Order Effects and Question Wording of Gender and Sexuality Blocks on Surveys: Experimental Evidence in the US and UK

Christina Pao (Princeton University)

PRINCETON UNIVERSITY

DEPARTMENT OF SOCIOLOGY

OPR
OFFICE OF POPULATION RESEARCH

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

7.2
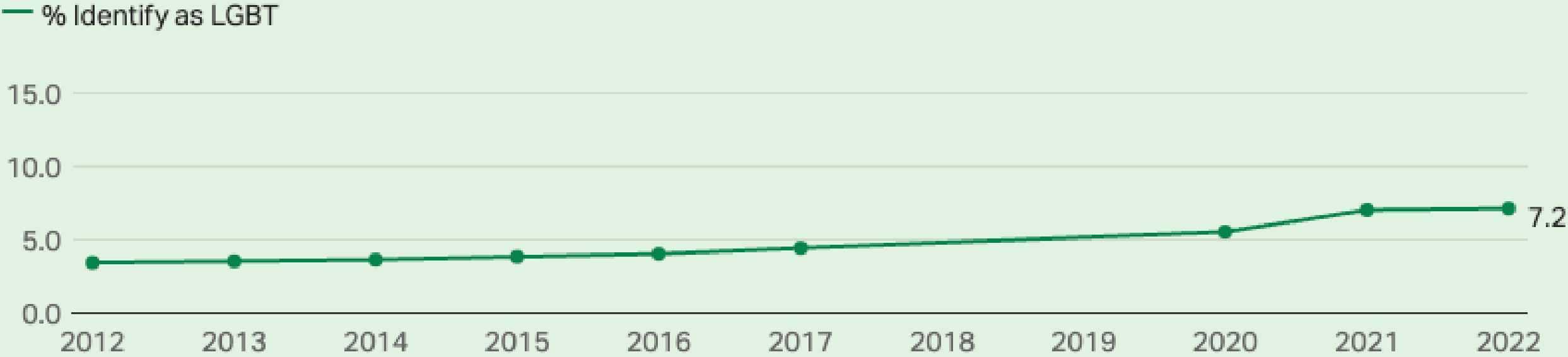
Respondents who volunteer another identity (e.g., queer, same-gender-loving; pansexual) are recorded as "Other LGBT" by interviewers. These responses are included in the LGBT estimate.
Data were not collected in 2018 and 2019.
2012-2013 wording: Do you, personally, identify as lesbian, gay, bisexual or transgender?

Get the data • Download image

GALLUP

Contention and the Two-Step

Fair Play For Women v The Office For National Statistics

census 2021

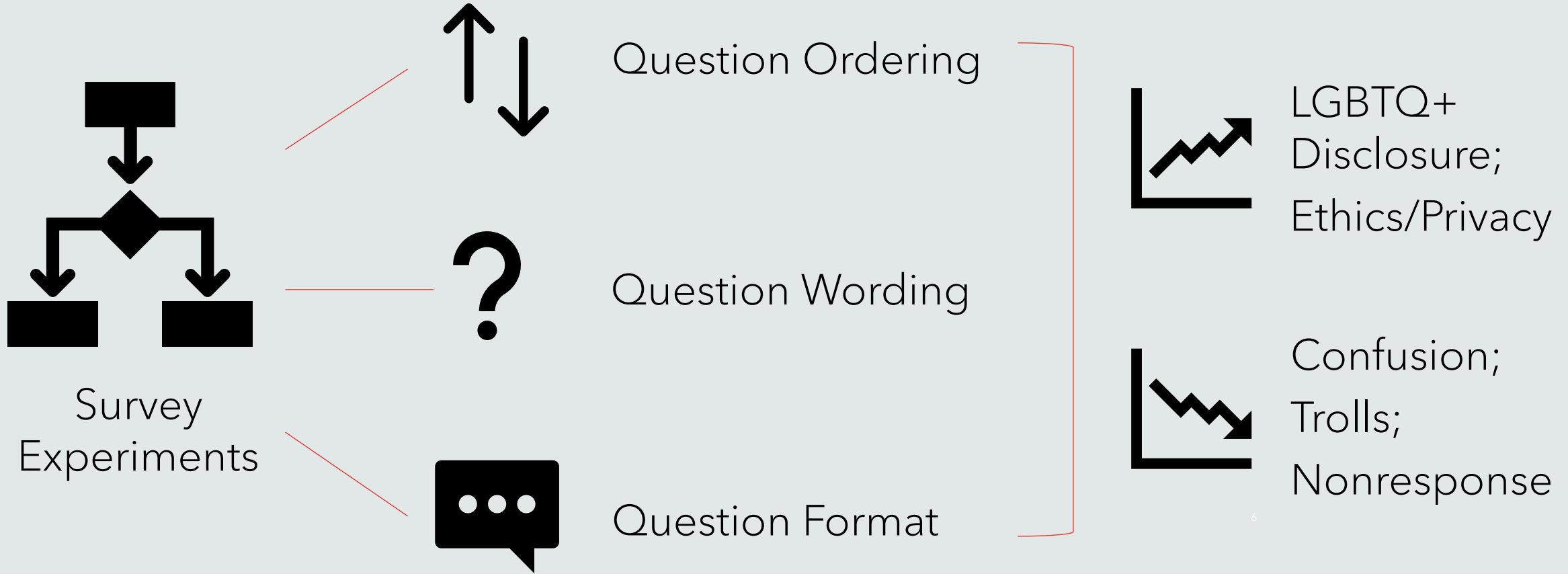| What is your sex? Note: A questions about your gender will follow if you are aged 16 or over (Female/Male) | Is your gender the same as the sex you were registered at birth? Note: This question is voluntary (Yes; No, write in gender) |

No consensus on queering the Census

	England and Wales 2021 Census	Scotland 2021 Census	United States 2018 GSS
Gender Identity	Is the gender you identify with the same as your sex registered at birth?	Do you consider yourself to be trans, or have a trans history?	What is your current gender?

	England & Wales + Scottish Censuses	New Zealand 2023 Census	United States 2018 GSS
Sexual Orientation	"Straight/heterosexual"	"Heterosexual/straight"	"Straight, that is, not gay"
	"Other sexual orientation, please write in:"	"Another identity. Please state below:"	No "other" option

Measurement Debate

Pros	Cons
Great inclusivity of minority identities (representation)	Need to determine which identities are measured (still exclusive)
New data to understand emerging social realities	Data harmonization to understand longstanding inequalities
Better counts of minority populations	Potentially worse counts for majority populations
Detailed data	Privacy, undercounts

Survey Experiments

Question Ordering

Question Wording

Question Format

LGBTQ+ Disclosure; Ethics/Privacy

Confusion; Trolls; Nonresponse

MY INTERVENTION
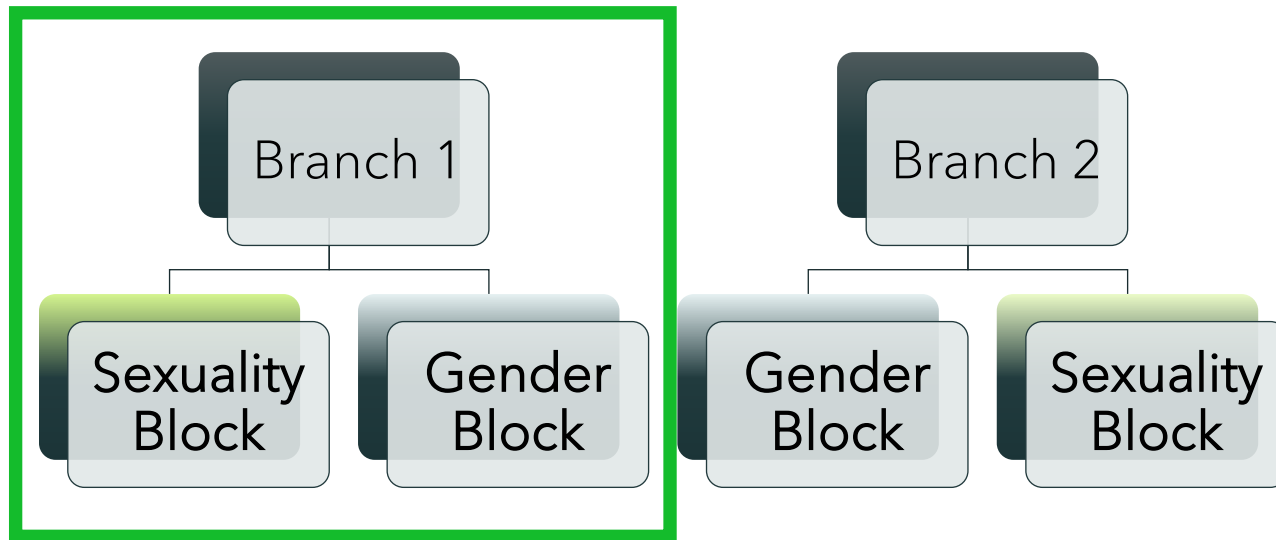
Empirics

Data	Methods
• COVID States (2022) (n = 20,743) • Original Pilot Survey (2021) (n = 2,518)	• Survey Experiments • OLS / Corrected T-Tests

Does the block order affect disclosure and nonresponse patterns?

Branch 1
Sexuality Block
Gender Block

Branch 2
Gender Block
Sexuality Block

Data: Pilot Study

Finding 1: Sexuality before gender increased noncisgender disclosure. (~1 percentage point)

Finding 2: Ordering had no significant difference on nonresponse.

8

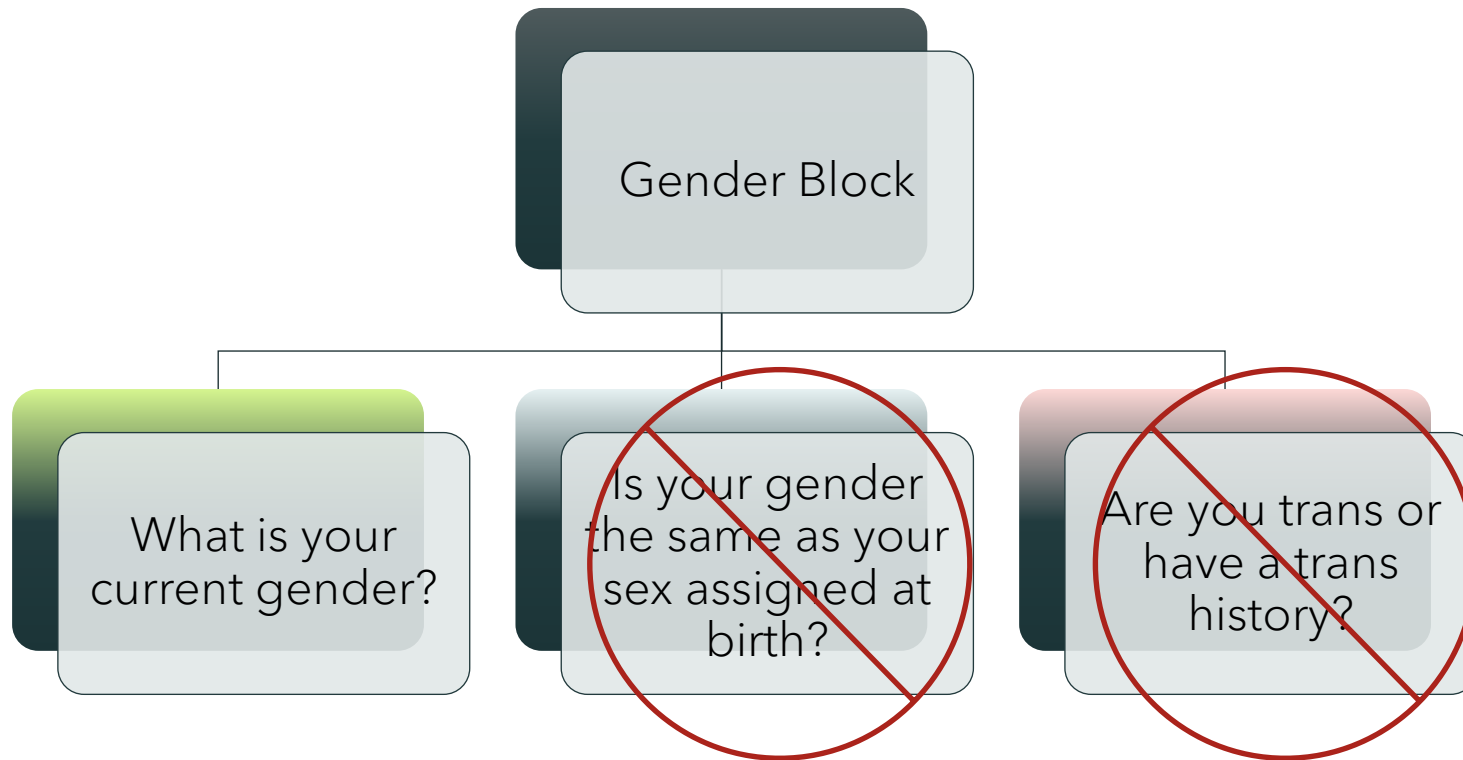# Disclosure by Block Randomization (Pilot Study)

Noncisgender Disclosure

T-Statistic	P-Value	Mean Difference	Lower CI	Upper CI
2.63	0.009	0.01	0.003	0.02

Sexuality before gender = 1 percentage point increase in noncisgender disclosure

Nonresponse

T-Statistic	P-Value	Mean Difference	Lower CI	Upper CI
.06	0.953	0.0003	-0.011	0.012

Does the wording of the two-step gender question affect disclosure and nonresponse?

Gender Block

What is your current gender?

Is your gender the same as your sex assigned at birth?

Are you trans or have a trans history?

Data: COVID States Project

Finding 3: Asking for gender in relation to sex significantly *reduced* noncisgender disclosure. (~.5 percentage points)
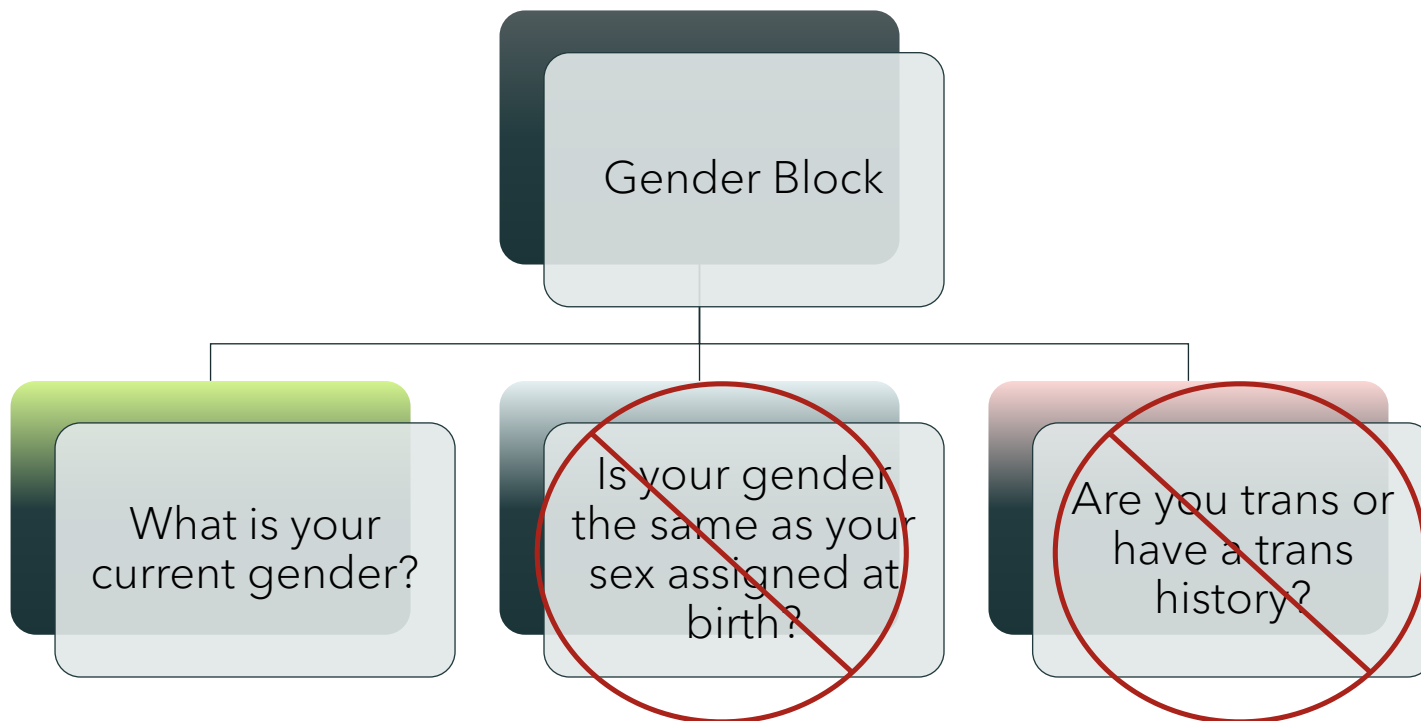
Finding 4: Asking about gender in relation to sex and trans status significantly *increased* nonresponse. (~.2 percentage points)

Disclosure by Two-Step (COVID States)

	Cisgender	Noncisgender
"What is your gender?"	6791 (0.980)	142 (0.020)
"Is your gender the same as your sex?"	6759 (0.985)	104 (0.015)
"Are you trans/have a trans history?"	6743 (0.982)	117 (0.017)

	Response	Nonresponse
"What is your gender?"	6897 (0.994)	45 (0.006)
"Is your gender the same as your sex?"	6823 (0.987)	87 (0.013)
"Are you trans/have a trans history?"	6812 (0.989)	79 (0.011)

Does the content/frequency of the free text responses change based on the question language and order?

Gender Block

What is your current gender?

Is your gender the same as your sex assigned at birth?

Are you trans or have a trans history?

Finding 5: "What is your current gender" significantly reduced the quantity of free-text responses.
(~1 percentage point)

Finding 6: Including a non-binary option would reduce free-text responses.

Data: COVID States Project

12

Free Text: Cis Backlash & Trans Feedback

	Backlash	Feedback
	Attack Helicopter equipped with machine guns and homing missles	Mostly, but also there are few times where it's No. So I may be genderfluid?
	Created UNAPPOLOGETICALLY FEMALE by God	non-binary (also "trans status" and "trans history" are societally incorrect and offensive. please simply say "do you consider yourself to be trans, or have a non conforming identity" followed by "yes, please describe your preferred gender"

Free-Text Disclosure (COVID States)

	No Text	Free Text
"What is your gender?"	6930 (0.998)	12 (0.002)
"Is your gender the same as your sex?"	6830 (0.988)	80 (0.012)
"Are you trans/have a trans history?"	6797 (0.986)	94 (0.014)

	Non-troll	Troll
"What is your gender?"	12 (0.583)	5 (0.417)
"Is your gender the same as your sex?"	80 (0.938)	5 (0.062)
"Are you trans/have a trans history?"	94 (0.947)	5 (0.053)
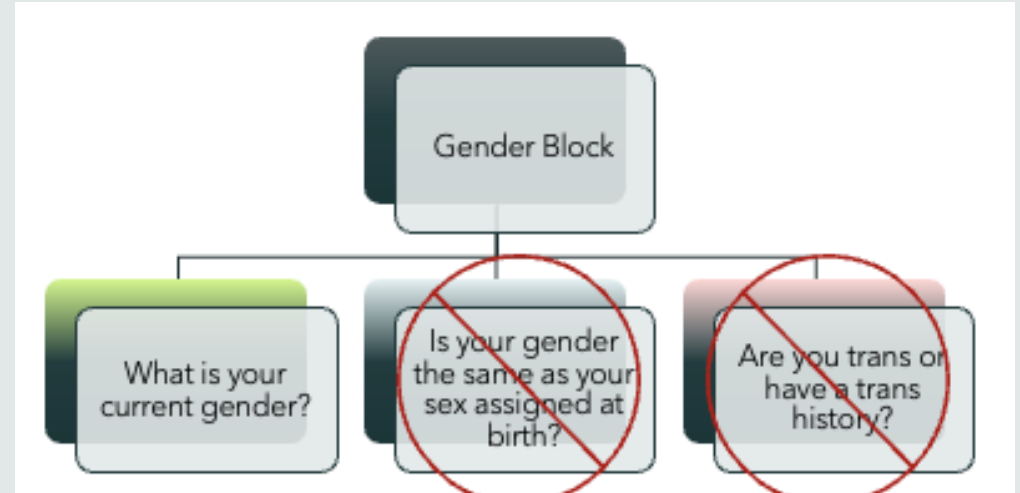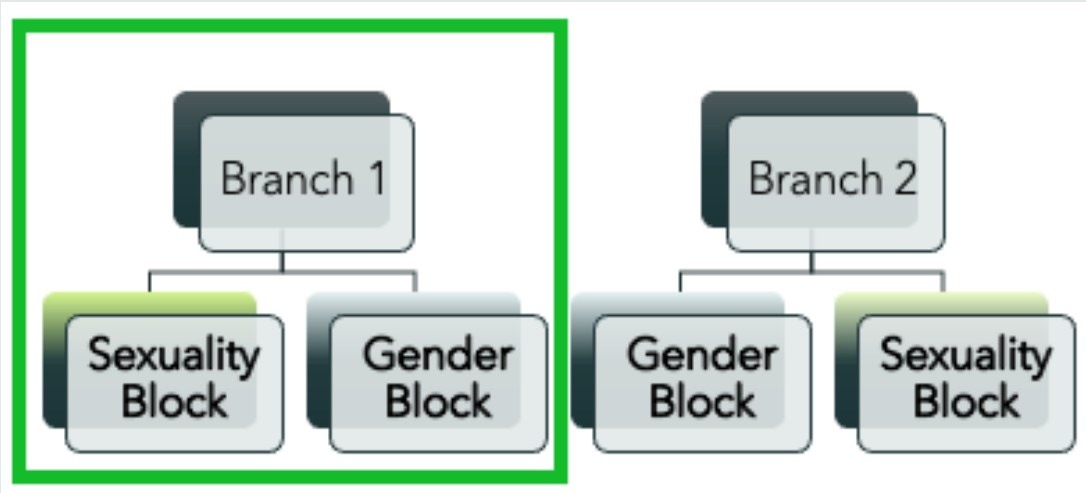
1. Ask sexuality before gender (with caveats)

2. "What is your current gender?" with a non-binary response option

Summary / Main Findings

Thank you!

Email: christina.pao@princeton.edu

Site: https://christina.pao.name/

Funding and data from:

- Nuffield College (Research Grant FE-21-7)
- Rhodes Trust (Warden's Discretionary Fund WDF MT21)
- COVID States Project Module Competition (https://www.covidstates.org/)

Much gratitude to:

- Kristopher Velasco and Ridhi Kashyap for advising
- Laurel Westbrook, Sanyu Mojola, and the OPR Works-in-Progress attendees for feedback on earlier paper drafts & presentation iterations

PRINCETON UNIVERSITY

DEPARTMENT OF SOCIOLOGY

OPR
OFFICE OF POPULATION RESEARCH

Appendix Slides

Two-Step Gender Question Format

Q2.1

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

○ Male

○ Female

○ Other

○ Prefer not to say

Q3.2

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

○ Yes

○ No (You may enter your gender identity below)
[   ]

○ Prefer not to say

Q2.2

What is your current gender?

○ Woman

○ Man

○ Non-binary / third gender

○ A gender not listed here
[   ]

Q4.2

Do you consider yourself to be trans, or have a trans history?

○ No

○ Yes, please describe your trans status (for example, non-binary, trans man, trans woman)
[   ]

○ Prefer not to say

Sexuality Question Format

☐ **Q1.1**

Which of the following best describes you?

○ Gay, lesbian, or homosexual

○ Bisexual

○ Heterosexual or straight

○ 🚫 Don't know

○ Prefer not to say

Block Randomization (Pilot Study)

Differences in Non-Straight Disclosure with Sexuality and Gender Block Randomization

	T-Statistic	Degrees of Freedom	P-Value	Bonferroni-Corrected P-Value	Mean Difference	Lower CI	Upper CI
Limited Sexuality Question	1.43	2500.9	0.15	0.80	0.02	-0.01	0.04
Expansive Sexuality Question	1.25	2513.4	0.21	1	0.02	-0.01	0.04

Differences in Non-Cisgender Disclosure with Sexuality and Gender Block Randomization

T-Statistic	Degrees of Freedom	P-Value	Bonferroni-Corrected P-Value	Mean Difference	Lower CI	Upper CI
2.63	2047.5	0.009	0.04	0.01	0.003	0.02

Two-Step Randomization (Pilot Study)

- With sample size, tested only: "what is your gender?" and "is your gender the same as your sex?"
- No significant differences in non-cis disclosure across the questions
- Asking about gender the same as sex significantly increased sex and gender nondisclosure

Tl;dr: We knew we needed larger N to test for differences in non-cisgender disclosure.

Non-straight Disclosure (COVID States)

	Missing	Gay	Bi	Straight	DK	PNTS
What is your gender?	4	246	514	5972	85	121
Sex = gender?	1	260	526	5933	64	126
Trans Status	5	255	467	5965	83	116

```
        Pearson's Chi-squared test

data:  df$twostep_merge and df$sexuality
X-squared = 10.72, df = 10, p-value = 0.3798
```

- Data challenges with randomization in COVID States