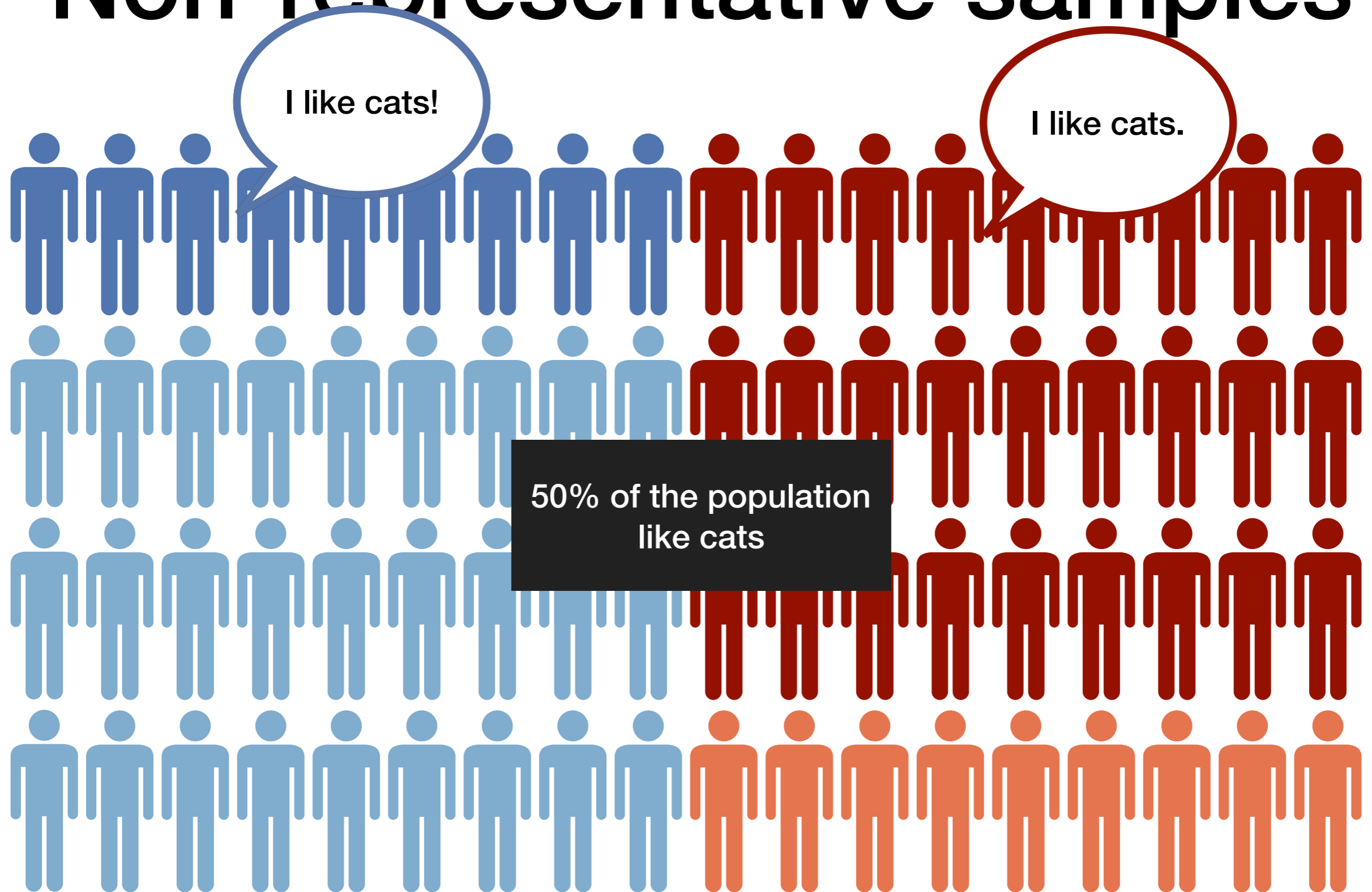




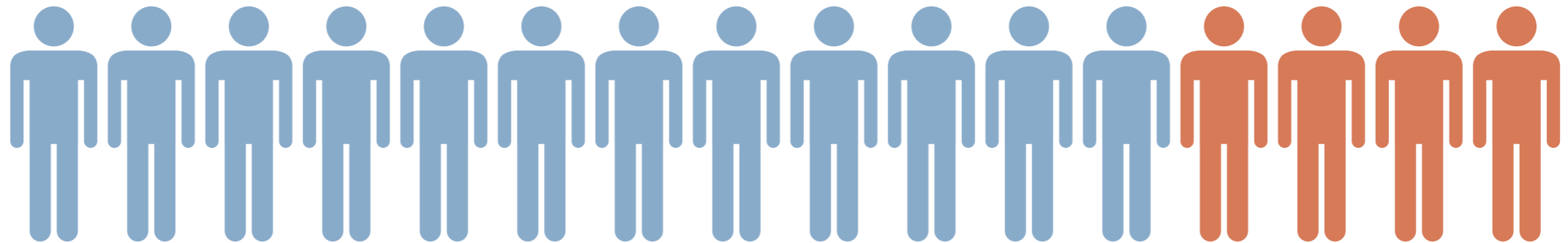
# Non-representative samples



# Non-representative samples



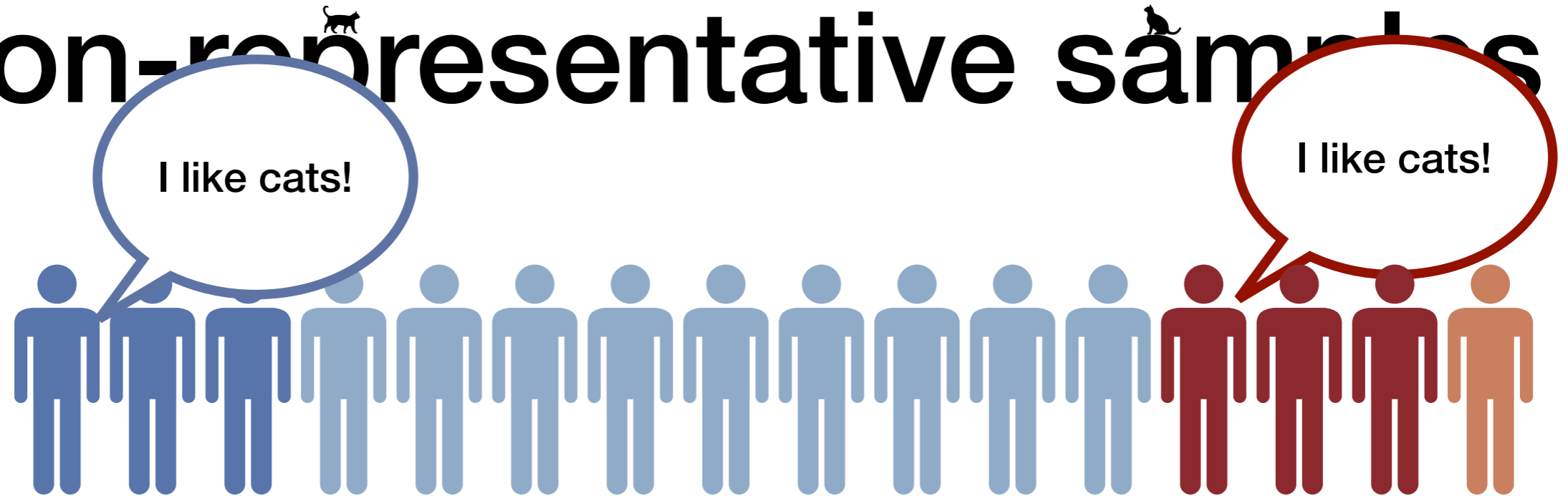
# Non-representative samples



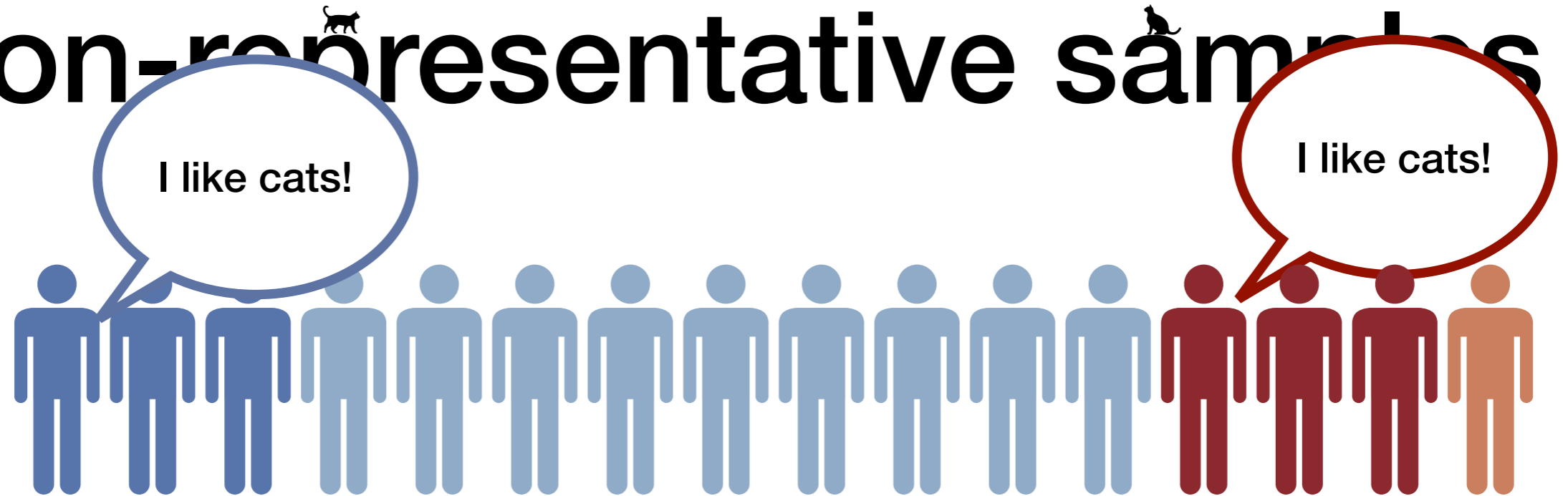
# Non-representative samples



# Non-representative samples



# Non-representative samples

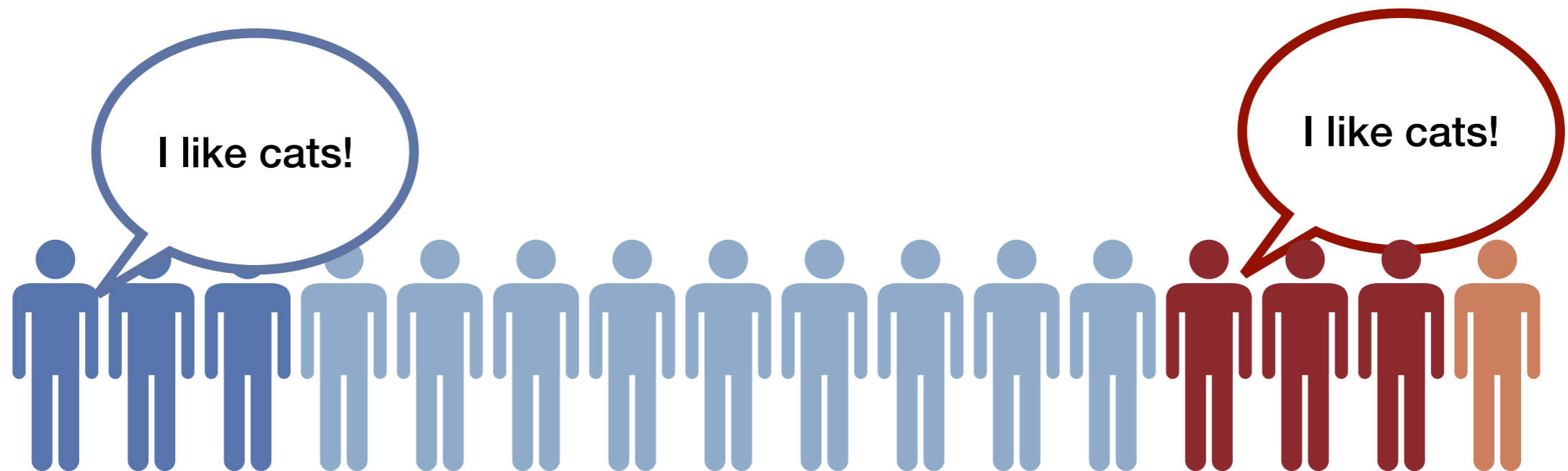


37.5% of the sample  
like cats



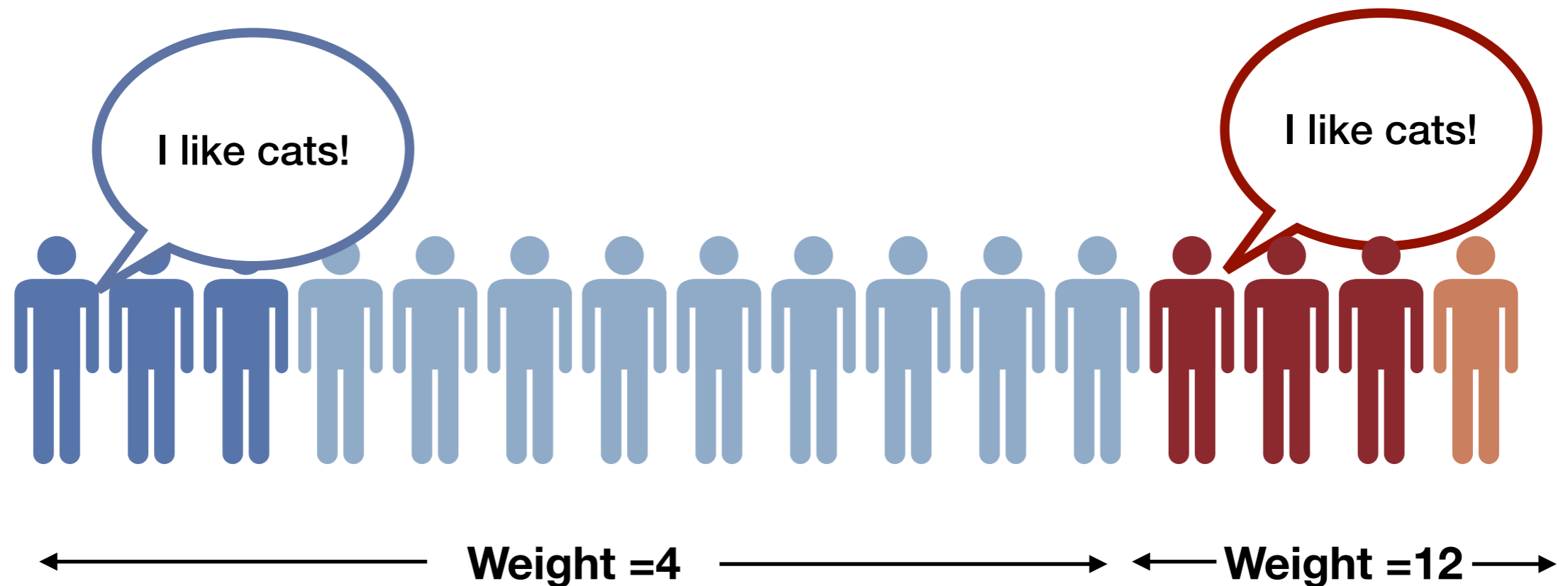
# Population statistics from non-representative sample

- Traditionally a vector of weights



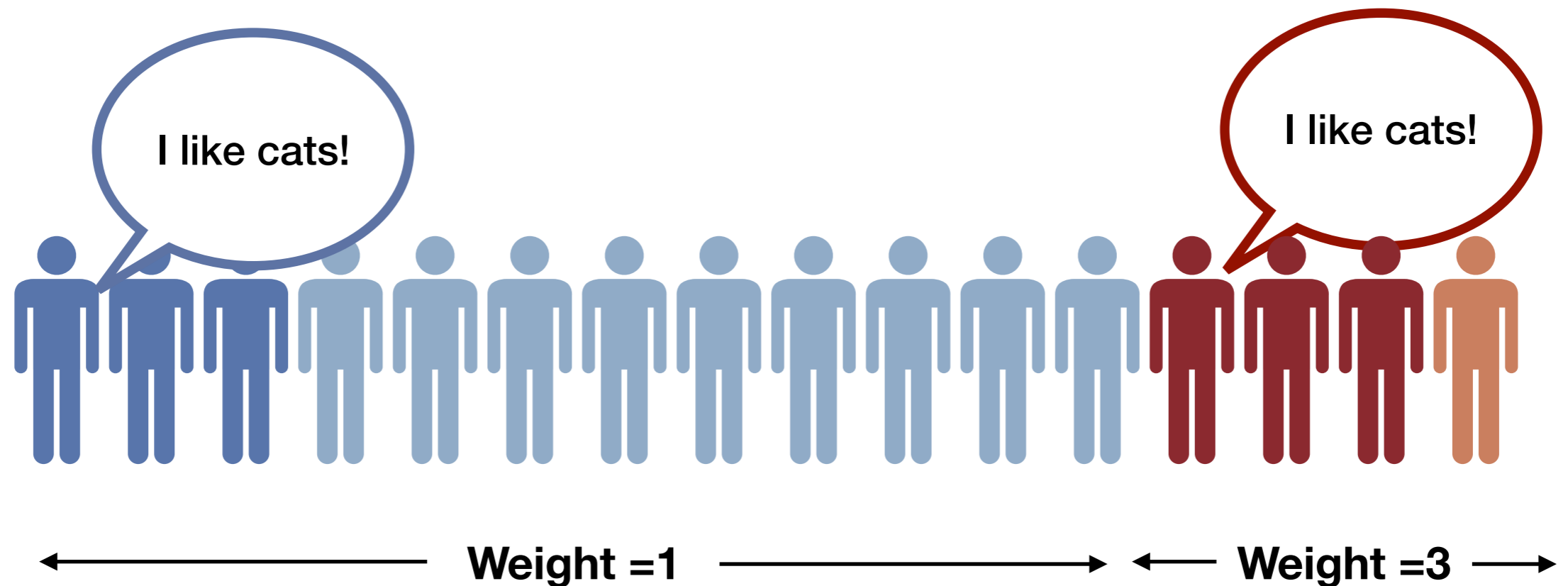
# Population statistics from non-representative sample

- Traditionally a vector of weights



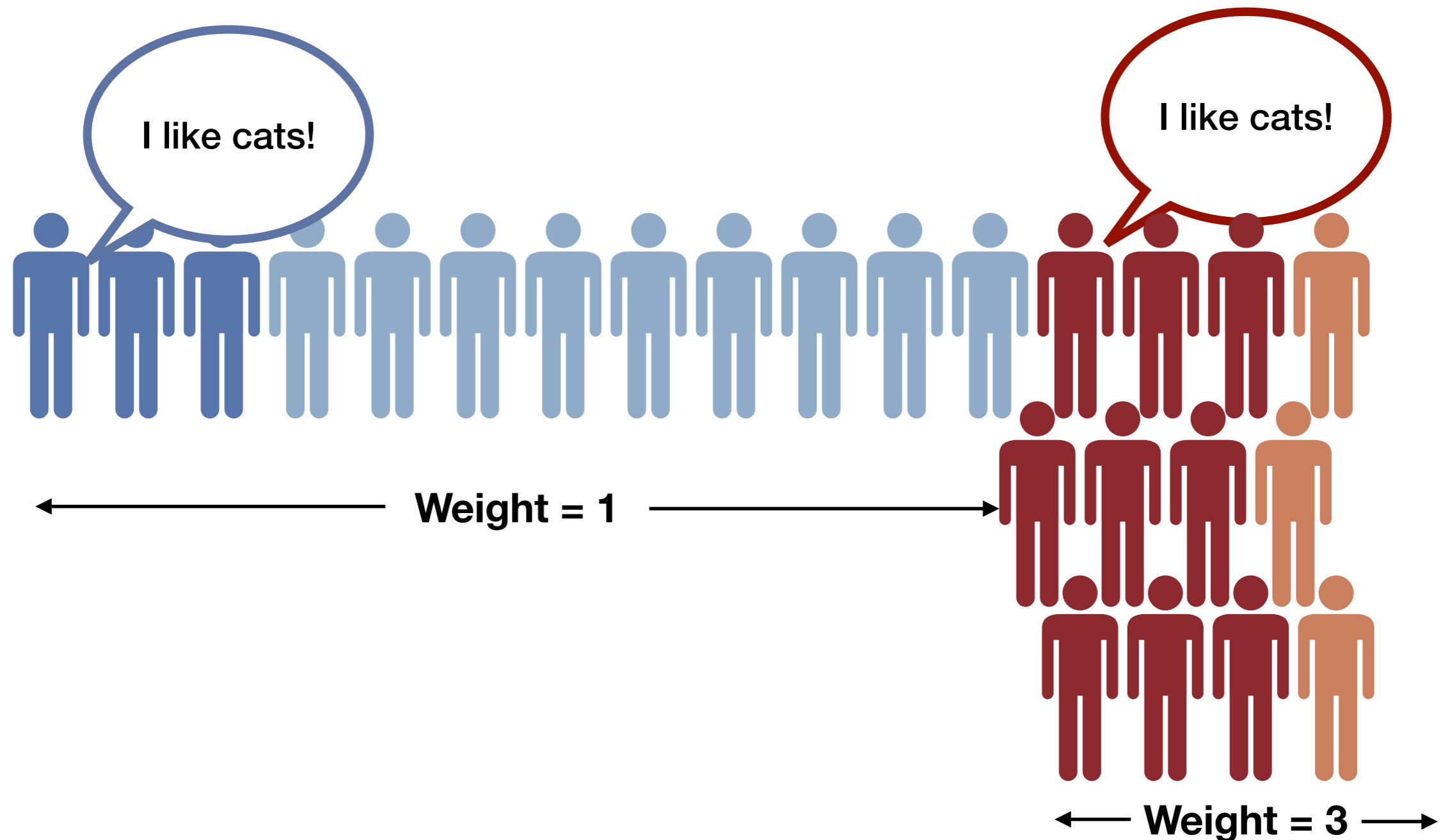
# Population statistics from non-representative sample

- Traditionally a vector of weights



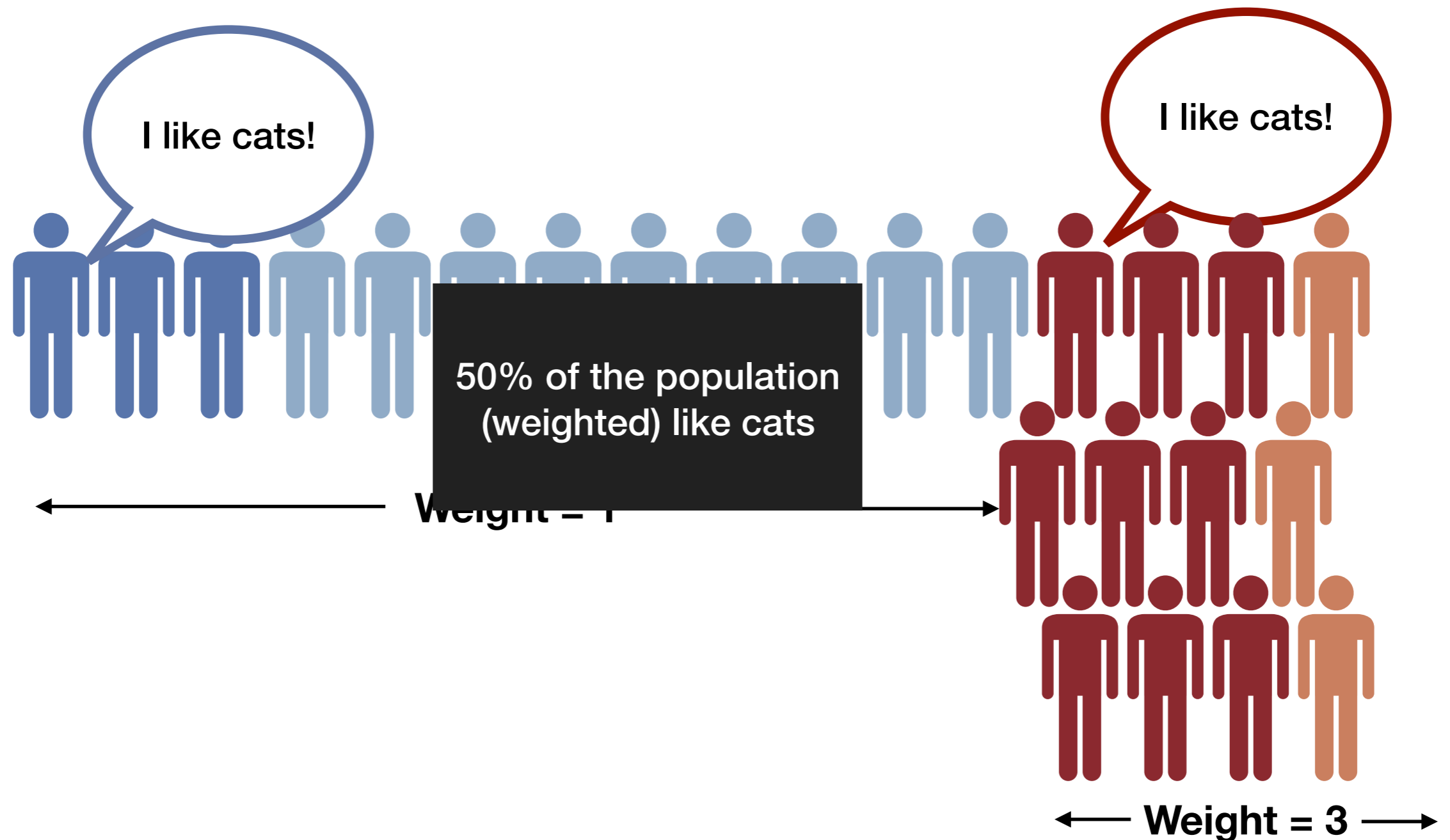
# Population statistics from non-representative sample

- Traditionally a vector of weights



# Population statistics from non-representative sample

- Traditionally a vector of weights



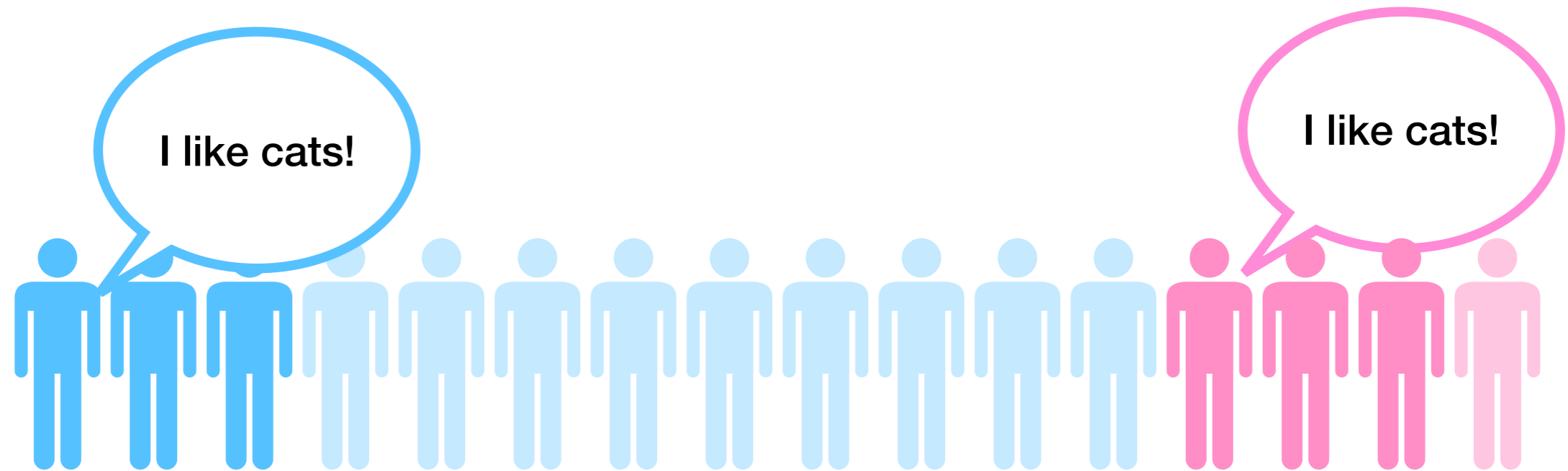
**Simple right?**

# Simple right?

- What if we don't measure the exact same thing?

# Simple right?

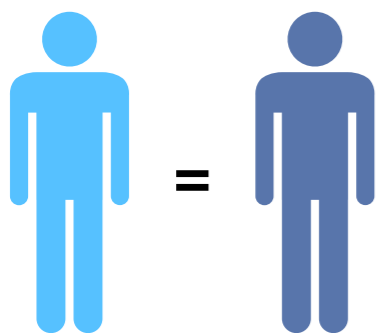
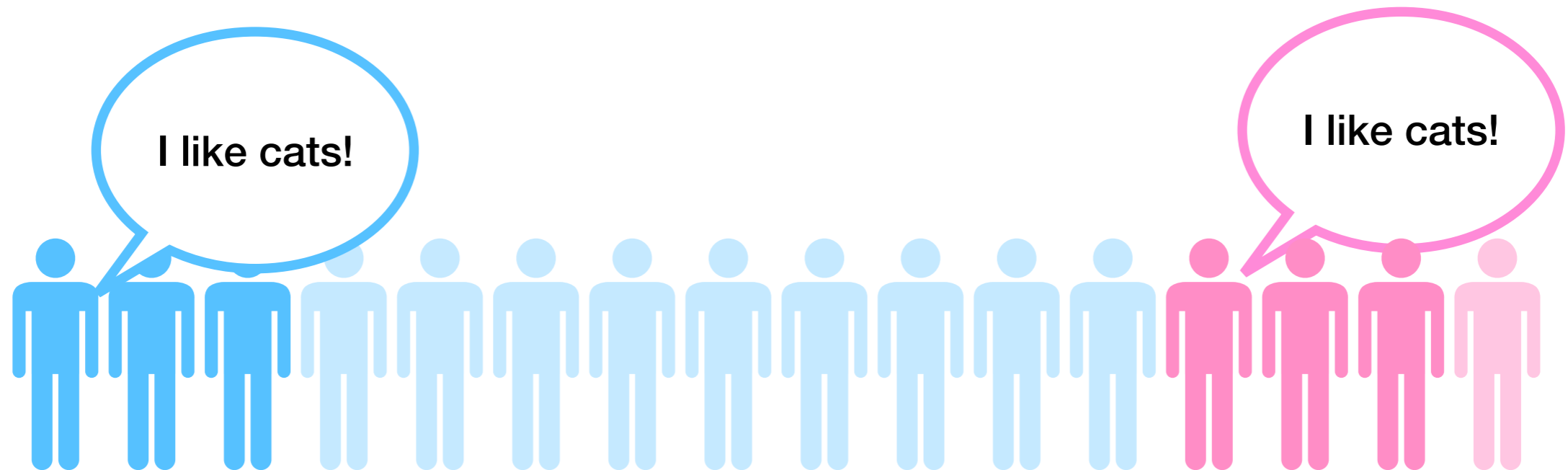
- What if we don't measure the exact same thing?





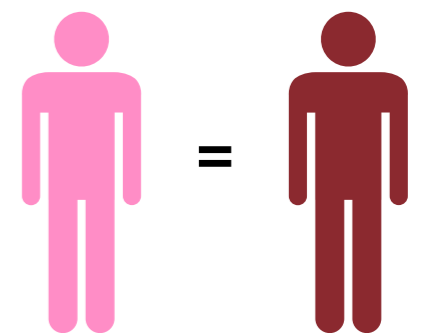
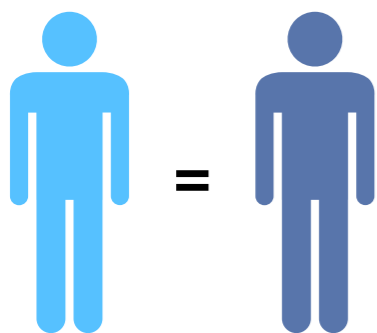
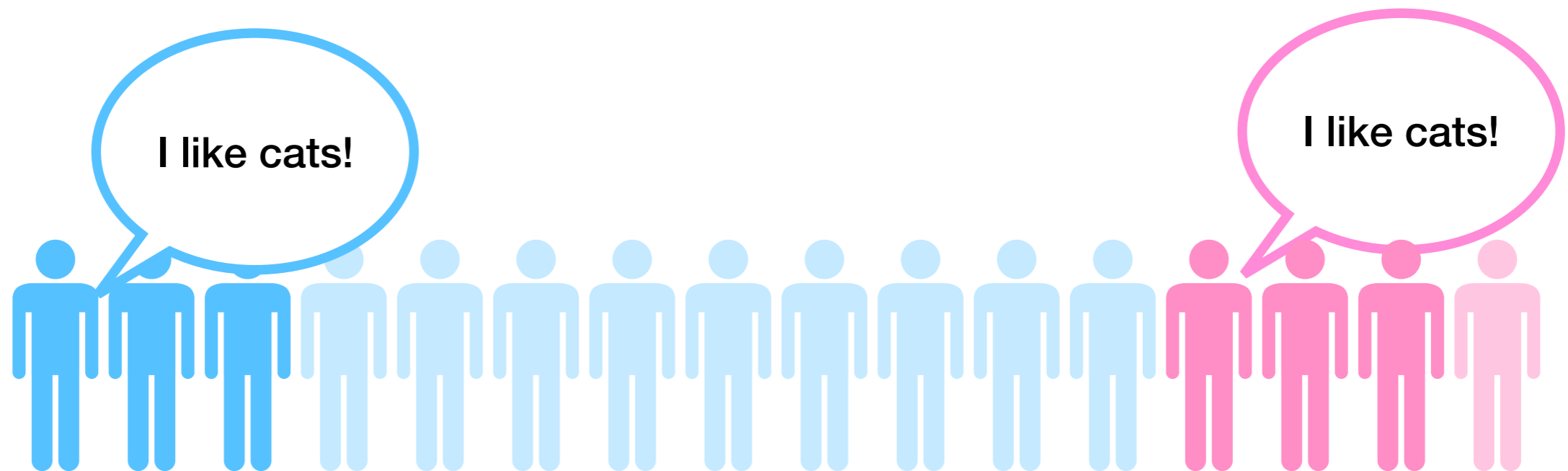
# Simple right?

- What if we don't measure the exact same thing?



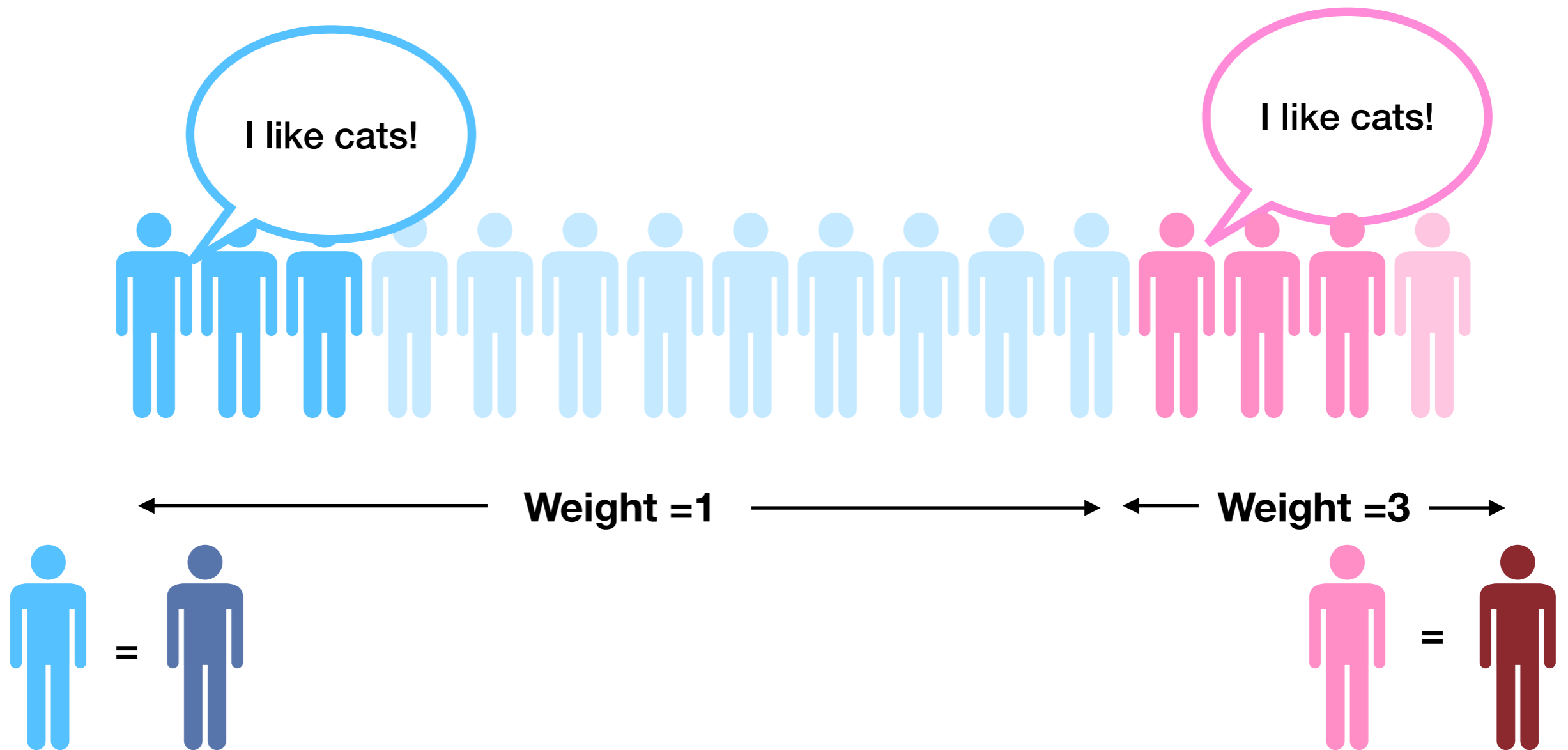
# Simple right?

- What if we don't measure the exact same thing?



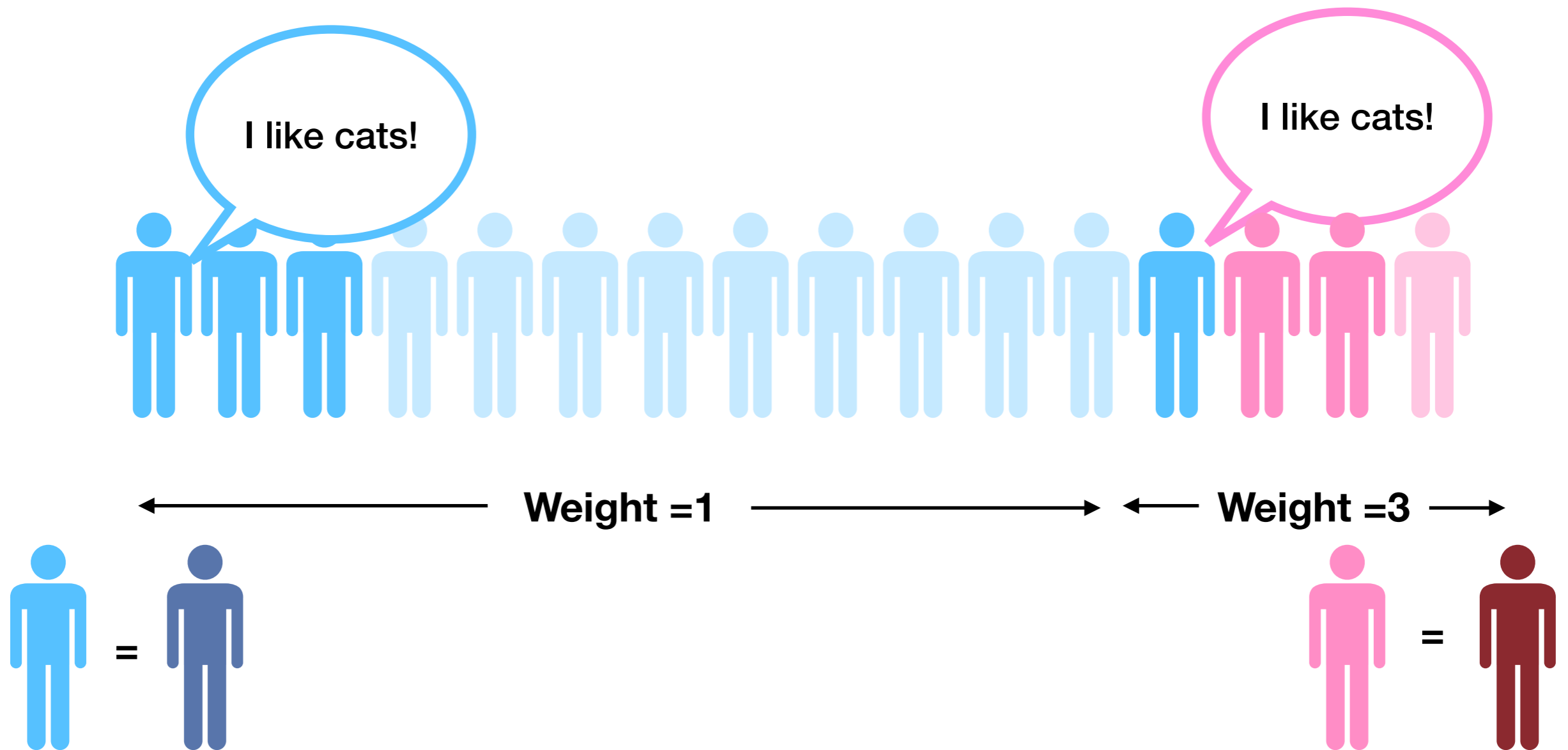
# Simple right?

- What if we don't measure the exact same thing?



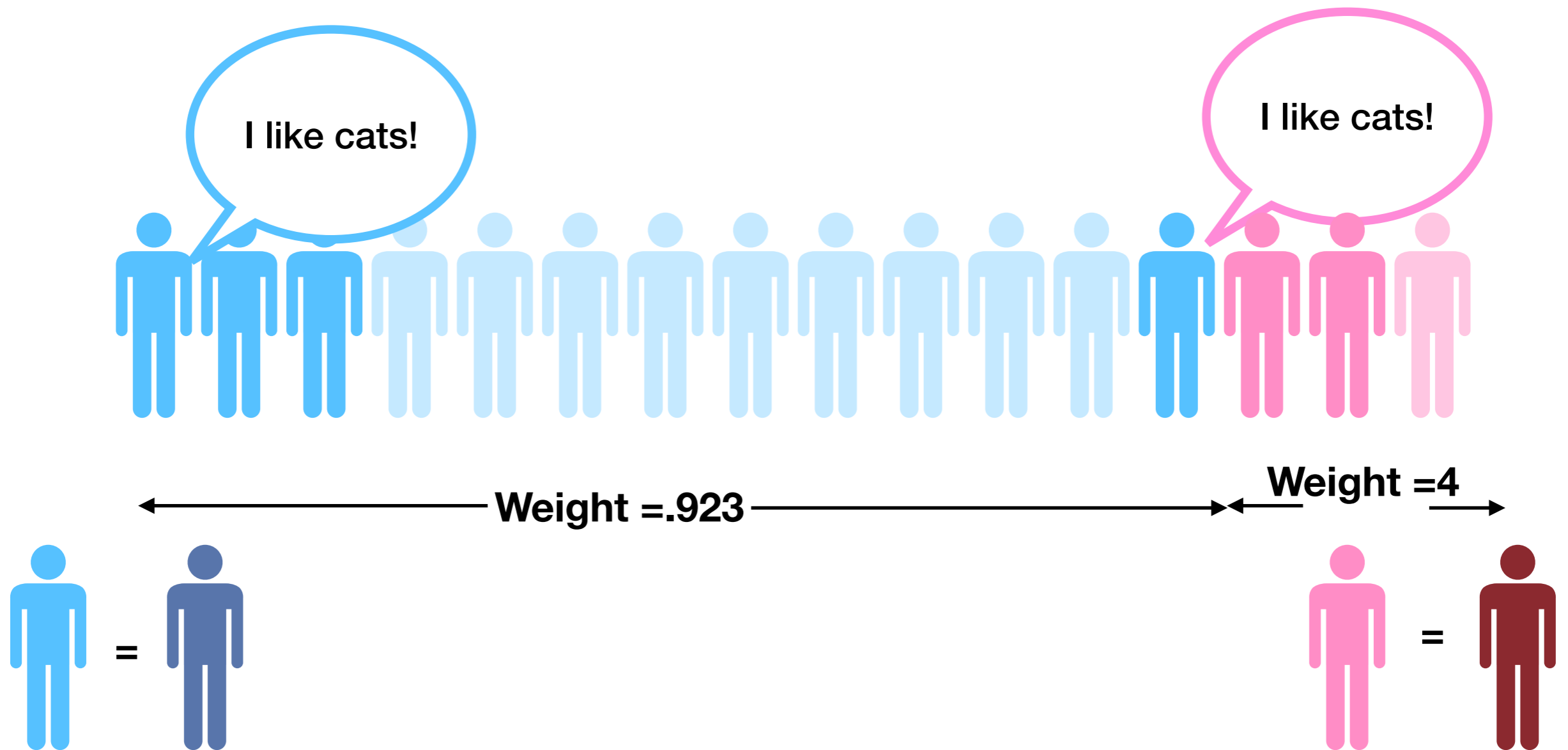
# Simple right?

- Potential construct error leads to errors in weights



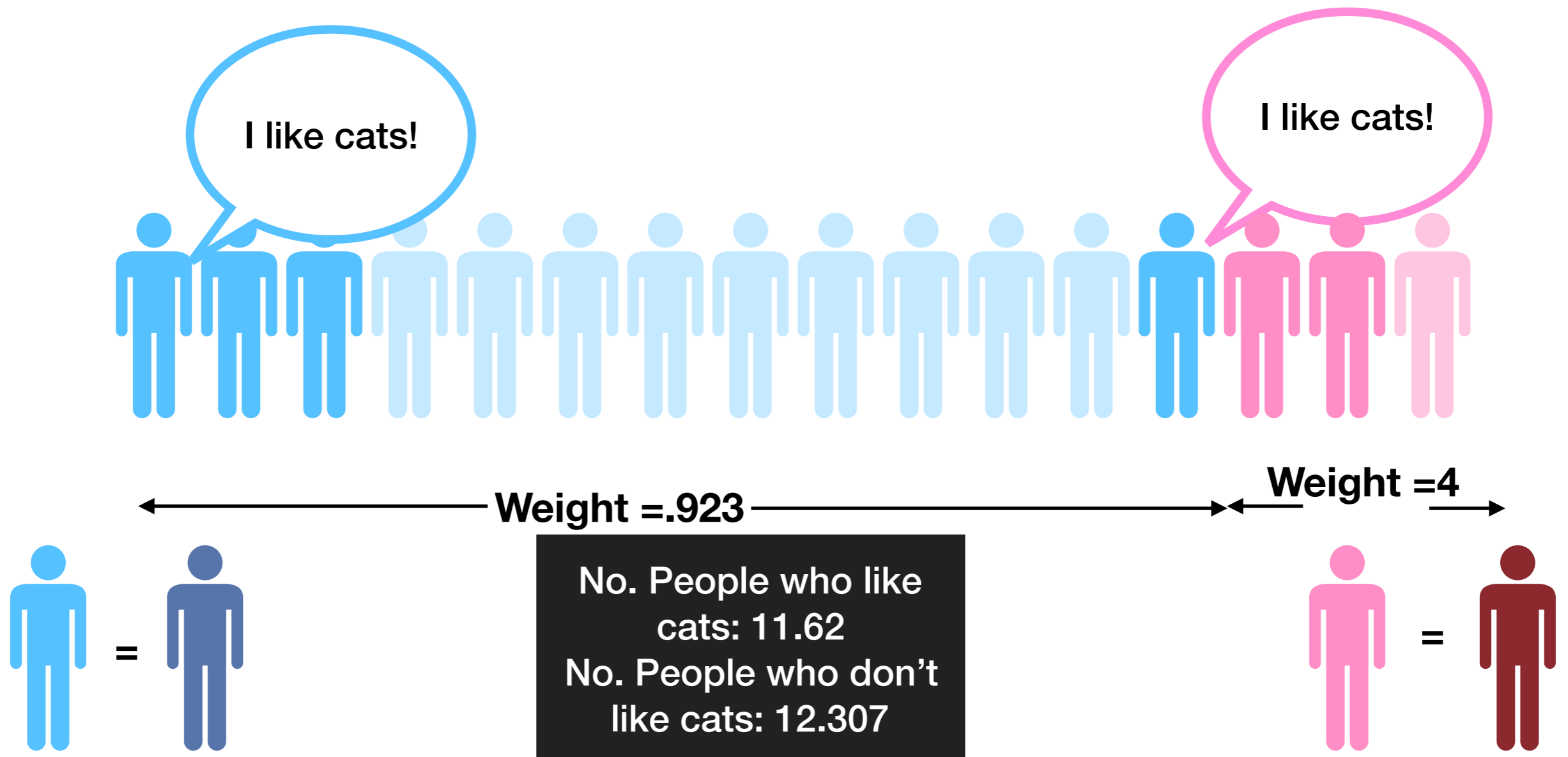
# Simple right?

- Potential construct error leads to errors in weights



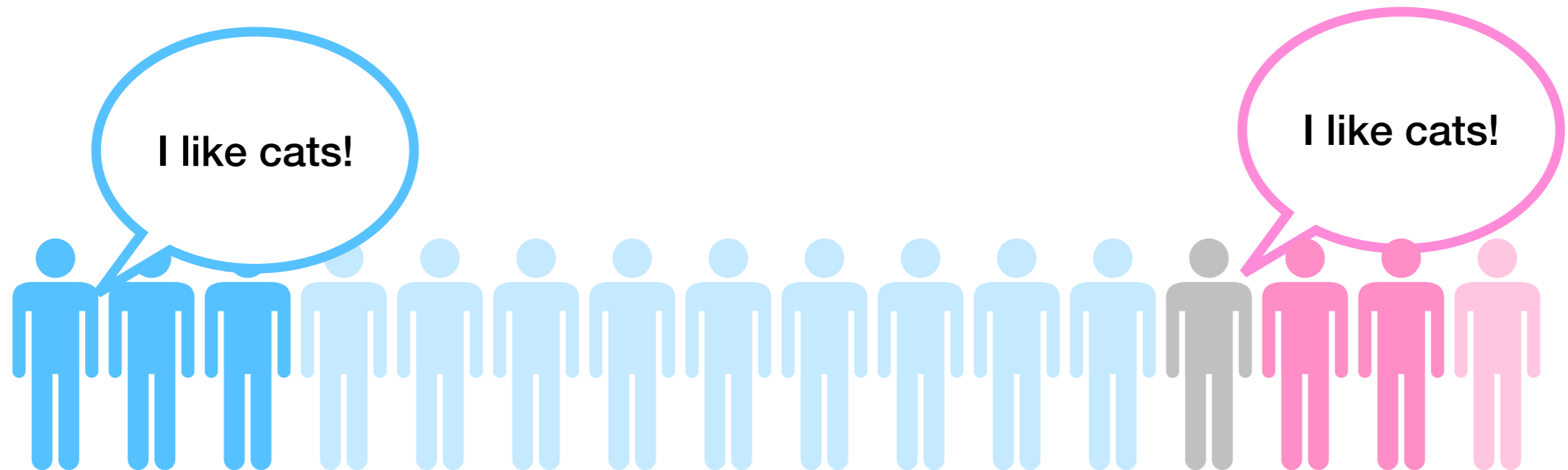
# Simple right?

- Potential construct error leads to errors in weights



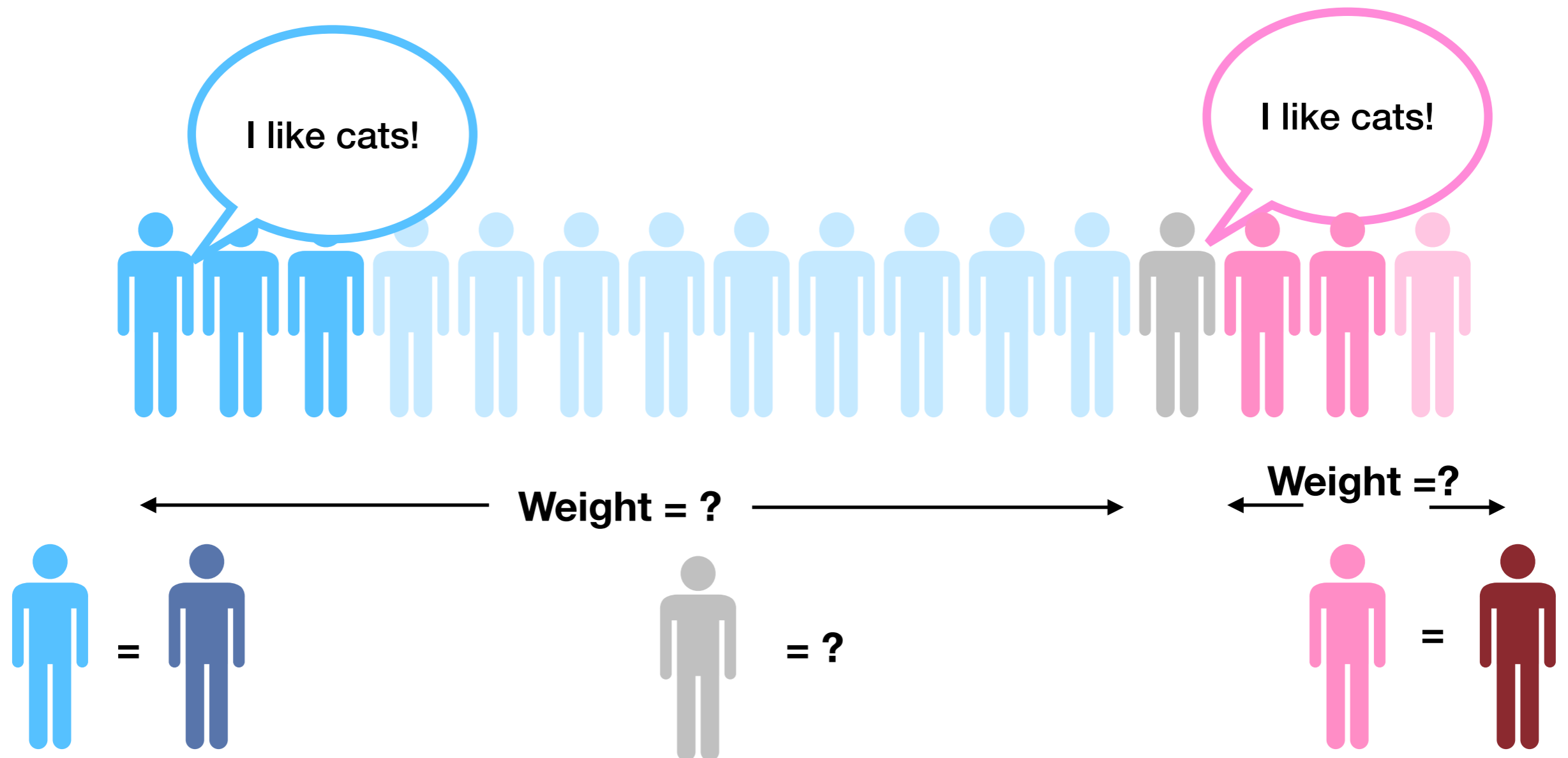
# Simple right?

- Different measurement categories highlight this issue



# Simple right?

- Different measurement categories highlight this issue





# Are we measuring sex or gender?

- Sex is measured in the US census, ACS and other official statistics in the US
- Social surveys have increasingly asked for gender but with two options (M/F)
- Recent improvements in gender measurement has led to the addition of (at least one) additional category. We refer to those who choose this category as non-binary.

We use non-binary as a category name for those who identify as non-binary, agender, gender fluid, and other gender identities outside of female and male. Although other is commonly used, we specifically do not use this term to avoid othering those who do not identify as male or female.

# What about outside the US?

- The 2016 Census counted 1,260 sex and/or gender diverse people in Australia.
- This count is not considered to be an accurate count, due to limitations around the special procedures and willingness or opportunity to report as sex and/or gender diverse. People who have been treated with disrespect, abuse and discrimination because of their sex or gender may be unwilling to reveal their sex in an official document.

- It was the first Australian Census to have a response option available on the online form for sex other than male or female, via a special online form with an 'Other' response option to the Sex question.

- Respondents generally needed to take extra steps and have knowledge of special procedures to report their sex as other than male or female. An 'Other' response category was not included as a standard response option because of uncertainty about how the general public would respond to a question with an explicit 'other' option. The reaction of the general public was explored through the pilot test

<https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/>

2071.0~2016~Main%20Features~Sex%20and%20Gender%20Diversity%20in%20the%202016%20Census~100

# What about outside the US?

## What is your sex?

A question about gender identity will follow later on in the questionnaire

Female

Male

Figure 2: Online question on gender identity in English and Welsh

Is the gender you identify with the same as your sex registered at birth?

This question is voluntary

Yes

No

Enter gender identity

<https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/sexandgenderidentityquestiondevelopmentforcensus2021#questions-recommended-for-census-2021>

\* Proposed questions for the 2021 census

# What about outside the US?

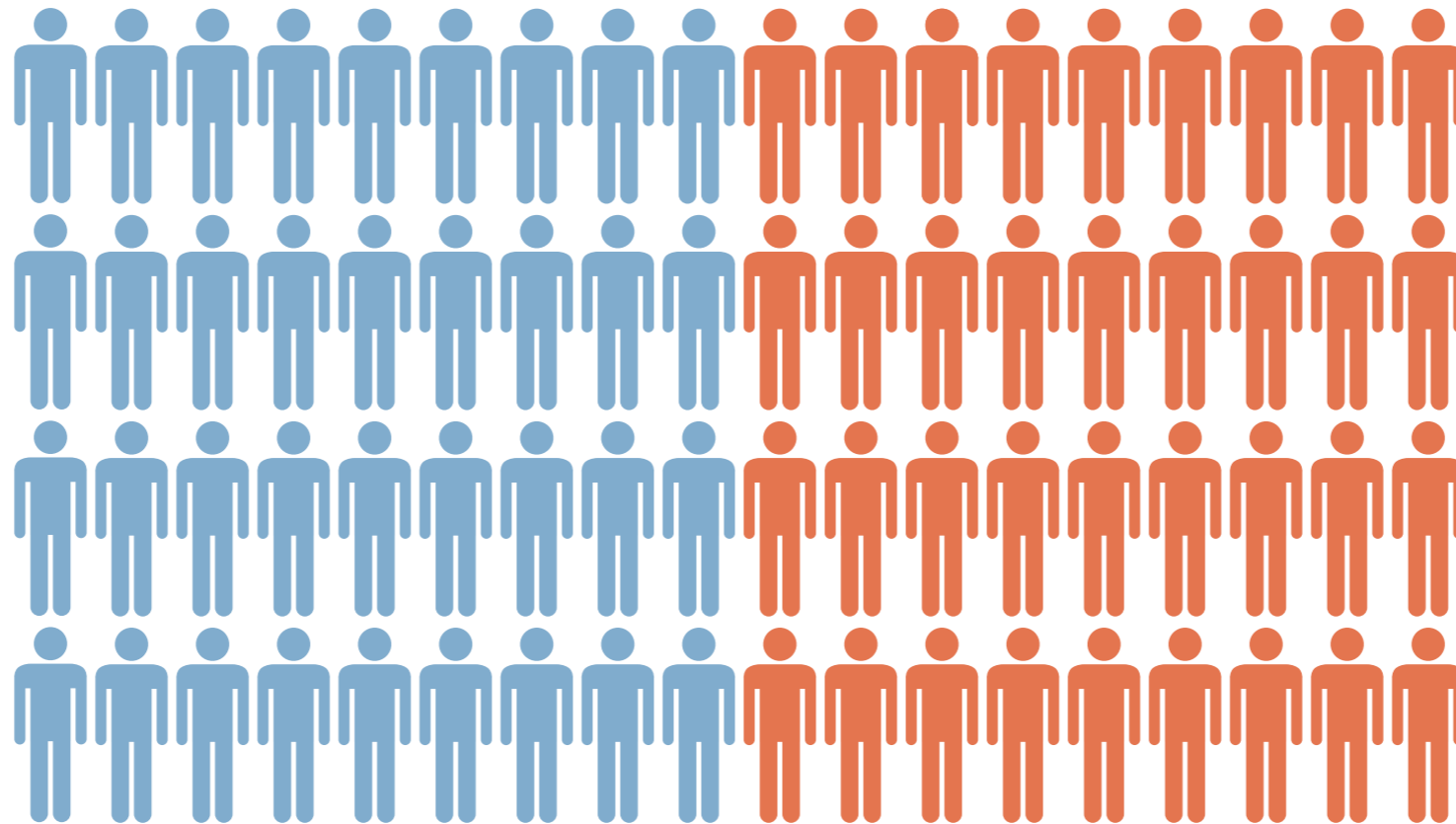
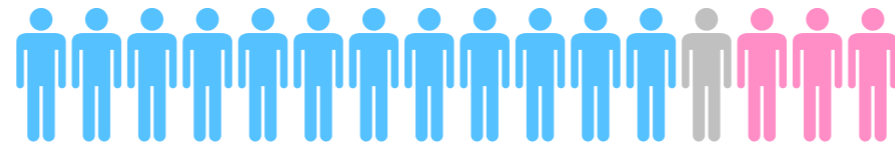
Consultation respondents and federal stakeholders suggested that the census questionnaire collect information on **sex at birth** and **gender**.

As an early action in response to the demands from users and through consultations, Statistics Canada developed and released new standards on sex and gender.

# Adjusting by sex/gender

- Sociological concerns
  - Historic underrepresentation
  - Current response patterns
  - Ensure those with low societal power are represented
- Statistical concerns
  - Relationship between sex/gender and many outcomes

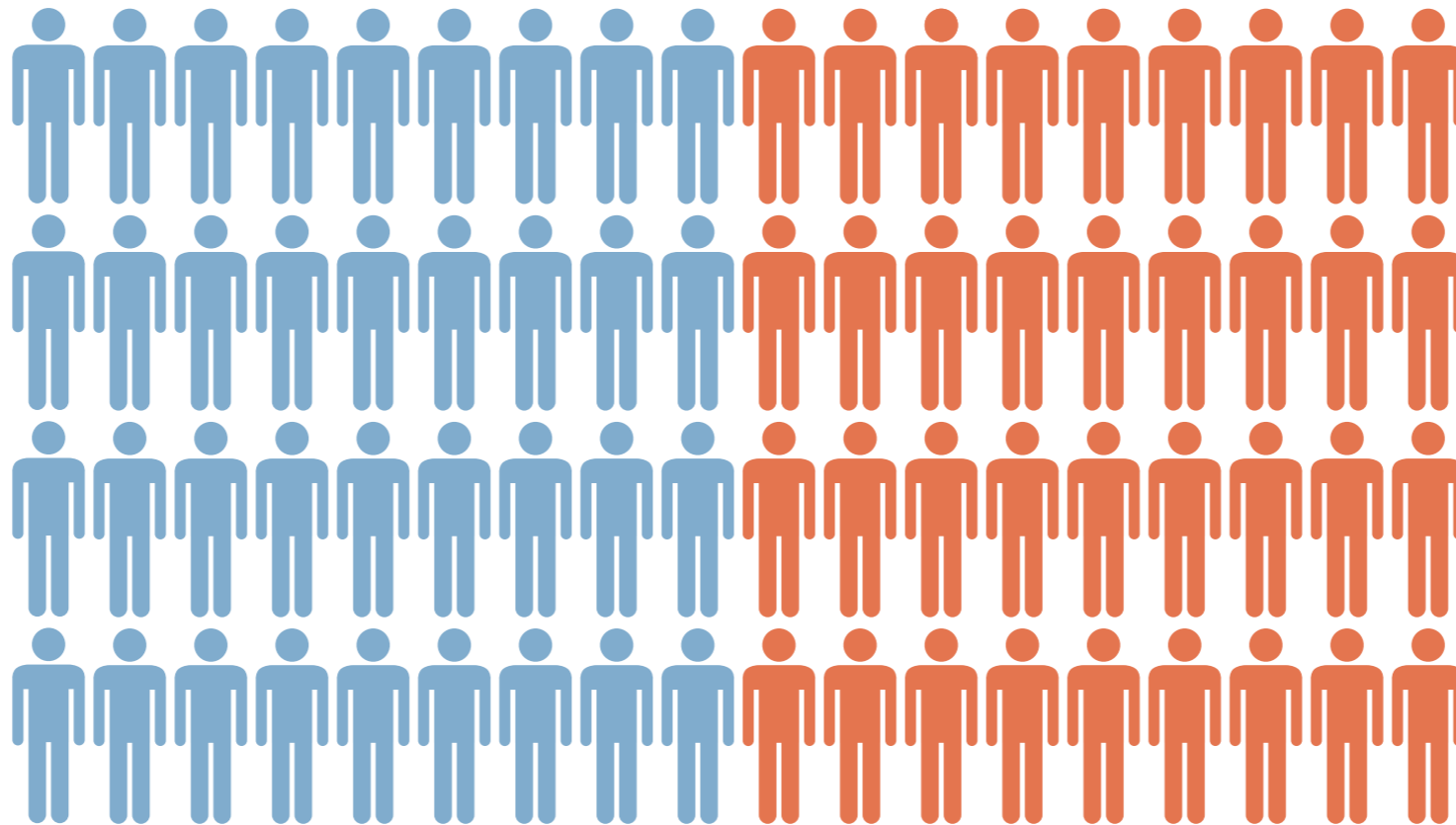
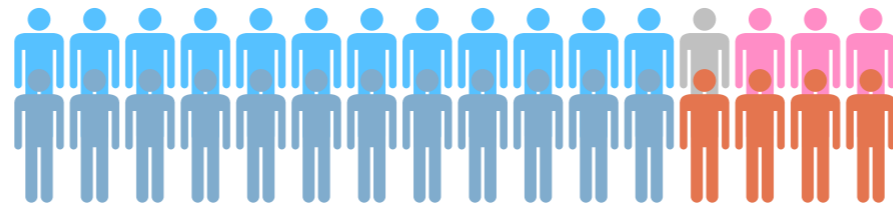
# Options



# Options

- 1. Impute sample sex assuming a known population counts**

# Options





# Benefits

- Easily explained to other researchers and respondents
- Backwards compatible with previous data collection
- Few infrastructure changes

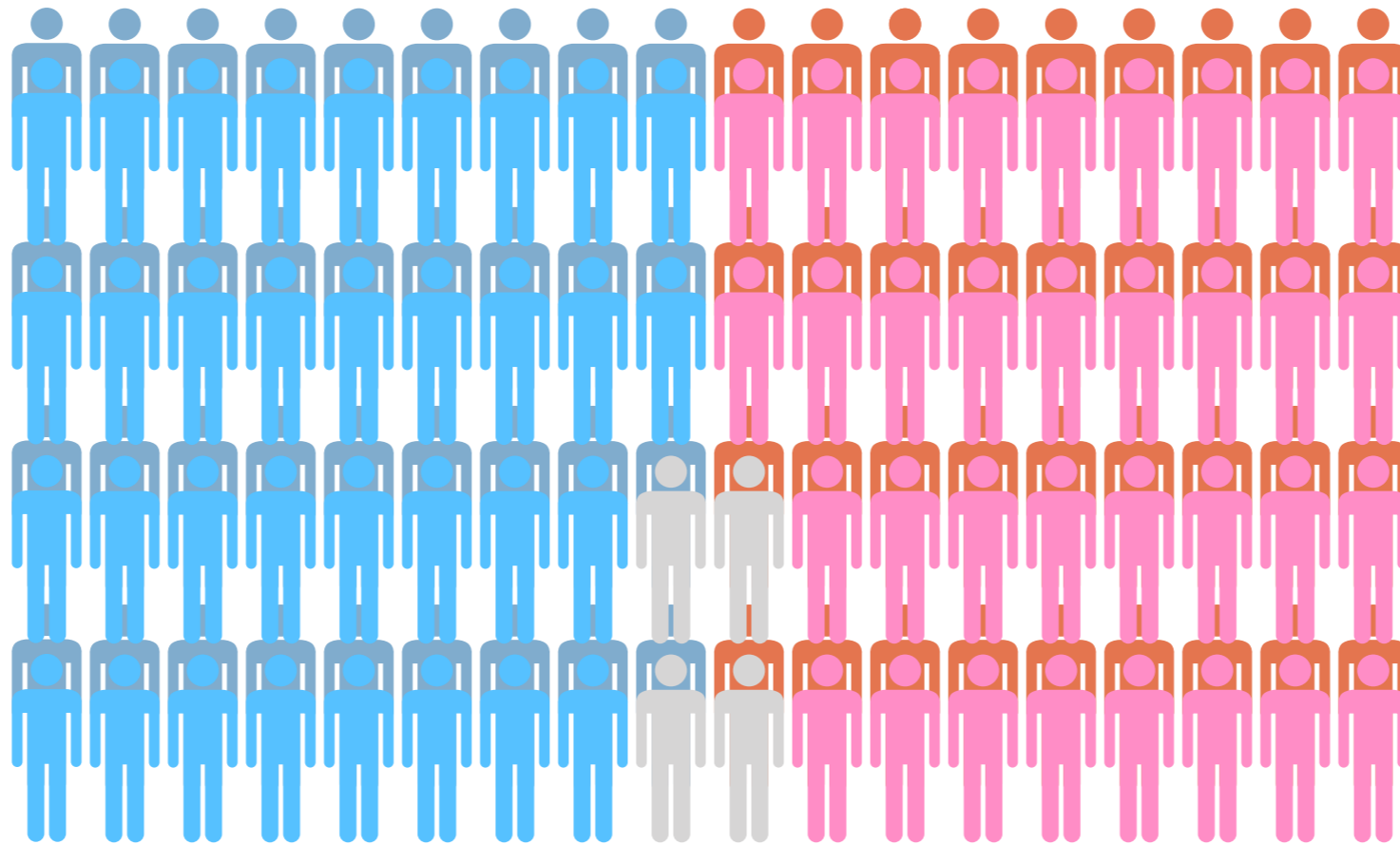
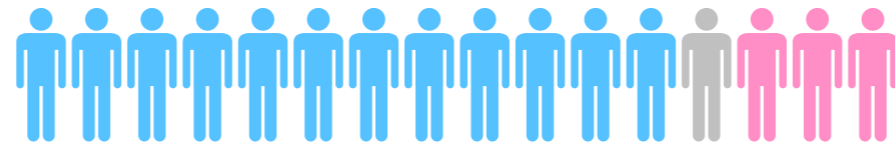
# Concerns and challenges

- Impute male gender as male sex and female gender as female sex.
- All involve imputing non-binary individuals as either male or female sex.
- Although we can be clear as statisticians we're imputing a potential response, the confusion between sex and gender makes this practice appear as statistical misgendering.
- Imputation error is not the same for all genders
- Uncertainty difficult to incorporate

# Options

1. Impute sample sex assuming a known population counts
- 2. Impute the population using a known population counts**

# Options



# Benefits

- Easily explained to other researchers and respondents
- Avoids touching the original survey data
- Avoids imputing individual cases

# Concerns and challenges

Scenario 1: Large population, large cells

- Doesn't involve imputing individual gender, rather splitting cell count by proportion

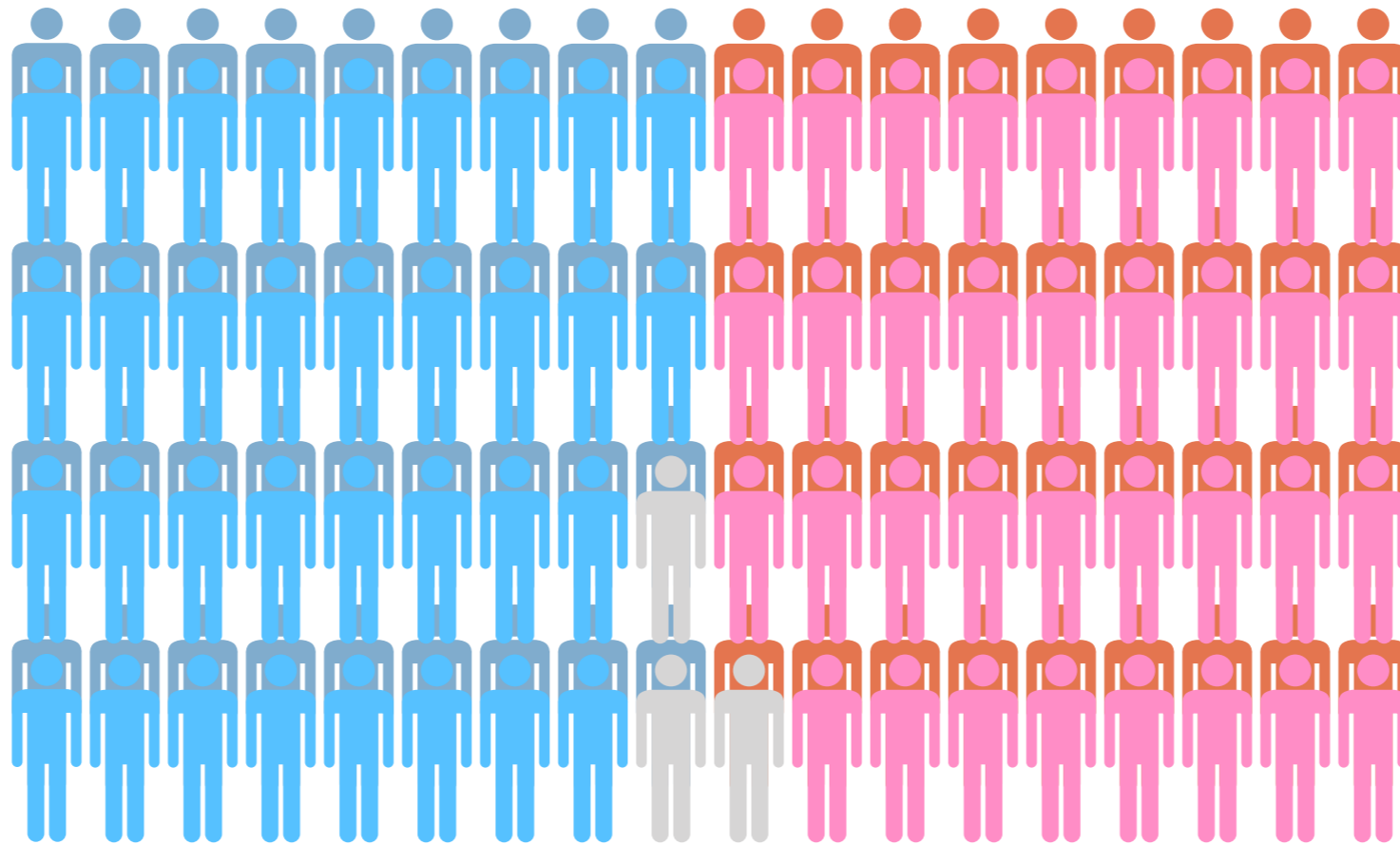
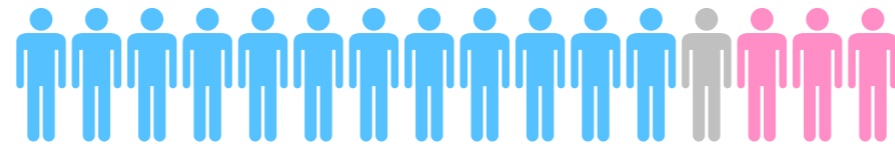
Scenario 2: Small cells in a large population, or small population

- Imputing individual gender and/or
- Difficult to split a cell by a small proportion
- For both cases - uncertainty difficult to incorporate

# Options

1. Impute sample sex assuming a known population counts
2. Impute the population using a known population counts
3. **Impute the population distribution using auxiliary data**

# Options





# Benefits

- Avoids touching the original survey data
- Avoids imputing individual cases
- Potentially more accurate
- More intricate imputation strategy allows us to move beyond imputing male sex as male gender always.

# Concerns and challenges

Scenario 1: Large population, large cells

- Doesn't involve imputing individual gender, rather splitting cell count by proportion

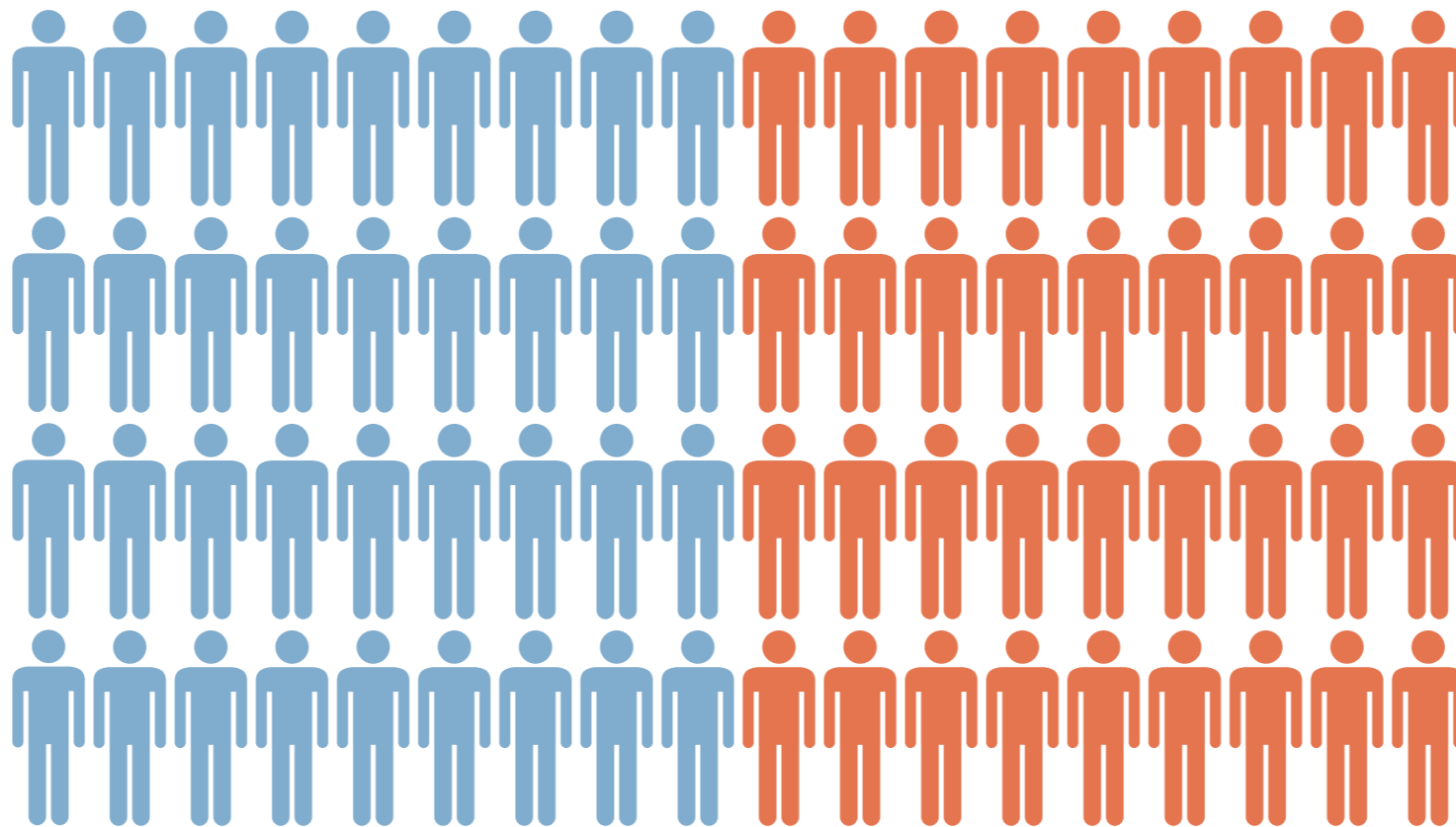
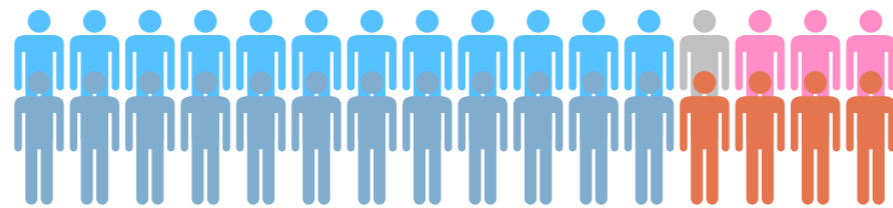
Scenario 2: Small cells in a large population, or small population

- Imputing individual gender and/or
- Difficult to split a cell by a small proportion

# Options

1. Impute sample sex assuming a known population counts
2. Impute the population using a known population counts
3. Impute the population distribution using auxiliary data
4. **Impute sample sex using auxiliary data**

# Options



# Benefits

- Easily explained to other researchers and respondents
- Backwards compatible with previous data collection
- Few infrastructure changes
- More intricate imputation allows us to impute male gender as non-male sex if necessary
- Potentially more accurate than simple population counts

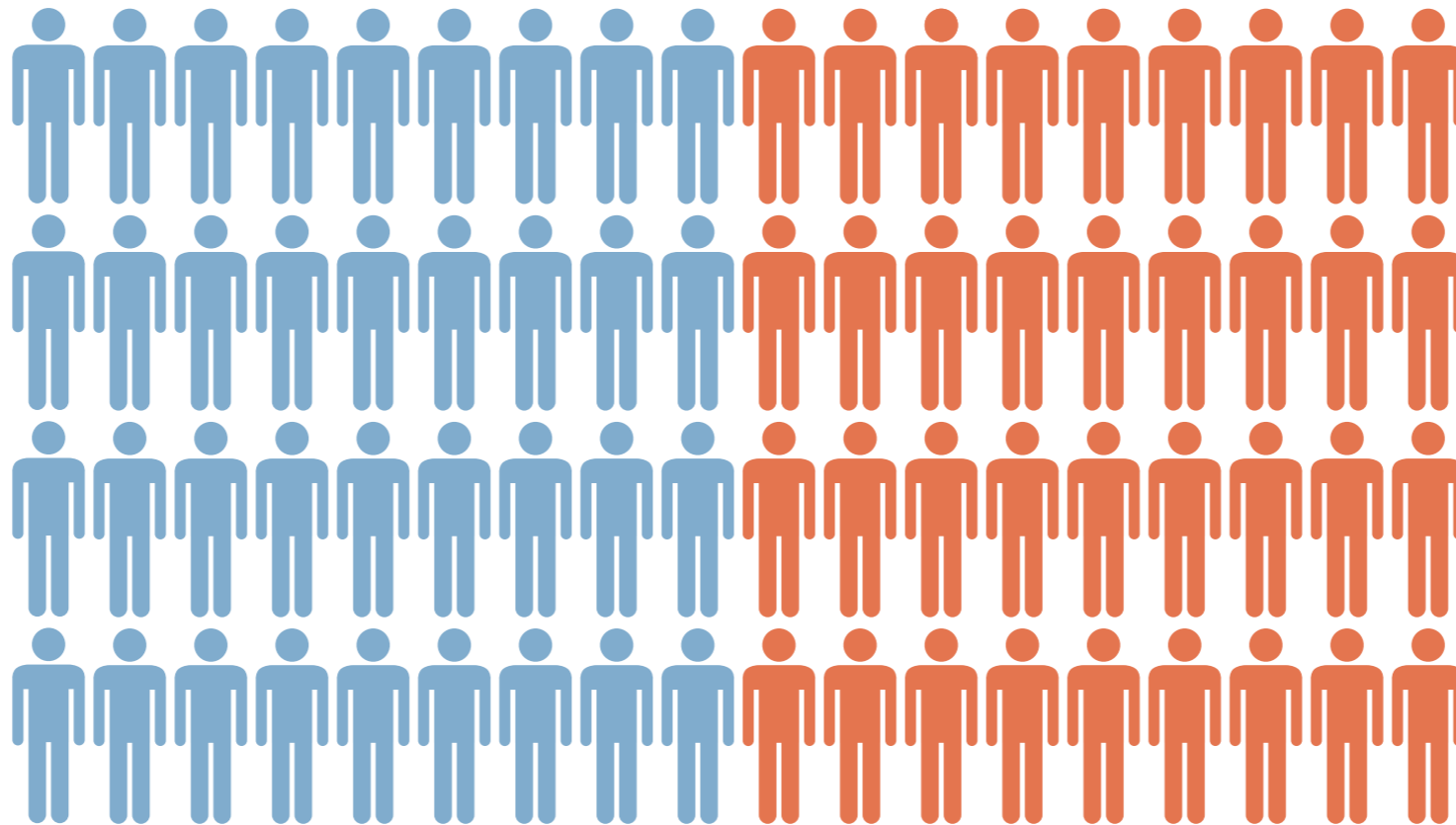
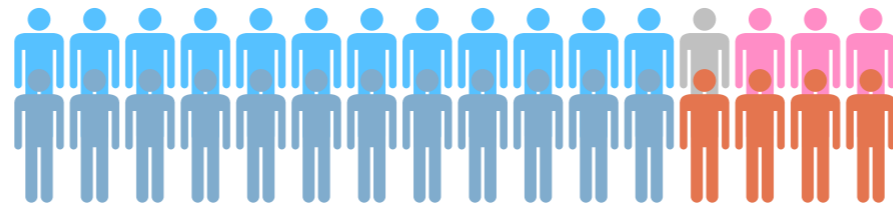
# Concerns and challenges

- Potentially doesn't impute male gender as male sex and female gender as female sex.
- Involves imputing non-binary individuals as either male or female sex (statistical misgendering)
- Imputation error is not the same for all genders
- Auxiliary data needs to be carefully chosen - difference between demographics and other auxiliary data (such as interviewer coder)

# Options

1. Impute sample sex assuming a known population counts
2. Impute the population using a known population counts
3. Impute the population distribution using auxiliary data
4. Impute sample sex using auxiliary data
- 5. Impute all non-male respondents as female**

# Options





# Benefits

- Easily explained to other researchers and respondents
- Theoretical use of constructs as systematic discrimination against due to societal structures.
- Collapsing cells based on most similar outcomes

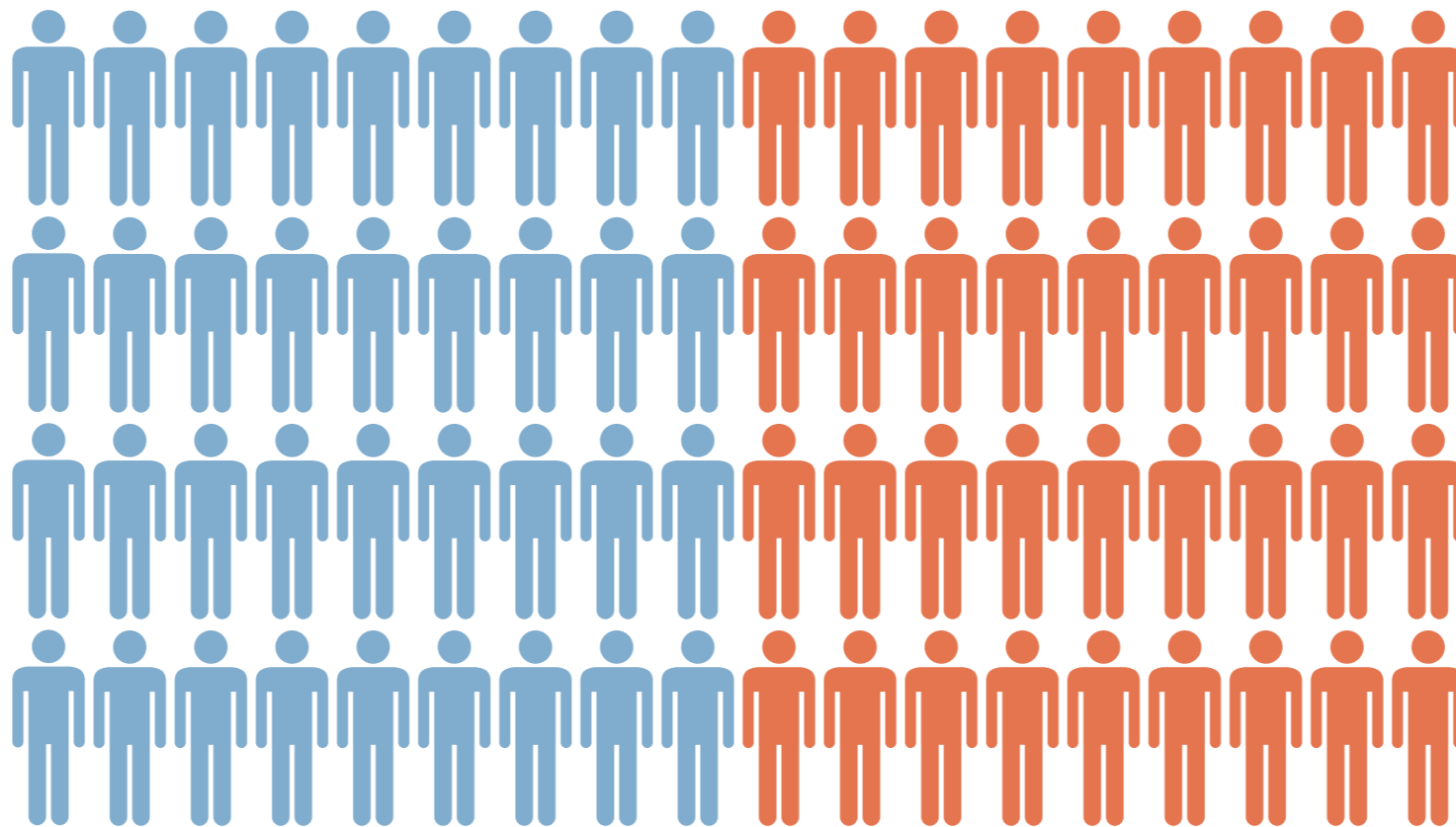
# Concerns and challenges

- Further confuses sex and gender
- Confusing for survey respondents
- Could potentially be a solution for surveys when richer population gender data is available if small cells are an issue

# Options

1. Impute sample sex assuming a known population counts
2. Impute the population using a known population counts
3. Impute the population distribution using auxiliary data
4. Impute sample sex using auxiliary data
5. Impute all non-male respondents as female
6. **Remove respondents**

# Options



# Benefits

- Easily explained to other researchers and respondents
- Statistically simple
- ?

# Concerns and challenges

- Could be a very small proportion (<1% by some sources)
- Unless very different from male/female in terms of outcomes omission is unlikely to impact population level estimates.
- Important and valuable people don't get included in a survey
- Cost to survey participation that isn't justified if responses aren't used

# Simulation study

Assume known population proportions

- Assume a gender distribution in the population - we use 49% M, 2% NB, 49% F but difficult without a census!
- 2% is likely to be an overestimate
- Use to split the post stratification cells, but can't propagate error.

# Simulation study

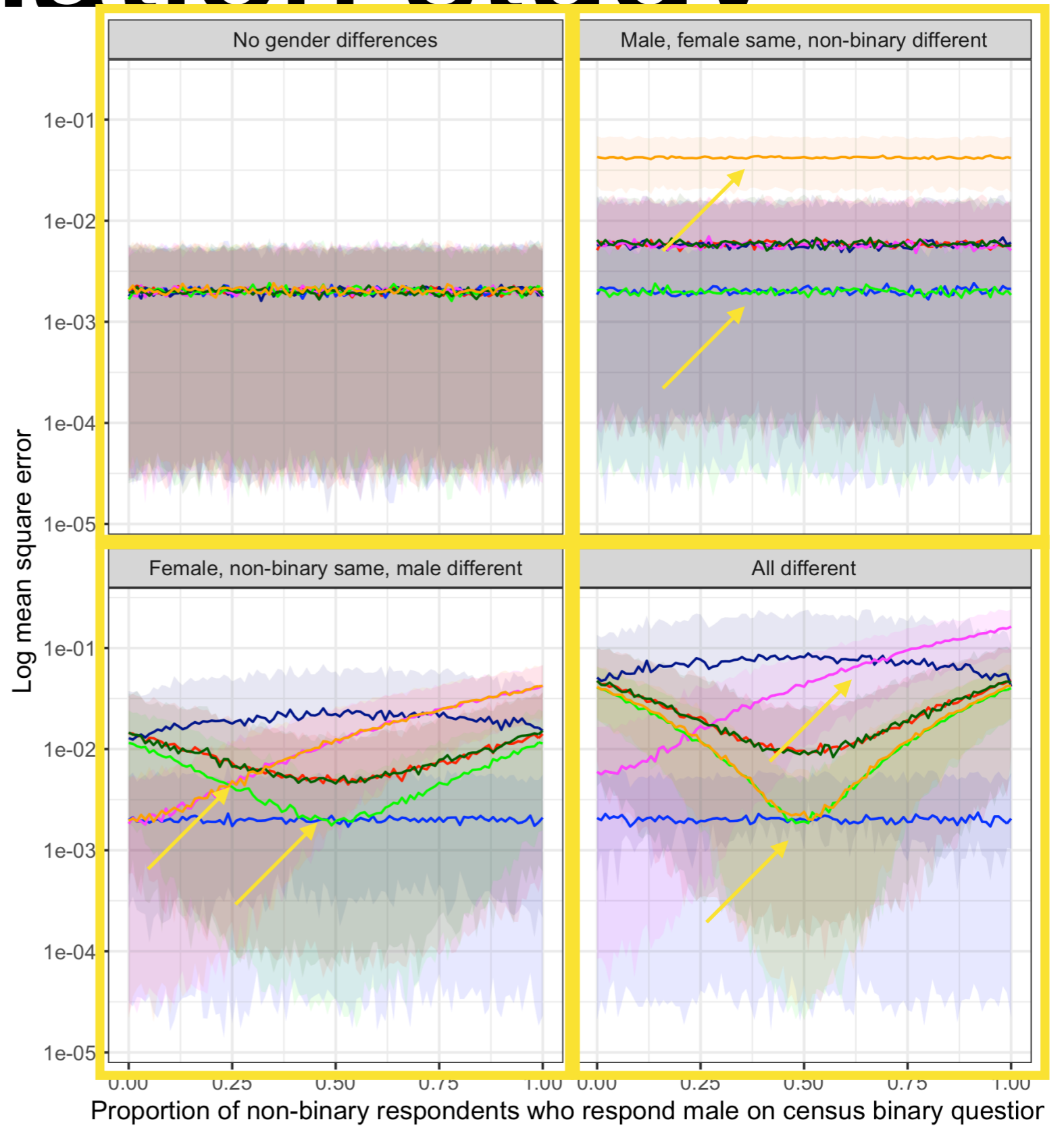
Condition	Male $\mu$	Female $\mu$	non-binary $\mu$
All same	0	0	0
Male, female same	10	10	0
Female, non-binary same	10	0	0
All different	10	-10	0

	Male sex	Female sex
Male gender	47%	2%
Female gender	2%	47%
Non-binary gender	$p \times 2\%$	$p \times 2\%$



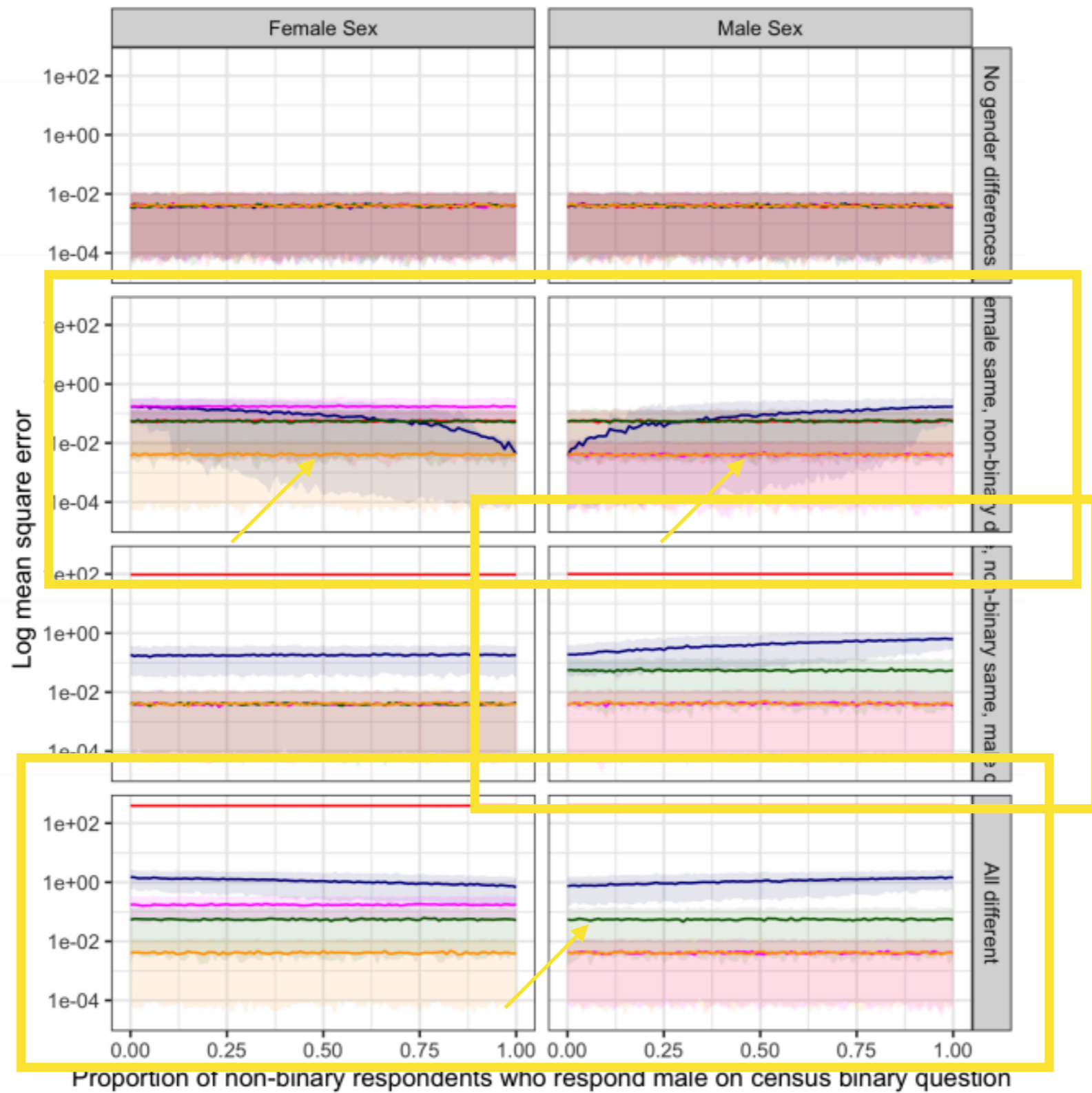
# Simulation study

- Estimate sex in sample, worst model
- Estimate sex in sample, best model
- Estimate gender in popn, by popn prop
- Remove those who respond non-binary
- Estimate gender in popn, best model
- Estimate female sex for those who respond other
- Estimate sex in sample, by popn prop



# Estimating small groups

- If gender is being measured with three (or more categories), great care needs to be taken to protect the privacy of those individuals who respond as non-binary
- Often in surveys of the general population these cells will be too small to release estimates on, and potentially should be removed/protected from public data releases
- However, there are still implications for estimation of female and male sex



- Estimate sex in sample, worst model
  - Estimate gender in popn, best model
  - Estimate sex in sample, best model
  - Estimate female sex for those who respond other
  - Estimate gender in popn, by popn prop
  - Estimate sex in sample, by popn prop
  - Remove those who respond non-binary
- Estimate gender in popn, best model
  - Estimate female sex for those who respond other
  - Estimate sex in sample, by popn prop

# Conclusion

- There's no solution that will work for all situations
- The context for the survey (and why gender/sex was measured) needs to be taken into account
- However, post stratification based survey methods to be careful and understand the implications of mapping responses of constructs.



Statistics > Applications

[Submitted on 30 Sep 2020]

# Using sex and gender in survey adjustment

[Lauren Kennedy](#), [Katharine Khanna](#), [Daniel Simpson](#), [Andrew Gelman](#)

Accounting for sex and gender characteristics is a complex, structural challenge in social science research. While other methodology papers consider issues surrounding appropriate measurement, we consider how gender and sex impact adjustments for non-response patterns in sampling and survey estimates. We consider the problem of survey adjustment arising from the recent push toward measuring sex or gender as a non-binary construct. This is challenging not only in that response categories differ between sex and gender measurement, but also in that both of these attributes are potentially multidimensional. In this manuscript we reflect on similarities to measuring race/ethnicity before considering the ethical and statistical implications of the options available to us. We do not conclude with a single best recommendation but rather an awareness of the complexity of the issues surrounding this challenge and the benefits and weaknesses of different approaches.

Comments: 19 Pages, 2 figures