Alignment Between Open- and Closed-ended Responses for Gender and Sexuality

Christina Pao and Dashram Pai

October 2024

Increase in LGBTQ+ Disclosure

Increase in movement to collect population-level LGBTQ+ data

Americans' Self-Identification as Lesbian, Gay, Bisexual, Transgender or Something Other Than Heterosexual, 2012-2022

Which of the following do you consider yourself to be? You can select as many as apply. Straight or heterosexual; Lesbian; Gay; Bisexual; Transgender
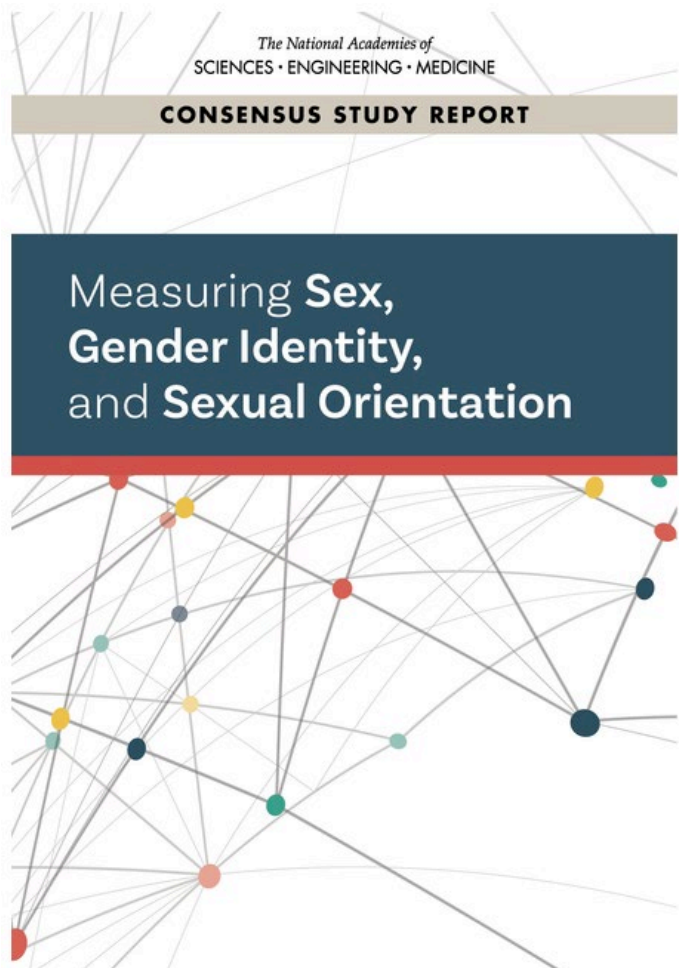
— % Identify as LGBT
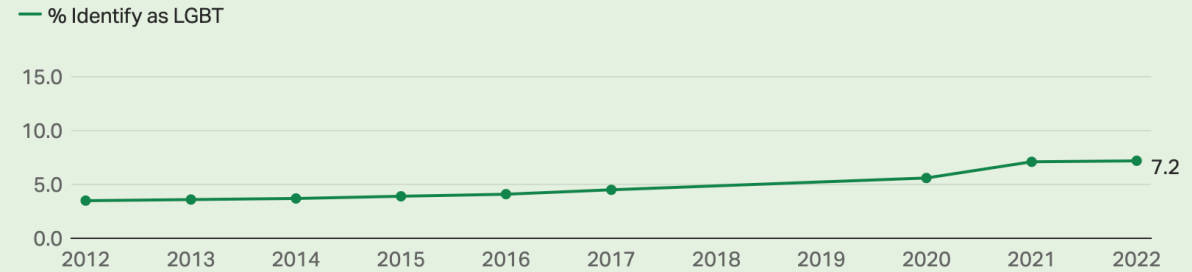
Respondents who volunteer another identity (e.g., queer, same-gender-loving; pansexual) are recorded as "Other LGBT" by interviewers. These responses are included in the LGBT estimate.
Data were not collected in 2018 and 2019.
2012-2013 wording: Do you, personally, identify as lesbian, gay, bisexual or transgender?

Get the data • Download image

GALLUP

be you.
be counted.

National LGBTQ Task Force

2

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

CONSENSUS STUDY REPORT

Measuring **Sex,**
Gender Identity,
and **Sexual Orientation**

Hidden No Longer: Expanding Our Knowledge on Sexual Orientation and Gender Identity

June 26, 2024
Written by: Robert L. Santos, Director, U.S. Census Bureau

Happy Pride Month, everyone! In June our nation celebrates the accomplishments and resilience of our lesbian, gay, bisexual, transgender, queer, intersex, and asexual (LGBTQIA+) community. To further that celebration, I want to talk about how we at the U.S. Census Bureau are expanding our knowledge on measuring sexual orientation and gender identity (SOGI).

Share

Facebook   X (Twitter)   LinkedIn

UCLA School of Law
Williams Institute

Topics   Publications   Experts

Publications   >   Data Collection

REPORT

Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys (GenIUSS)

September 2014

Two-Step Gender Question

Gradational Gender Scales

Q10.2

What sex were you assigned at birth? (For example, on your birth certificate.)

- ○ Female
- ○ Male
- ○ Other
- ○ Prefer not to say
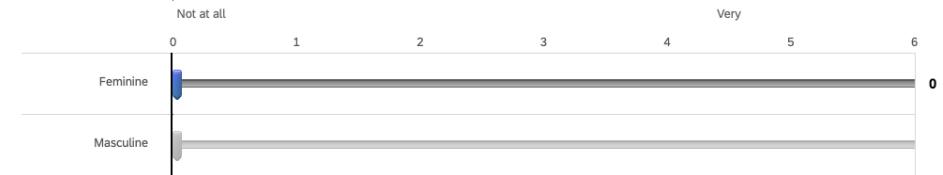
Q10.3

What is your current gender?

- ○ Woman
- ○ Man
- ○ Non-binary / third gender
- ○ A gender not listed here (please specify)

Q9.1

In general, how do **you** see **yourself**? Please answer on both scales below:

Not at all — Very
0 1 2 3 4 5 6
Feminine
Masculine

Q9.2

In general, how do **most people** see **you**? Please answer on both scales below:

Not at all — Very
0 1 2 3 4 5 6
Feminine
Masculine

Q1

What is your gender?

Q2

How would you describe your sexuality?

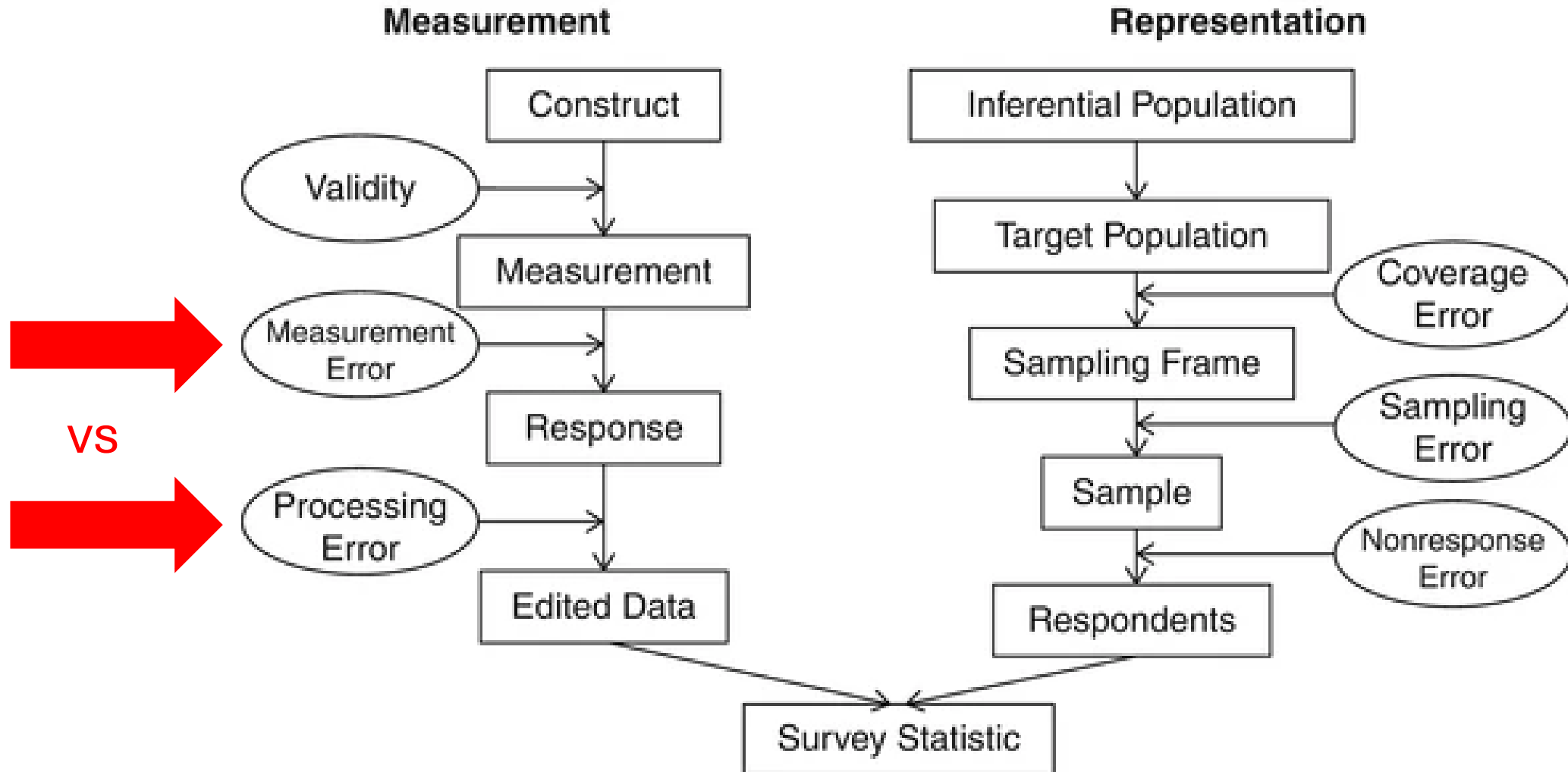# What about open-text?

Tradeoffs

Pros	Cons
Detailed responses	Coding logistics
Responsive to emergent identities	Concordance with old measures
Inclusivity	Trolls and respondent confusion

Total Survey Error Framework
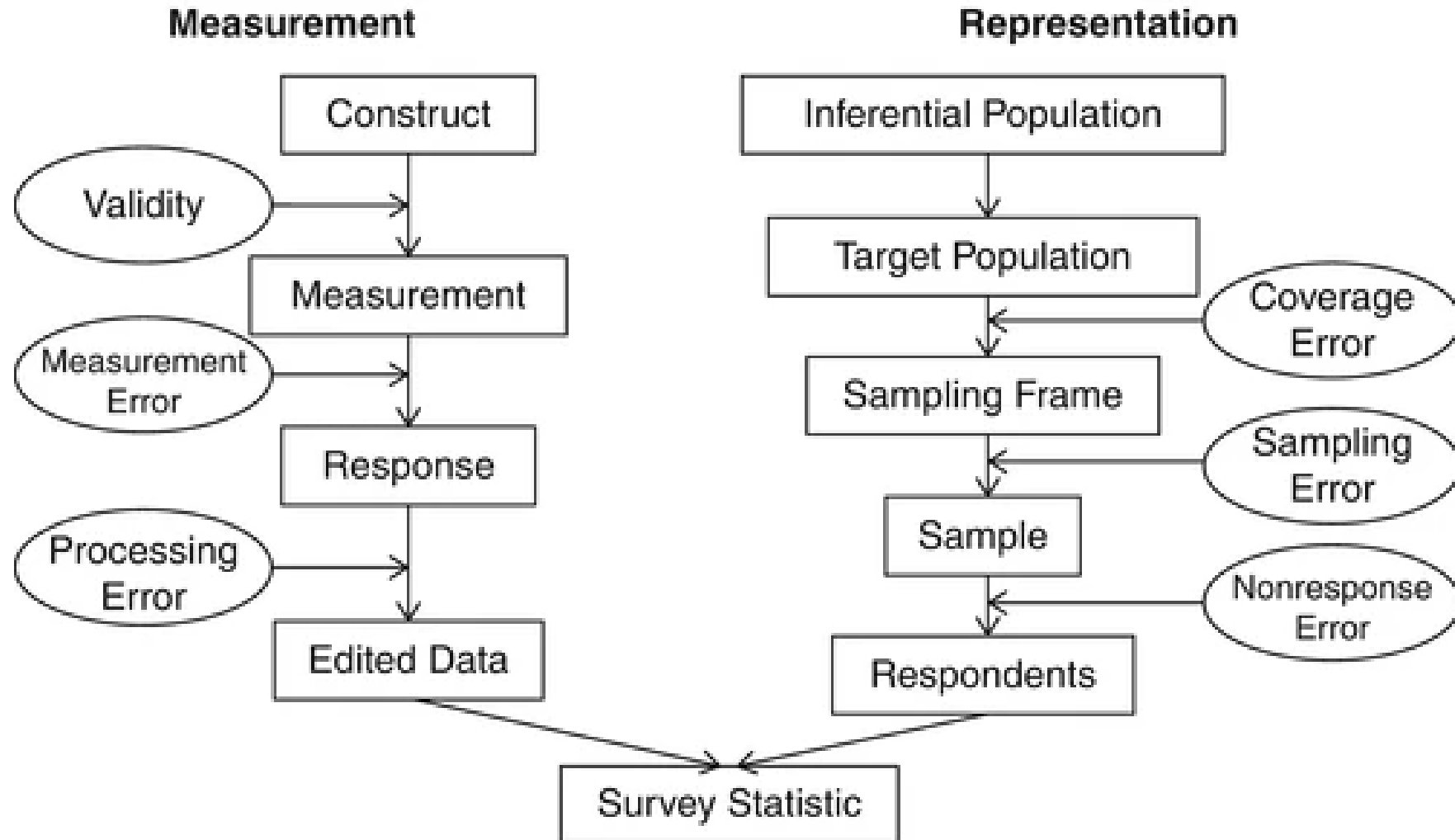
Total Survey Error Framework

RQ1: How well do open text measures align with traditional (closed-ended) categorical measures?

RQ2: How much new information do we gather about majority and minority groups, respectively, from open text?

Case Study: Gender Identity and Sexual Orientation

Data and Methods

Original Survey in US and UK (2022) N ~ 2,600

Data	Dependent Vars.
• **Lucid** (opt-in convenience sample) • **Quotas** (age, region, sex, education, race—in US sample)	• **Alignment:** correspondence between closed-ended and open-text responses • **New information**: information gleaned from open-ended that was not provided in closed-ended

Data collected in collaboration with Prof. Ridhi Kashyap (University of Oxford)

Original Survey in US and UK (2022)
N ~ 2,600

Gender

- Open Text

- Two Step Gender Question

  - Sex

  - Gender

Sample Two-Step

Q10.2

What sex were you assigned at birth? (For example, on your birth certificate.)

○ Female

○ Male

○ Other

○ Prefer not to say

Q10.3

What is your current gender?

○ Woman

○ Man

○ Non-binary / third gender

○ A gender not listed here (please specify)

Original Survey in US and UK (2022)
N ~ 2,600

Gender

- Open Text

- Two Step Gender Question
    - Sex
    - Gender

Sample Two-Step

Q10.2

What sex were you assigned at birth? (For example, on your birth certificate.)

○ Female

○ Male

○ Other

○ Prefer not to say

Q10.3

What is your current gender?

○ Woman

○ Man

○ Non-binary / third gender

○ A gender not listed here (please specify)

Original Survey in US and UK (2022)
N ~ 2,600

Sexuality

- Open Text

- UK census categories
  - straight, gay/lesbian, bisexual, prefer not to say
  - select one

- Expansive categories (used on university surveys)
  - 12 categories
  - select all that apply

Expansive List

Out of this expanded list, please select all that apply to you:

- ☐ Aromantic
- ☐ Asexual
- ☐ Bisexual
- ☐ Fluid
- ☐ Gay
- ☐ Lesbian
- ☐ Pansexual
- ☐ Queer
- ☐ Questioning or unsure
- ☐ Straight (heterosexual)
- ☐ Prefer not to disclose
- ☐ Additional category/identity not listed

Original Survey in US and UK (2022)
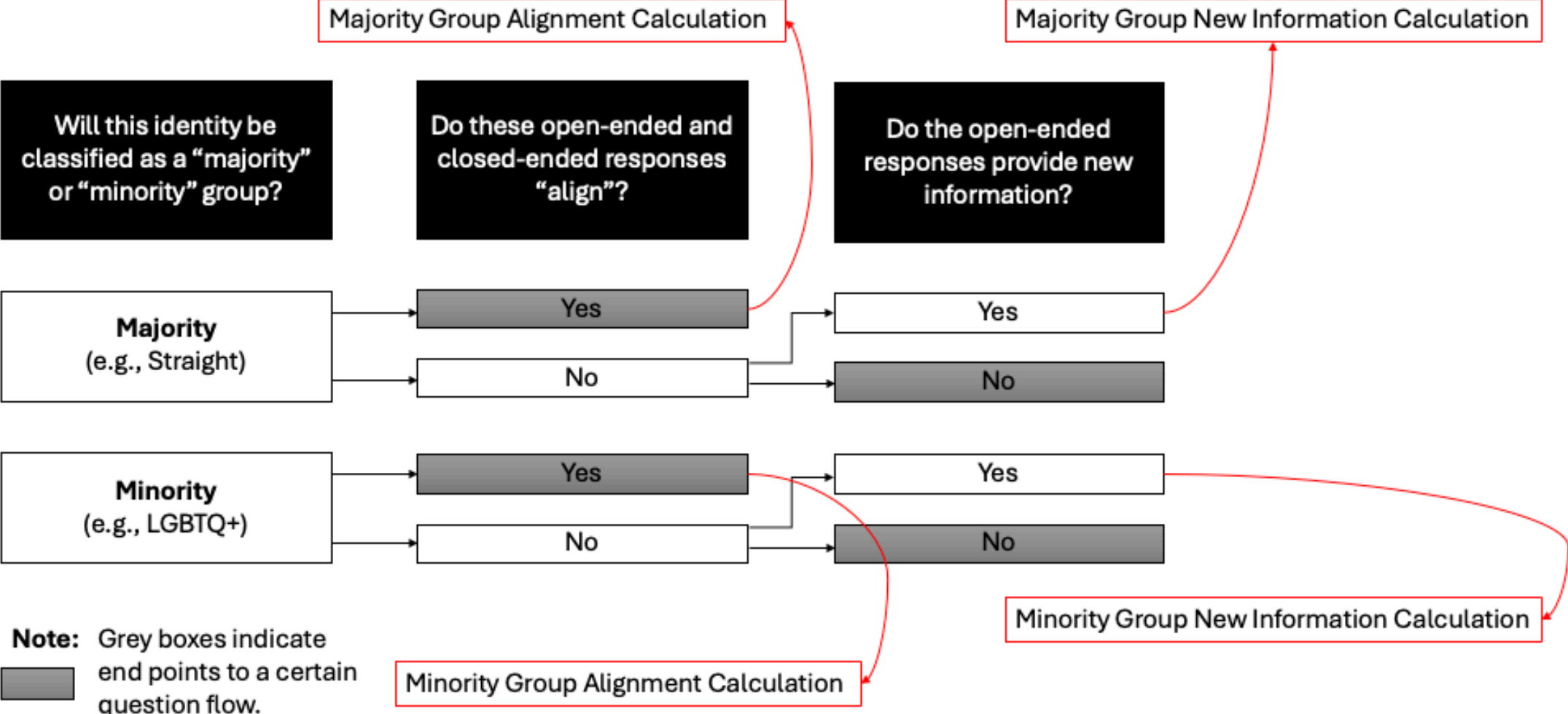# N ~ 2,600

Sexuality

- Open Text
- UK census categories
  - straight, gay/lesbian, bisexual, prefer not to say
  - select one

- Expansive categories (used on university surveys)
  - 12 categories
  - select all that apply

Expansive List

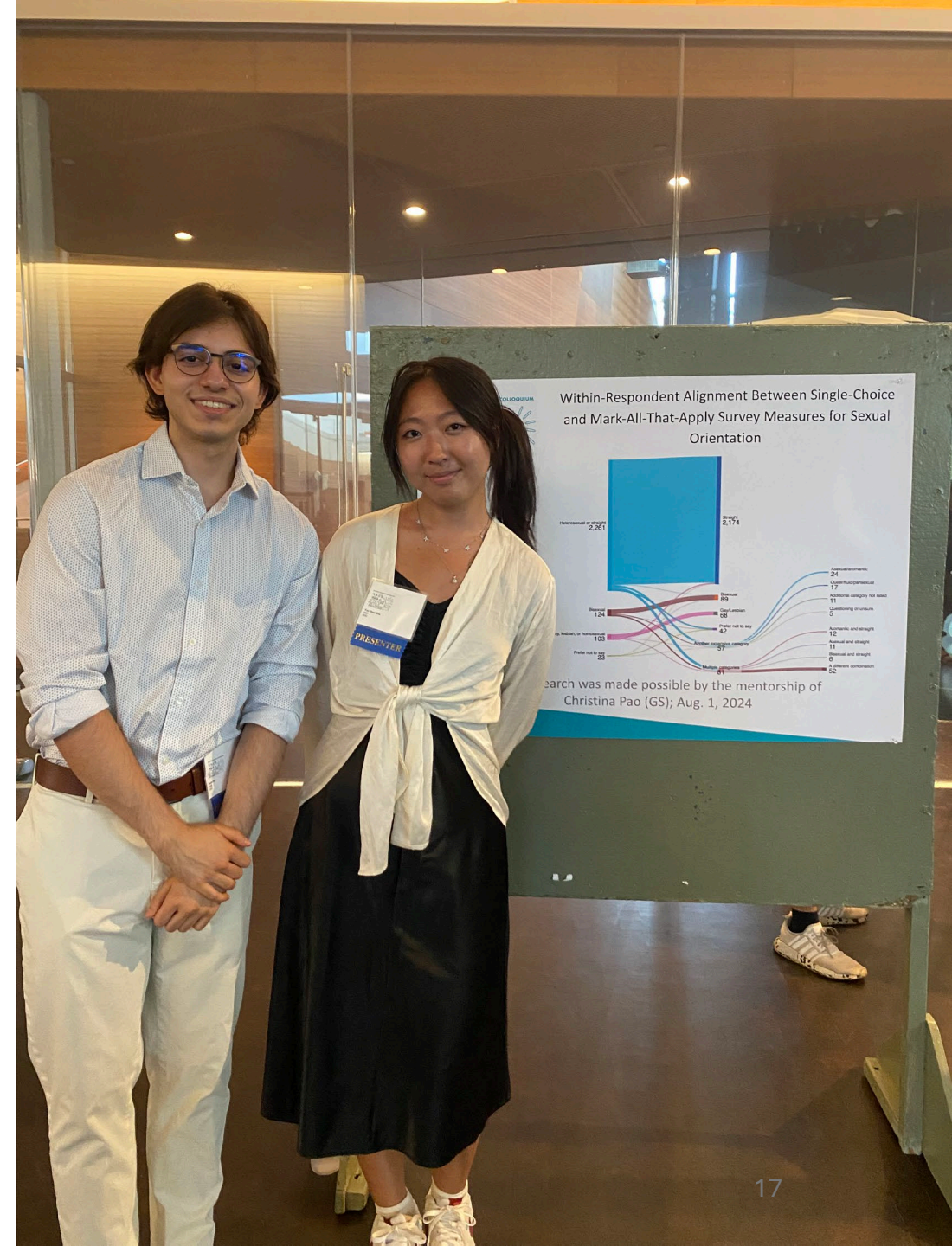Out of this expanded list, please select all that apply to you:

☐ Aromantic

☐ Asexual

☐ Bisexual

☐ Fluid

☐ Gay

☐ Lesbian

☐ Pansexual

☐ Queer

☐ Questioning or unsure

☐ Straight (heterosexual)

☐ Prefer not to disclose

☐ Additional category/identity not listed

Overview of Coding Open-ended Responses

Majority Group Alignment Calculation

Majority Group New Information Calculation

Will this identity be classified as a "majority" or "minority" group?	Do these open-ended and closed-ended responses "align"?	Do the open-ended responses provide new information?

Majority (e.g., Straight)
- Yes
- No

- Yes
- No

Minority (e.g., LGBTQ+)
- Yes
- No

- Yes
- No

Minority Group New Information Calculation

Minority Group Alignment Calculation

Note: Grey boxes indicate end points to a certain question flow.

16

Coding Process

Thanks to two undergraduate RAs, Dash Pai (left) and Yan Zhen Zhu (right)!

Coding Steps
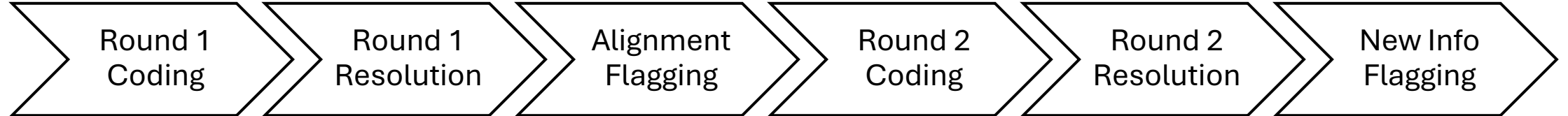
Round 1 Coding → Round 1 Resolution → Alignment Flagging → Round 2 Coding → Round 2 Resolution → New Info Flagging

Coding Steps

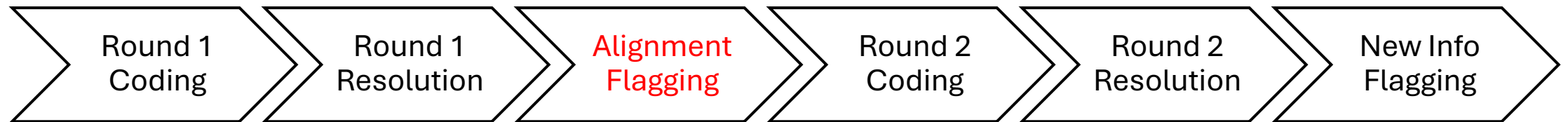| Round 1 Coding	Round 1 Resolution	Alignment Flagging	Round 2 Coding	Round 2 Resolution	New Info Flagging

- RAs coded open-ended responses according to pre-established codebooks
- Avg. 1.5 hours

Coding Steps
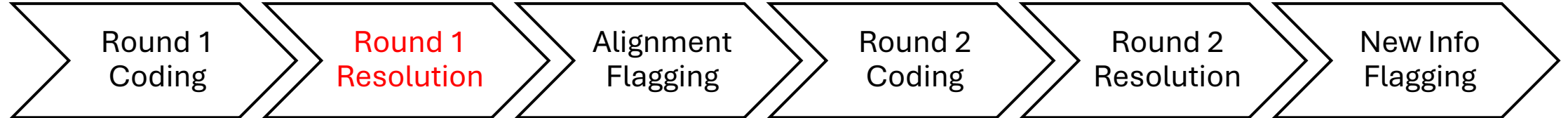
Round 1
Coding

Round 1
Resolution

Alignment
Flagging

Round 2
Coding

Round 2
Resolution

New Info
Flagging

- RAs coded open-ended responses according to pre-established codebooks
- Avg. 1.5 hours

- Finalizing codes that were nonaligned between RAs
- 50 minutes

Coding Steps

```
┌──────────┐  ┌──────────┐  ┌──────────┐  ┌──────────┐  ┌──────────┐  ┌──────────┐
│ Round 1  │  │ Round 1  │  │Alignment │  │ Round 2  │  │ Round 2  │  │ New Info │
│ Coding   │  │Resolution│  │ Flagging │  │ Coding   │  │Resolution│  │ Flagging │
└──────────┘  └──────────┘  └──────────┘  └──────────┘  └──────────┘  └──────────┘
```
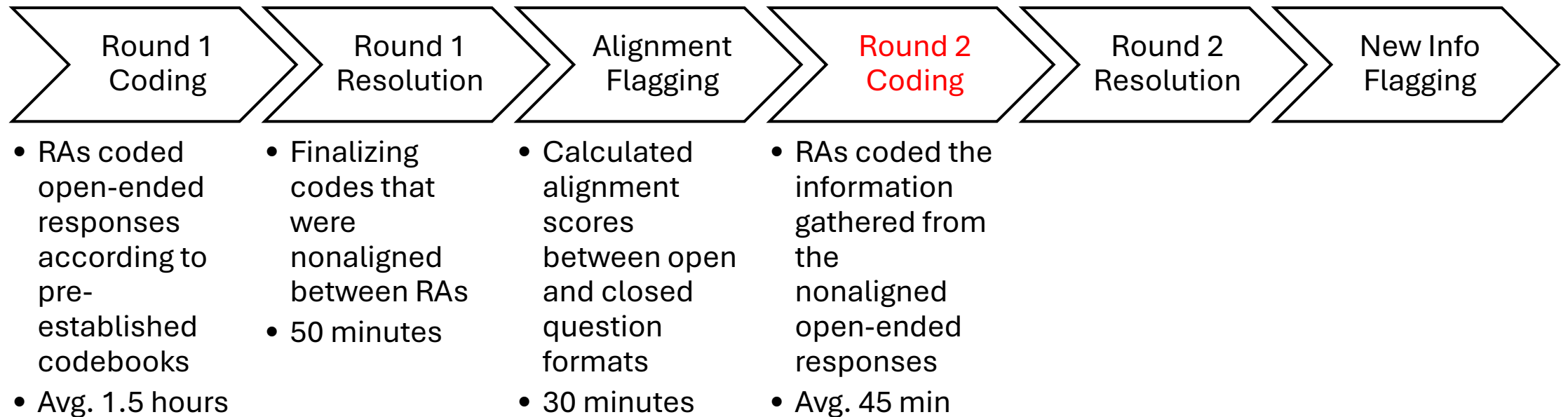
- RAs coded open-ended responses according to pre-established codebooks
- Avg. 1.5 hours

- Finalizing codes that were nonaligned between RAs
- 50 minutes

- Calculated alignment scores between open and closed question formats
- 30 minutes

Coding Steps

Round 1 Coding	Round 1 Resolution	Alignment Flagging	Round 2 Coding	Round 2 Resolution	New Info Flagging

- RAs coded open-ended responses according to pre-established codebooks
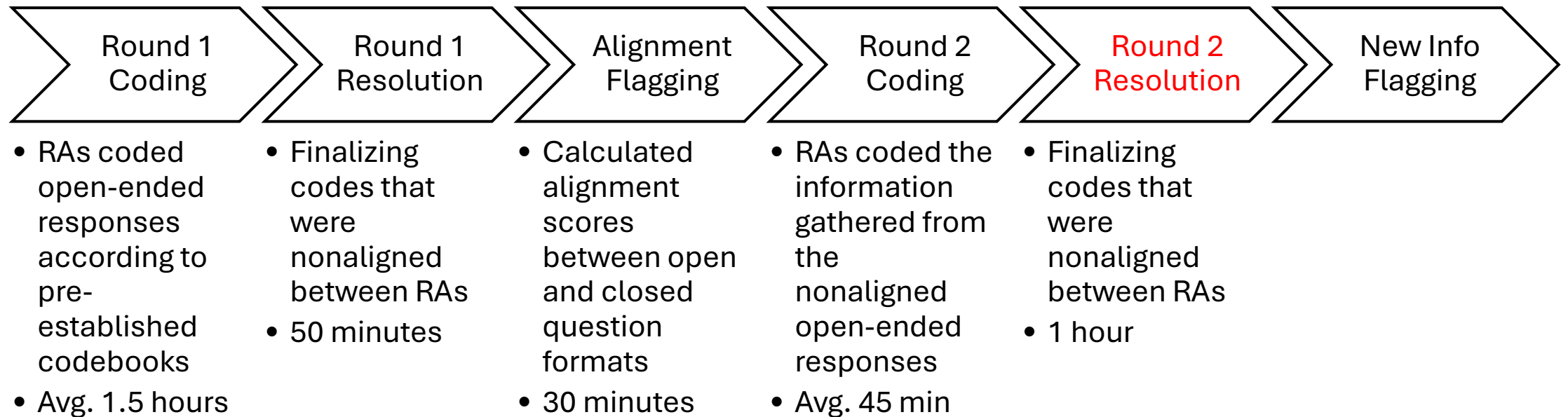- Avg. 1.5 hours

- Finalizing codes that were nonaligned between RAs
- 50 minutes

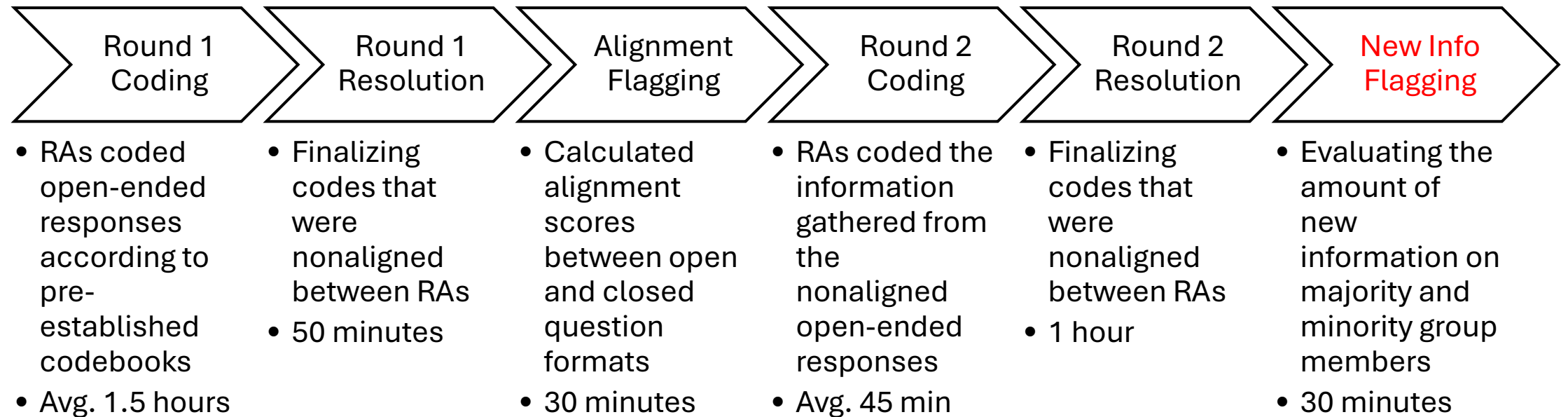- Calculated alignment scores between open and closed question formats
- 30 minutes

- RAs coded the information gathered from the nonaligned open-ended responses
- Avg. 45 min

Coding Steps

Round 1 Coding	Round 1 Resolution	Alignment Flagging	Round 2 Coding	Round 2 Resolution	New Info Flagging

- RAs coded open-ended responses according to pre-established codebooks
- Avg. 1.5 hours

- Finalizing codes that were nonaligned between RAs
- 50 minutes

- Calculated alignment scores between open and closed question formats
- 30 minutes

- RAs coded the information gathered from the nonaligned open-ended responses
- Avg. 45 min

- Finalizing codes that were nonaligned between RAs
- 1 hour

Coding Steps

| Round 1 Coding | Round 1 Resolution | Alignment Flagging | Round 2 Coding | Round 2 Resolution | New Info Flagging |

- RAs coded open-ended responses according to pre-established codebooks
- Avg. 1.5 hours

- Finalizing codes that were nonaligned between RAs
- 50 minutes

- Calculated alignment scores between open and closed question formats
- 30 minutes

- RAs coded the information gathered from the nonaligned open-ended responses
- Avg. 45 min

- Finalizing codes that were nonaligned between RAs
- 1 hour

- Evaluating the amount of new information on majority and minority group members
- 30 minutes

Findings

Finding 1

RQ1: How well do open text measures align with traditional (closed-ended) categorical measures?

Gender identity open text has **high alignment** with a closed-ended measure among men and women.

Sexual orientation open text has **low alignment** with a closed-ended measure, particularly among non-straight respondents.
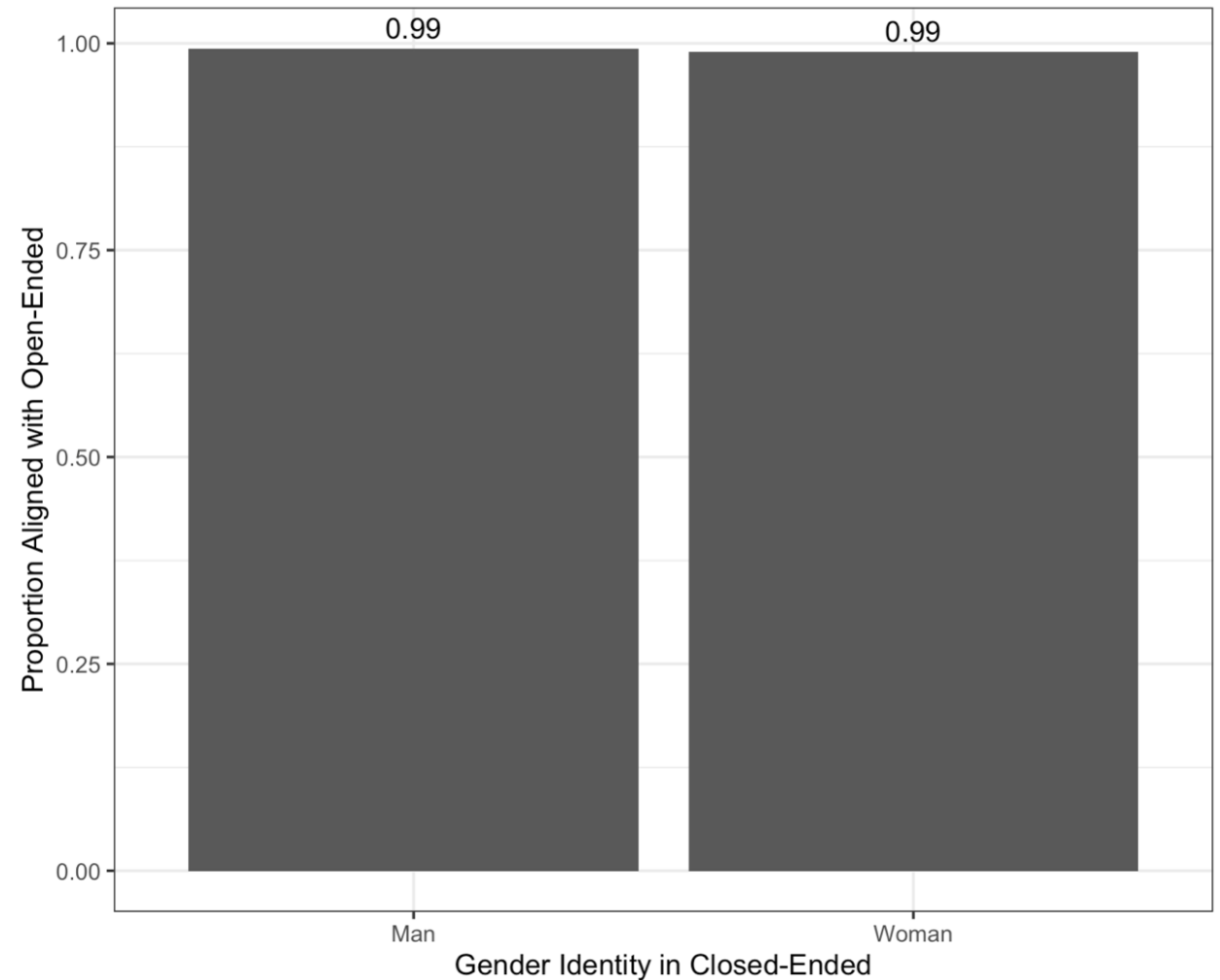
Gender Response Alignment

What is your gender?

99%

of responses align between open text and closed-ended responses within respondents.

Gender Response Alignment

What is your gender?

High alignment among both men and women between closed- and open- ended formats

Sexuality Response Alignment

How would you describe your sexuality?

80%

of responses align between open text and closed-ended responses within respondents.

Sexuality Response Alignment

How would you describe your sexuality?

However, there are large differences in alignment between majority and minority group members.
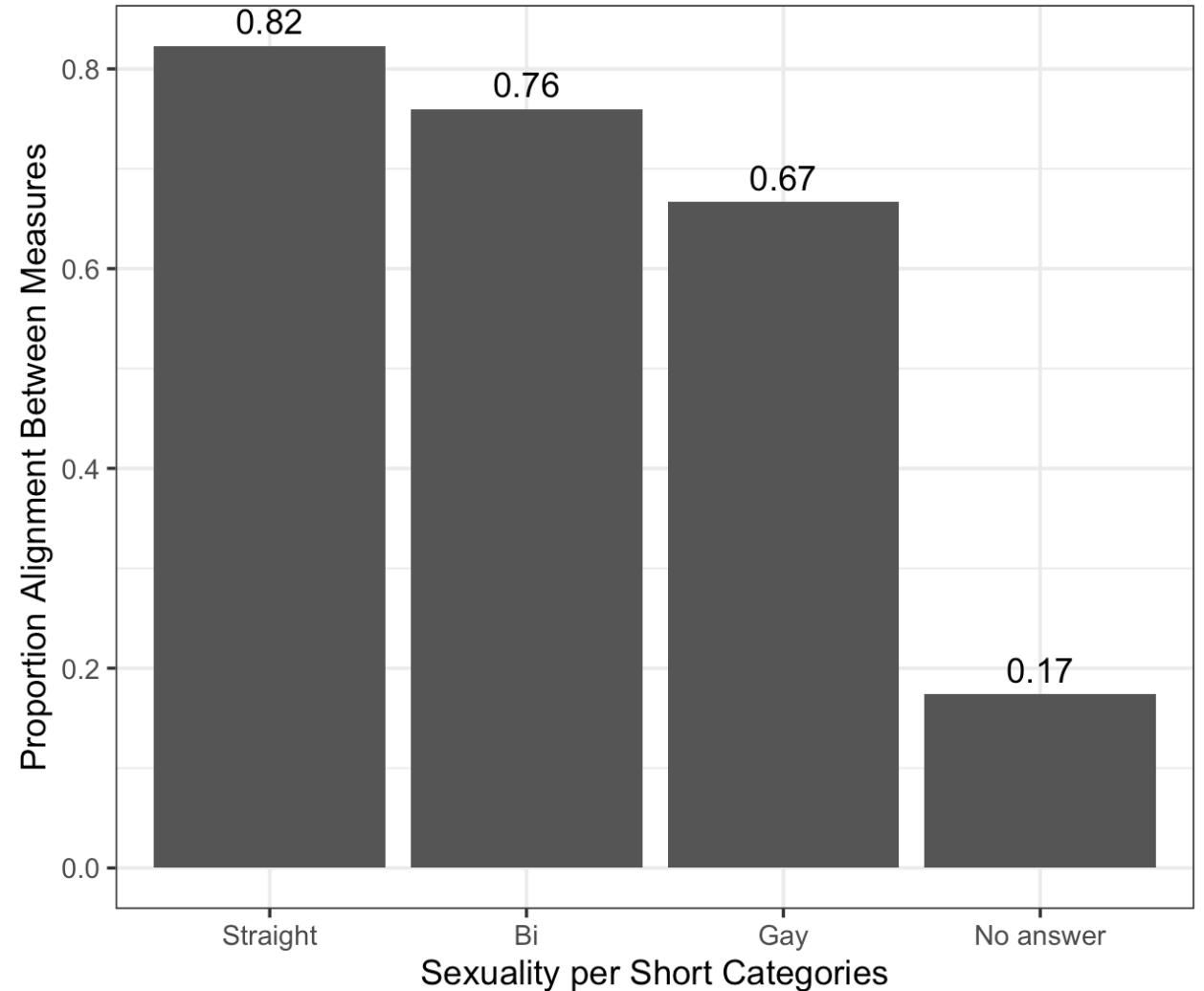
Finding 2

RQ2: How much new information do we gather about majority and minority groups, respectively, from open text?

Gender identity open text does not allow us to gain much information about the majority group, but it does allow us to gain information about the minority group.

Sexual orientation open text allows us to gain information about both the majority and minority group.

Gender Responses: New information

What is your gender?

Out of those who did not fit into "man" or "woman" categories...

| Code	#	Examples
Trans, nonbinary, or other gender term	11	"nonbinary", "Gender queer", "Androgynous", "non-binary", "Genderfluid"
Prefer not to say or don't know	3	"unsure", "Prefer not to say", "Gender is what the individual thinks it is"
Missing data or invalid response (e.g., a numeric)	14	"43", "Straight", "bcde", "mn,m."

Sexuality Responses: New information

What is your sexuality?
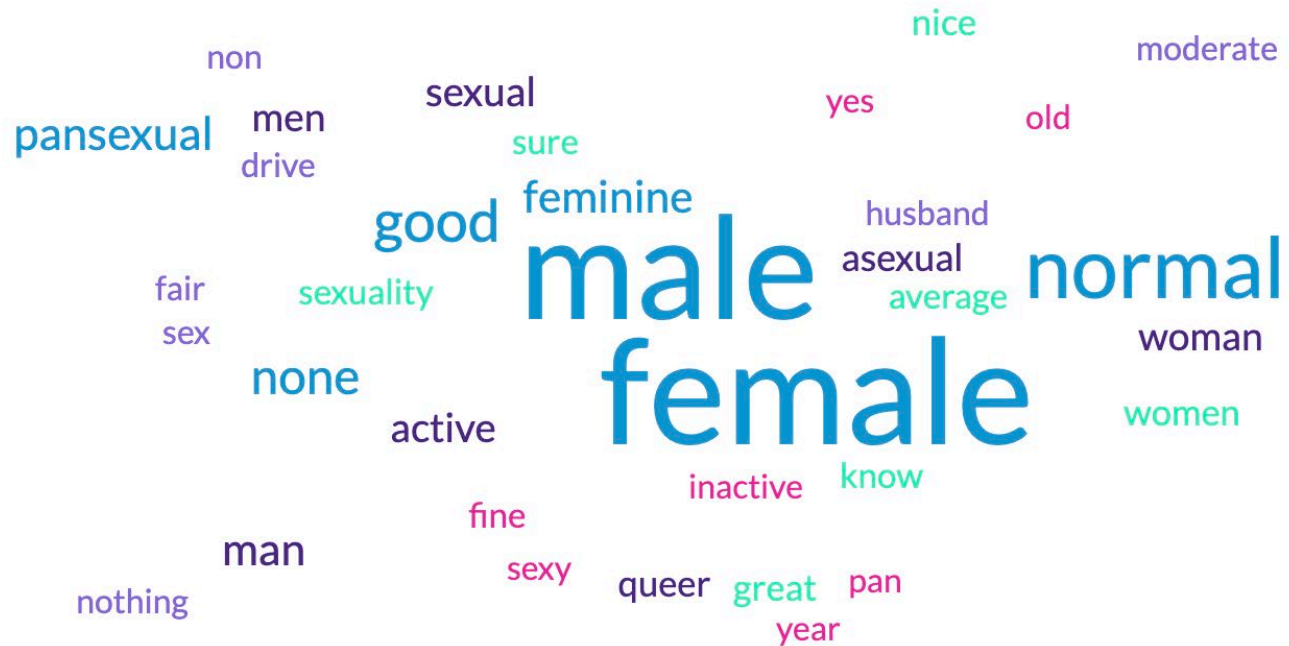
Sexuality Responses: New information

What is your sexuality?
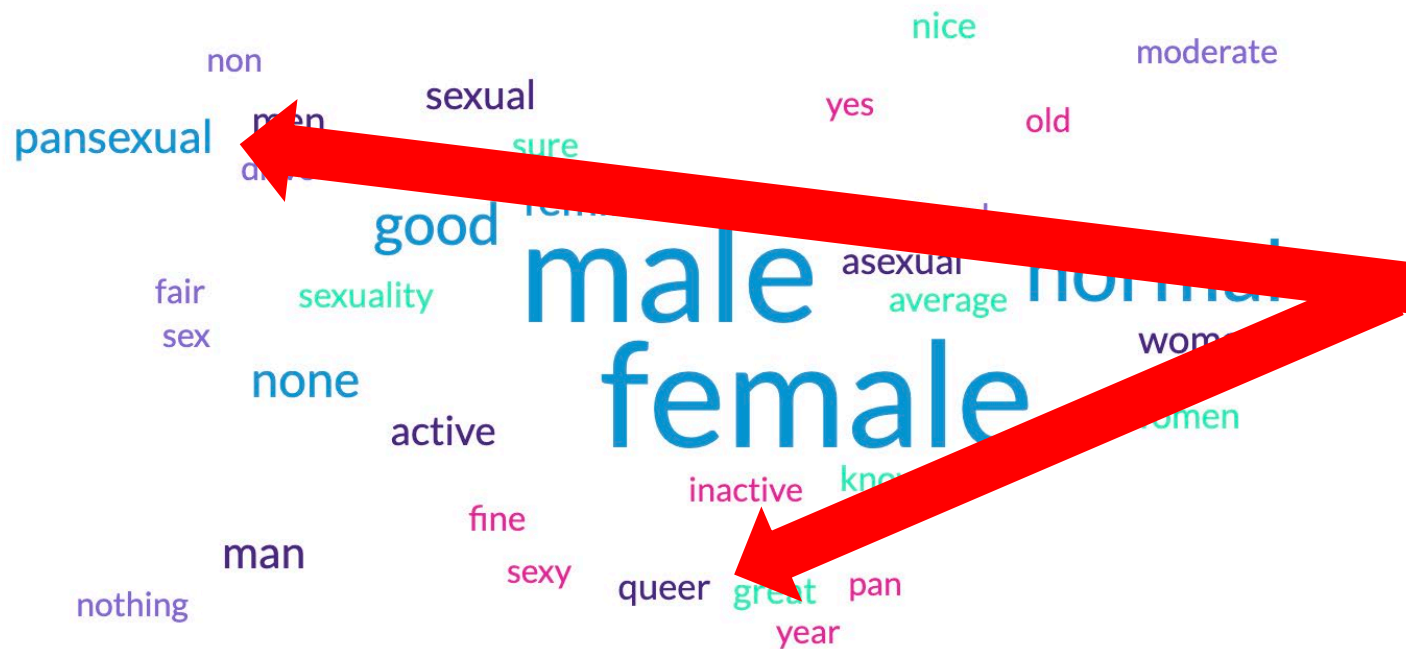
Other expansive sexuality terms

Sexuality Responses: New information

What is your sexuality?

Conflation of Sex, Gender, and Sexuality Terms
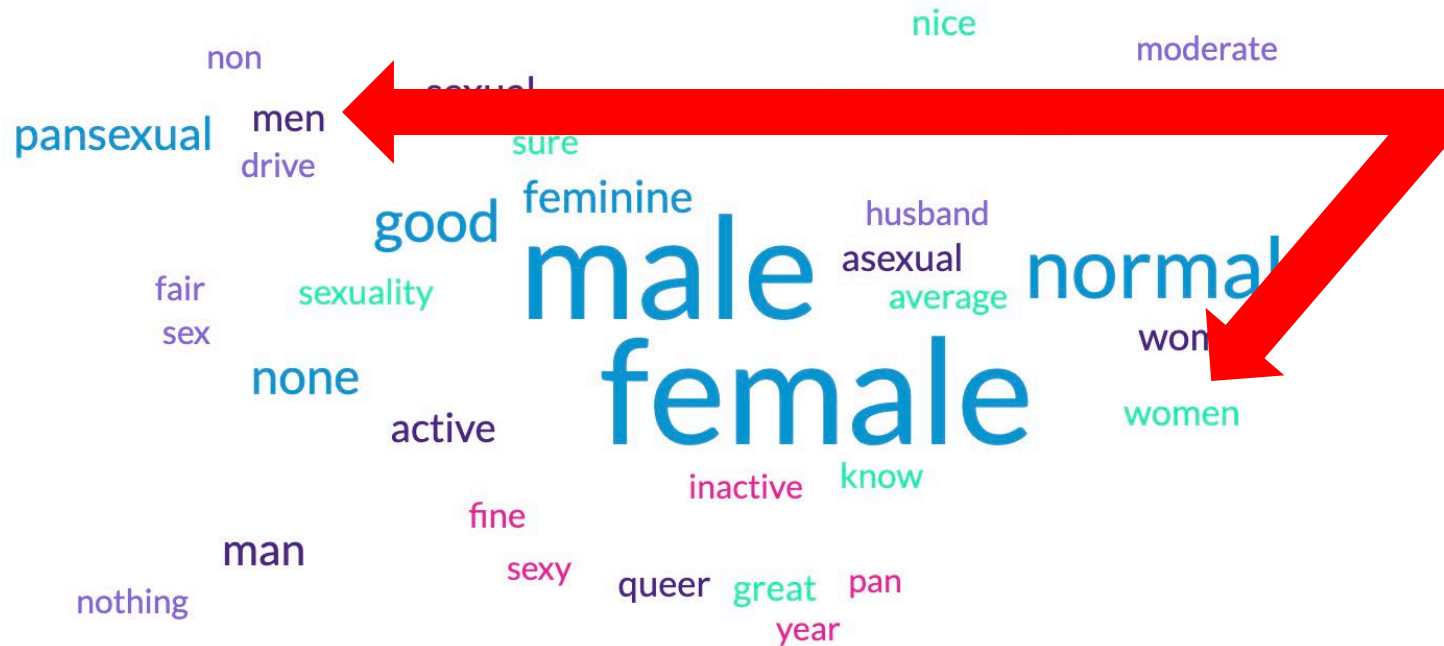
Sexuality Responses: New information

What is your sexuality?

Substitution of marital status

Sexuality Responses: New information

What is your sexuality?

Discussion of sexual behavior

Sexuality Responses: New information

What is your sexuality?

Naming the gender of sexual attraction (instead of own self-ID label)

Off-topic or revealing?

Circumstantial Sexuality

- "No longer an important part of my life. I am old, bald and overweight, not the posterboy for sexy"
- "a middle-aged, single, gay man trying to manage family obligation while watching his dating life pass him by"

Attraction/Behavior vs. Identity

- "straight until I have a little liquor"
- "straight with lesbian tendencies"
- "straight but like girls sometimes"

Intra-straight heterogeneity?

- "MANAGOMOUS"
- "Fucking Straight"
- "HETERAL SEXIUAL"

In summary...

Recapitulation

There are tradeoffs to using open-ended and closed-ended measures, and this is a particularly unique time to study gender and sexuality.

We found that there was high alignment between open-ended and closed-ended formats of gender questions, but not of sexuality questions.

The new information gleaned from *nonaligned* open-ended responses reveals particularly interesting information for sexuality.

At scale, gender open-ended text questions are likely feasible.

In small samples, sexuality open-ended text questions may be interesting and tenable, but they will also yield some respondent confusion.

Recapitulation

There are tradeoffs to using open-ended and closed-ended measures, and this is a particularly unique time to study gender and sexuality.

We found that there was high alignment between open-ended and closed-ended formats of gender questions, but not of sexuality questions.

The new information gleaned from *nonaligned* open-ended responses reveals particularly interesting information for sexuality.

At scale, gender open-ended text questions are likely feasible.

In small samples, sexuality open-ended text questions may be interesting and tenable, but they will also yield some respondent confusion.

Recapitulation

There are tradeoffs to using open-ended and closed-ended measures, and this is a particularly unique time to study gender and sexuality.

We found that there was high alignment between open-ended and closed-ended formats of gender questions, but not of sexuality questions.

The new information gleaned from *nonaligned* open-ended responses reveals particularly interesting information for sexuality.

At scale, gender open-ended text questions are likely feasible.

In small samples, sexuality open-ended text questions may be interesting and tenable, but they will also yield some respondent confusion.

Recapitulation

There are tradeoffs to using open-ended and closed-ended measures, and this is a particularly unique time to study gender and sexuality.

We found that there was high alignment between open-ended and closed-ended formats of gender questions, but not of sexuality questions.

The new information gleaned from *nonaligned* open-ended responses reveals particularly interesting information for sexuality.

At scale, gender open-ended text questions are likely feasible.

In small samples, sexuality open-ended text questions may be interesting and tenable, but they will also yield some respondent confusion.

Recapitulation

There are tradeoffs to using open-ended and closed-ended measures, and this is a particularly unique time to study gender and sexuality.

We found that there was high alignment between open-ended and closed-ended formats of gender questions, but not of sexuality questions.

The new information gleaned from *nonaligned* open-ended responses reveals particularly interesting information for sexuality.

At scale, gender open-ended text questions are likely feasible.

In small samples, sexuality open-ended text questions may be interesting and tenable, but they will also yield some respondent confusion.

Thank you!

Christina Pao (PhD Student, Princeton University)

christina.pao@princeton.edu

Dashram Pai (Undergraduate, Princeton University)

dashpai@princeton.edu

And stay tuned! Forthcoming work evaluating alignment between short-form and expansive mark-all-that-apply options for sexuality in *Socius*.