

 NYU | STEINHARDT

How Invalid and Mischiefous Survey Responses Bias Estimates of LGBTQ-Heterosexual Youth Risk Disparities

Joseph Cimpian
March 16, 2022



LGBQ-heterosexual youth relative risk disparities (YRBS 2017)

Outcome	Males (<i>N</i> = 6,778; 8.5% LGBQ)	Females (<i>N</i> = 7,211; 20.4% LGBQ)
	Unadjusted RR	Unadjusted RR
57 - IV drug use (ever)	4.32***	7.09***
55 - Steroid use (ever)	3.27***	2.59**
51 - Heroin use (ever)	5.86***	4.24***
52 - Methamphetamine use (ever)	4.95***	3.99***
28 - Attempted suicide	4.11***	3.03***
15 - Skipped school for safety	2.25***	1.32
27 - Made a suicide plan	3.16***	2.92***
24 - Bullying online	2.35***	1.47***
25 - Feeling sad or hopeless	2.15***	1.76***
26 - Considered suicide	3.14***	2.79***
23 - Bullying at school	2.10***	1.49***

But... are the data valid?

Adoptee ex. (Triangulation)

Students given a SAQ (464 “adoptees”)

Then, **parents** asked:

“Is [child’s name] your biological child?”

“Does at least 1 biological parent live with [child’s name]?”

Results:

88 of 464 “adoptees” misreported (**19%**)

<i>Outcome Variable</i>	<i>True Adoptees (n = 376) vs. Nonadoptees (n = 14,662)</i>	<i>False Adoptees (n = 88) vs. Nonadoptees (n = 14,662)</i>
School grades (+)	-0.11	-0.24
School troubles	-0.10	0.53
Positive school feelings (+)	0.05	-0.99
Skipping school	0.02	1.97
Smoking	0.05	0.94
Drinking	-0.03	1.45
Drunk	-0.02	1.68
Self-esteem (+)	-0.01	-0.96
Emotional distress	0.12	1.02
Future hope (+)	-0.01	-1.36
Health problems	0.06	1.08
Physical problems	0.06	1.73
Sickness	-0.08	1.15
Fighting	0.05	1.36
Lie to parents	-0.11	0.84
Mean of absolute effect size	0.06	1.09

This is *not* solely an AddHealth concern...

Bigger problem: most studies do not have a clear way to assess the extent of the concern

“Mischievous responder”

Truth

Not adopted

Heterosexual

Cisgender

Suicidal: **No**

Skips school: Yes

2+ children: No

Blind: No

Deaf: No

In a gang: No

Very short/tall: No

Not seen dr in 5+ yrs: No



What he reported

“Funny”

Adopted

Gay

Transgender

Suicidal: **Yes**

Skips school: Yes

2+ children: Yes

Blind: Yes

Deaf: Yes

In a gang: Yes

Very short/tall: Yes

Not seen dr in 5+ yrs: Yes

Item (and low-frequency response)	“Cisgender” (N = 11,625)	“Trans” (N = 204)
1. Provided a height in the top or bottom 2.5%	3.7%	41.7%
2. Provided a weight in the top or bottom 2.5%	4.1%	30.4%
3. Are you deaf or have a hearing impairment? (Yes)	1.3%	25.5%
4. Are you blind or have vision impairment? (Yes)	3.1%	31.4%
5. When was the last time you visited a dentist? (3 or more years ago)	3.5%	30.9%
6. How many times have you been pregnant or have gotten a girl pregnant? (2 or more times)	0.7%	25.0%
7. How many children do you have? (2 or more)	1.3%	23.0%
8. Is one or more of your family members in a gang? (Yes)	3.2%	41.7%
9. Are you in a gang? (Yes, currently)	1.9%	35.3%
10. In the past month, how many days have you carried a weapon to school? (6 or more days)	1.2%	28.4%

Item (and low-frequency response)	“Heterosexual” (N = 11,058)	“LGBQ” (N = 771)
1. Provided a height in the top or bottom 2.5%	3.7%	13.5%
2. Provided a weight in the top or bottom 2.5%	4.0%	13.1%
3. Are you deaf or have a hearing impairment? (Yes)	1.3%	8.9%
4. Are you blind or have vision impairment? (Yes)	2.9%	13.9%
5. When was the last time you visited a dentist? (3 or more years ago)	3.4%	12.8%
6. How many times have you been pregnant or have gotten a girl pregnant? (2 or more times)	0.7%	7.8%
7. How many children do you have? (2 or more)	1.3%	6.6%
8. Is one or more of your family members in a gang? (Yes)	3.1%	15.2%
9. Are you in a gang? (Yes, currently)	1.9%	11.3%
10. In the past month, how many days have you carried a weapon to school? (6 or more days)	1.1%	9.2%

Number of low-freq. responses

	"Cisgender" (N = 11,625)			"Transgender" (N = 204)	
0	9,755	83.91	71	34.80	
1	1,408	12.11	28	13.73	
2	287	2.47	25	12.25	
3	73	0.63	7	3.43	
4	31	0.27	7	3.43	
5	18	0.15	10	4.90	
6	13	0.11	9	4.41	
7	19	0.16	9	4.41	
8	12	0.10	14	6.86	
9	4	0.03	7	3.43	
10	5	0.04	17	8.33	

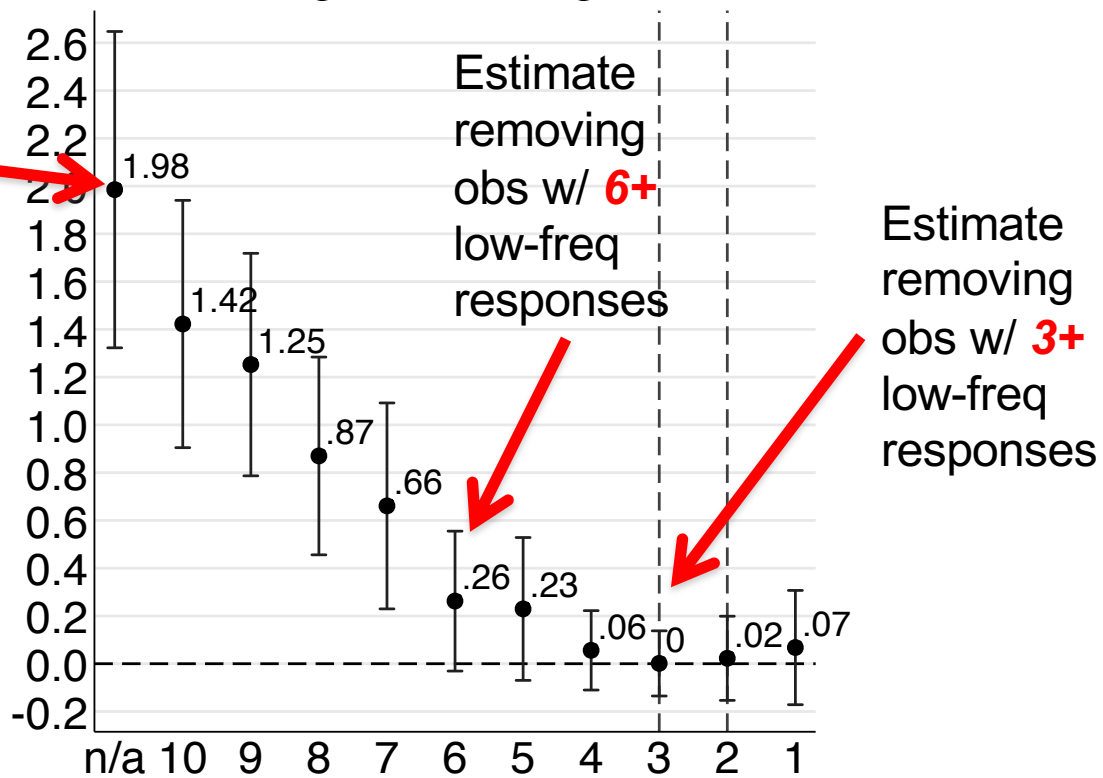
Extreme suicidal ideation

Estimate with **full** sample

Reporting thinking about suicide almost all the time

(shown here in SDs)

Transgender-Cisgender



Number of low-frequency responses required for exclusion

Different methods for validity-weighting

	Count-based	Boosted reg max. classification	Boosted reg min. screener differences	
Requires screener items	Y	Y	Y	<i>Human</i>
Requires screener responses specified	Y			
Simple	Y		ish	
Lots of possible estimates	Y	Y		
Researcher specified combinations	Y			
Data-driven interactions		Y	Y	<i>Data-driven, more flexible</i>
Missing data handled easily		Y	Y	
Matching/Reweighting-based			Y	
	Robinson & Espelage, <i>ER</i> , 2011; Robinson-Cimpian, <i>ER</i> , 2014	Cimpian et al., <i>AJPH</i> , 2018; Cimpian & Timmer, <i>AERA Open</i> , 2019	Cimpian, Timmer, & Kim, R&R	
Registered report (pre-specified hypotheses and methods)		Part (national); Y	Y	

Current Study: Registered Report

With the 2019 YRBS, combine a machine-learning approach (boosted regression) with propensity-score reweighting to **minimize the likelihood of potentially invalid responses** between LGBTQ and heterosexual youth responses.

Flexible: Can be mischievousness, inattention, satisficing, combination...

2017 YRBS shown here as “**pilot**” results. No analyses with 2019 YRBS yet.

Cimpian et al., *in-principle accept*

Research questions

1. What are the LGBTQ-heterosexual youth relative risks (RRs) for a wide range of outcomes, and **are the RRs robust** to adjustments for potentially invalid responses?
2. Looking across outcomes, **are low-incidence outcomes less robust** to adjustments?
3. Is the validity-checking **process itself robust** to researcher choices regarding how to identify potentially invalid responses?
4. How does this new method **compare with data-removal** boosted reg?
5. Does coding risk **dichotomously** (vs. **continuously**) diminish the effect of mischievous and invalid responses?

Process/Items used in screener

Perhaps counterintuitive, but...

Use items *theoretically unrelated* to LGBTQ status to predict reporting “LGBTQ” on survey

- Carrot eating
- Salad eating
- Fruit eating
- Potato eating
- Height
- Asthma
- Dentist visits
- Survey weight

Process/Items used in screener

Perhaps counterintuitive, but...

Use items *theoretically unrelated* to LGBTQ status to predict reporting “LGBQ” on survey

- Carrot eating
- Salad eating
- Fruit eating
- Potato eating
- Height
- Asthma
- Dentist visits
- Survey weight

Assess robustness by *removing* these items

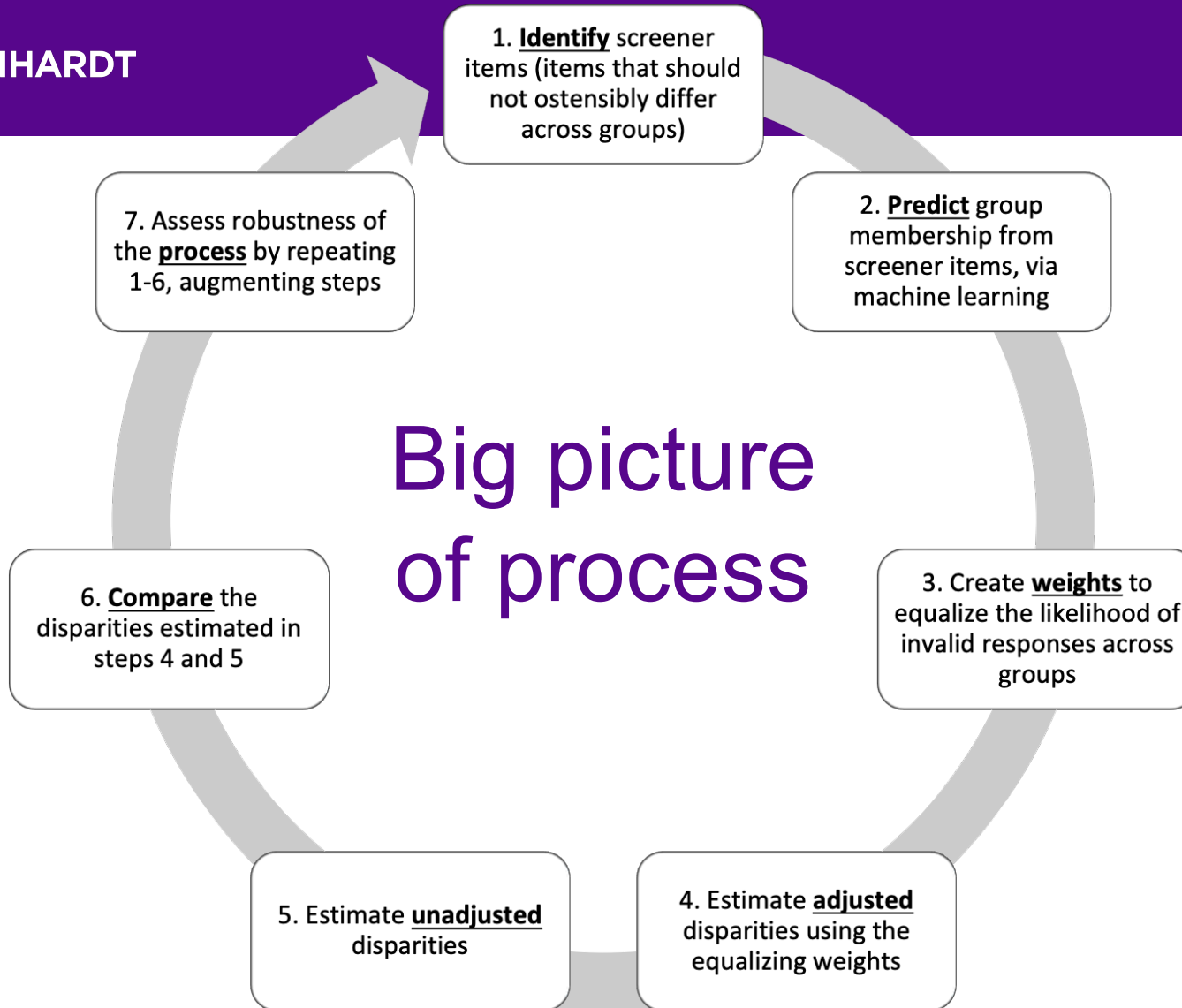
Process/Items used in screener

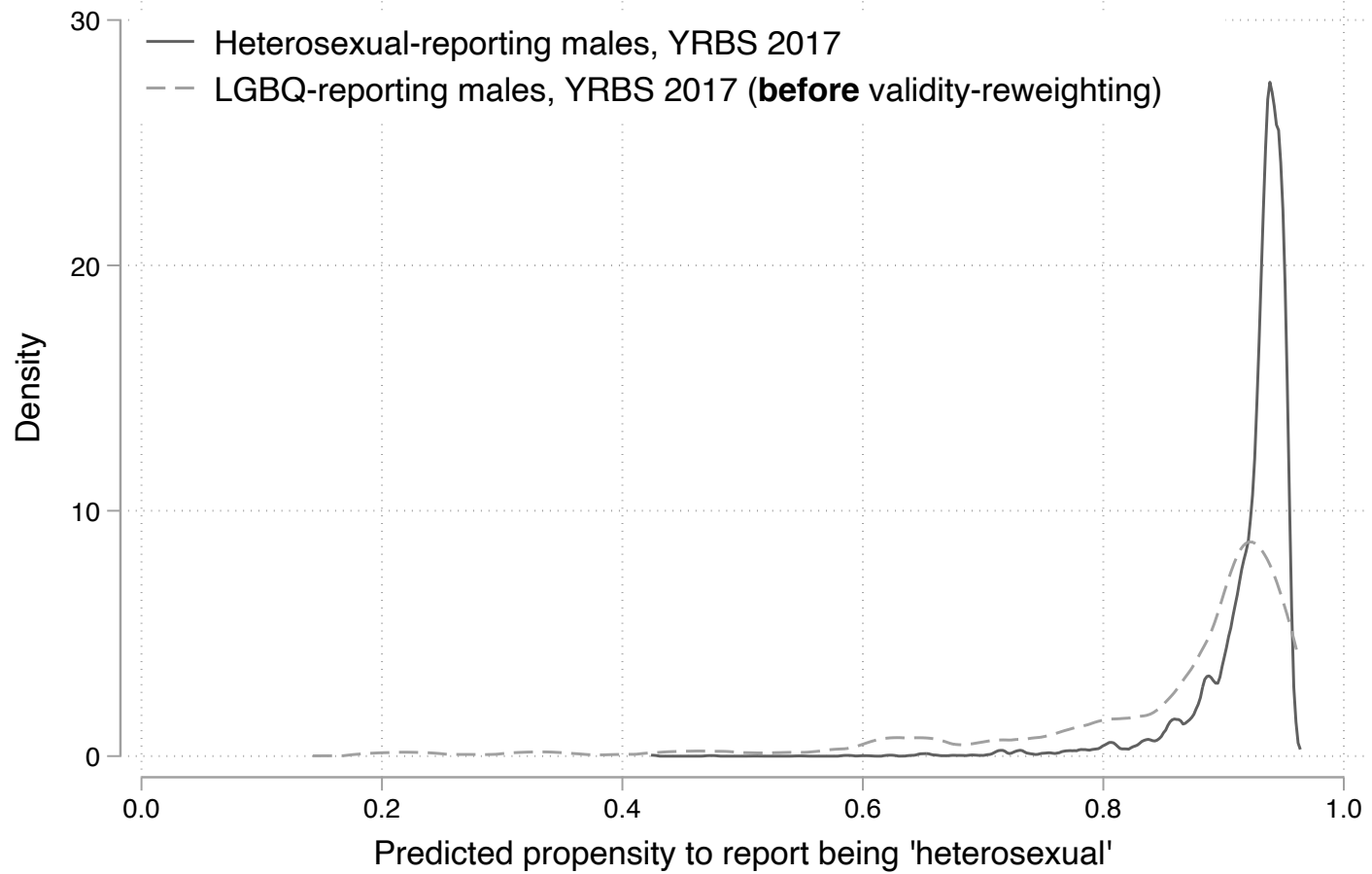
Perhaps counterintuitive, but...

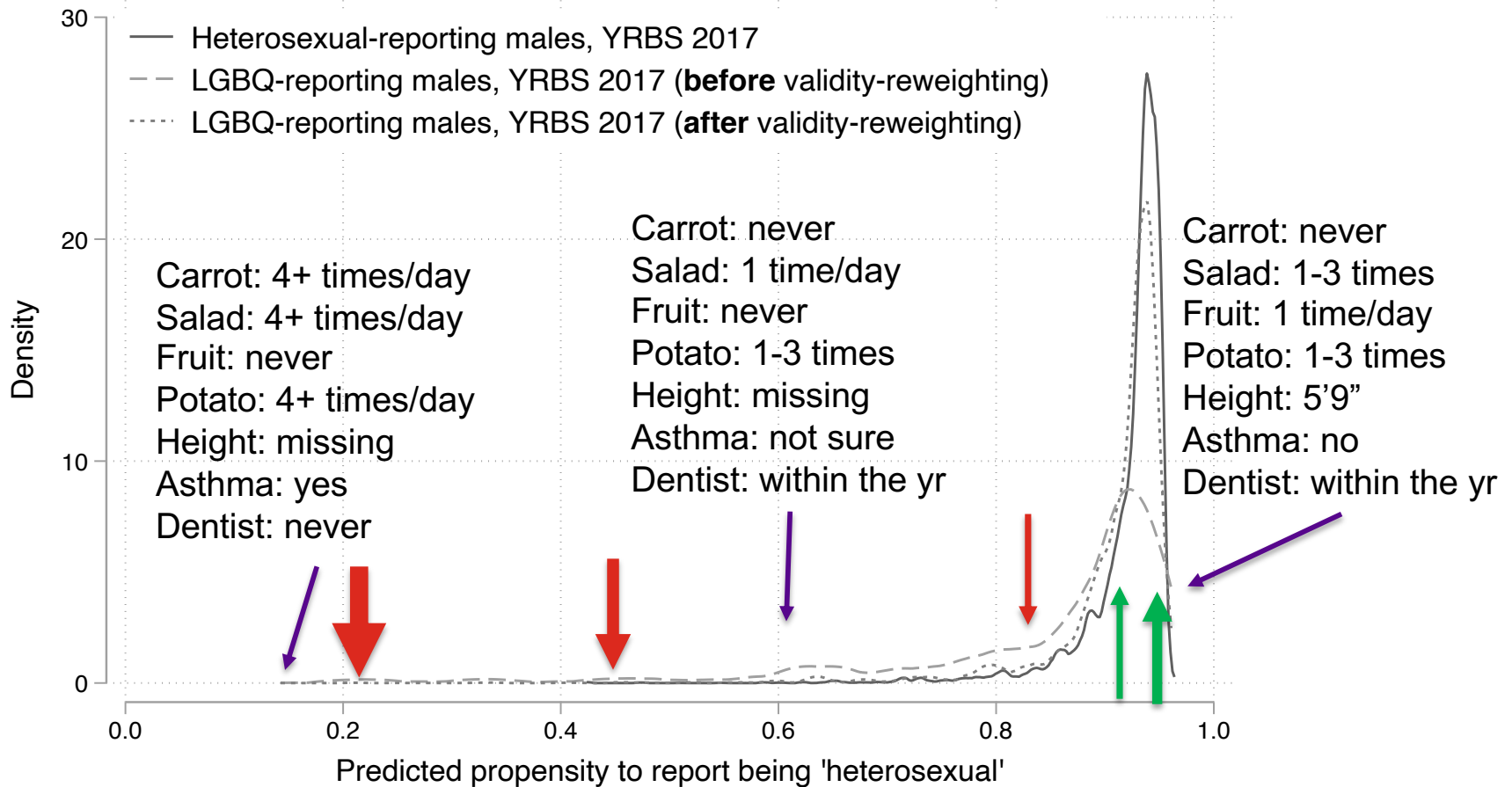
Use items *theoretically unrelated* to LGBTQ status to predict reporting “LGBQ” on survey

- Carrot eating
- Salad eating
- Fruit eating
- Potato eating
- Height
- Asthma
- Dentist visits
- Survey weight
- Bullying
- Heroin use

Assess robustness by *adding* these likely invalid screener items







LGBQ-heterosexual youth relative risk disparities (YRBS 2017)

Outcome	Males (<i>N</i> = 6,778; 8.5% LGBQ)	Females (<i>N</i> = 7,211; 20.4% LGBQ)
	Unadjusted RR	Unadjusted RR
57 - IV drug use (ever)	4.32***	7.09***
55 - Steroid use (ever)	3.27***	2.59**
51 - Heroin use (ever)	5.86***	4.24***
52 - Methamphetamine use (ever)	4.95***	3.99***
28 - Attempted suicide	4.11***	3.03***
15 - Skipped school for safety	2.25***	1.32
27 - Made a suicide plan	3.16***	2.92***
24 - Bullying online	2.35***	1.47***
25 - Feeling sad or hopeless	2.15***	1.76***
26 - Considered suicide	3.14***	2.79***
23 - Bullying at school	2.10***	1.49***

Male LGBTQ-heterosexual youth RR disparities (YRBS 2017)

Outcome	Males (<i>N</i> = 6,778; 8.5% LGBTQ)		
	Unadjusted RR	Adjusted RR	Reduction in RR
57 - IV drug use (ever)	4.32***	1.68	-61%***
55 - Steroid use (ever)	3.27***	1.32	-60%***
51 - Heroin use (ever)	5.86***	2.61**	-56%***
52 - Methamphetamine use (ever)	4.95***	2.40**	-51%***
28 - Attempted suicide	4.11***	2.65***	-36%***
15 - Skipped school for safety	2.25***	1.54	-31%***
27 - Made a suicide plan	3.16***	2.73***	-14%
24 - Bullying online	2.35***	2.12***	-10%
25 - Feeling sad or hopeless	2.15***	2.02***	-6%
26 - Considered suicide	3.14***	2.96***	-6%
23 - Bullying at school	2.10***	2.07***	-1%

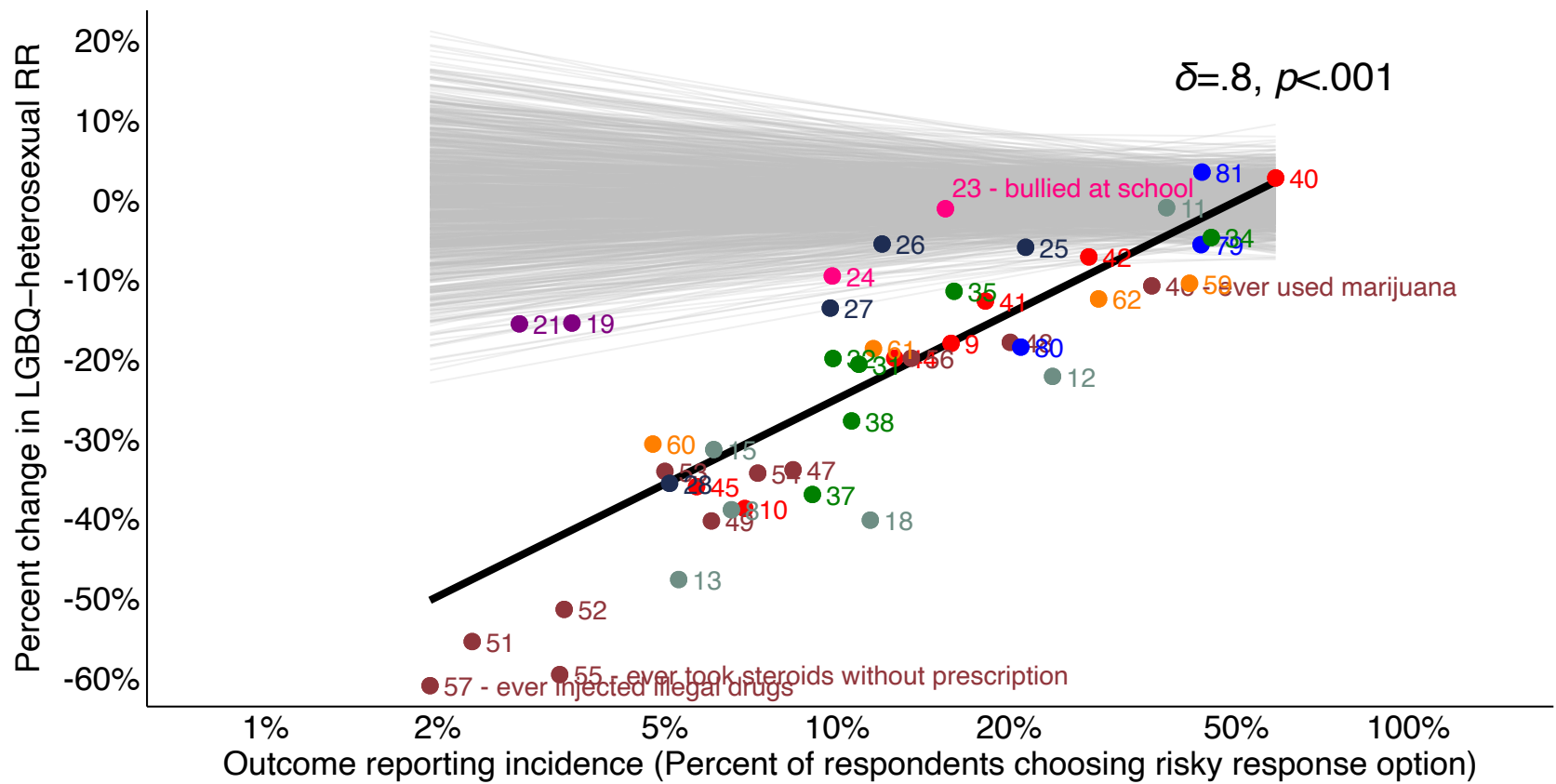
Female LGBTQ-heterosexual youth RR disparities (YRBS 2017)

Outcome	Females (<i>N</i> = 7,211; 20.4% LGBTQ)		
	Unadjusted RR	Adjusted RR	Reduction in RR
57 - IV drug use (ever)	7.09***	5.33***	-25%
55 - Steroid use (ever)	2.59**	1.97*	-24%
51 - Heroin use (ever)	4.24***	2.67**	-37%
52 - Methamphetamine use (ever)	3.99***	2.98***	-25%
28 - Attempted suicide	3.03***	2.88***	-5%
15 - Skipped school for safety	1.32	1.28	-3%
27 - Made a suicide plan	2.92***	2.80***	-4%
24 - Bullying online	1.47***	1.47***	1%
25 - Feeling sad or hopeless	1.76***	1.72***	-2%
26 - Considered suicide	2.79***	2.70***	-3%
23 - Bullying at school	1.49***	1.43***	-4%

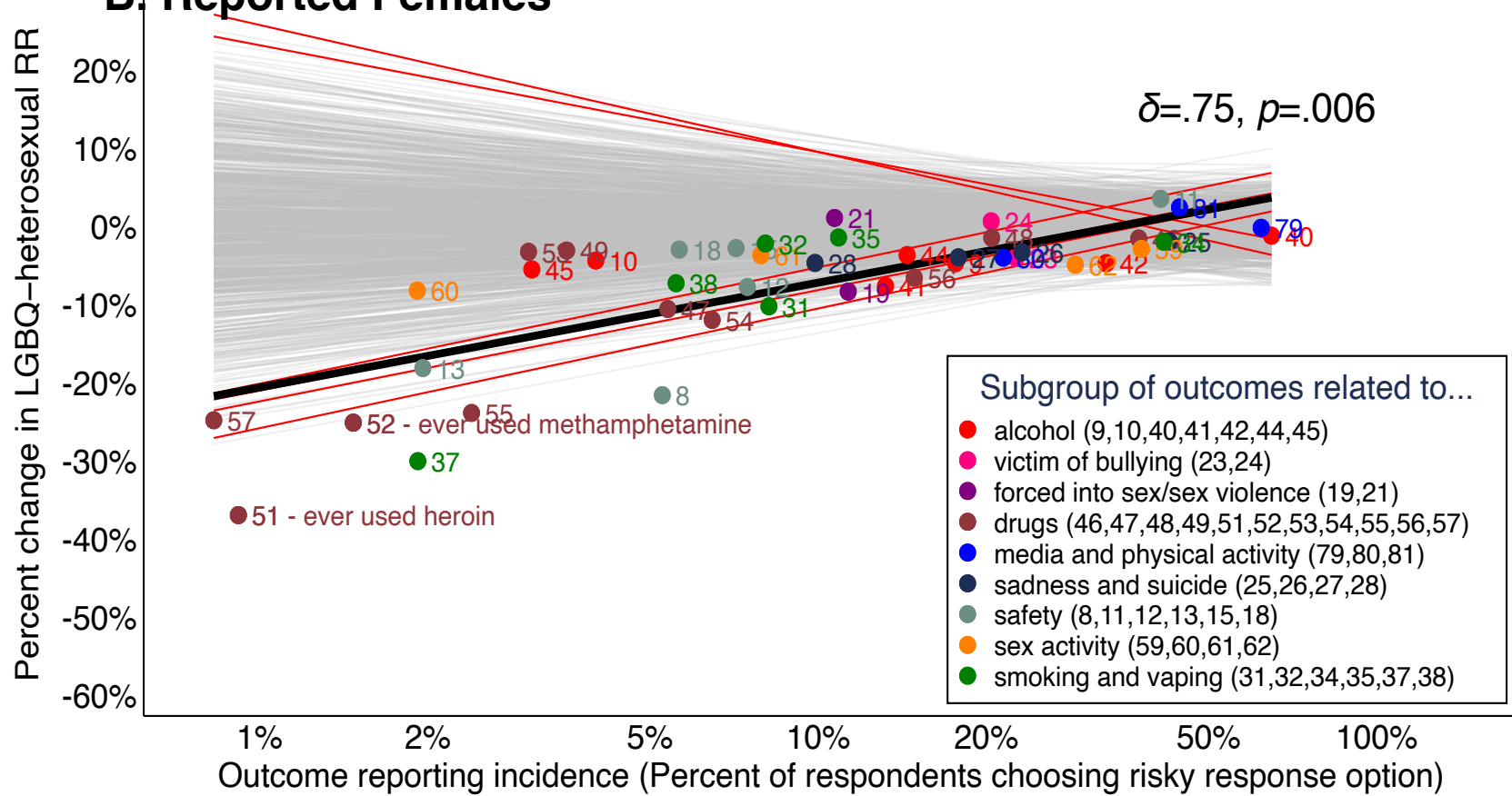
What *predicts* outcome variation?

Item response-option extremity?

A. Reported Males



B. Reported Females



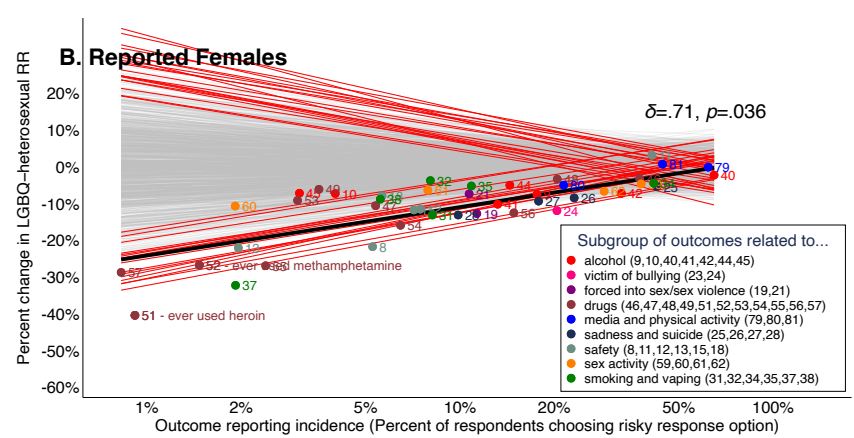
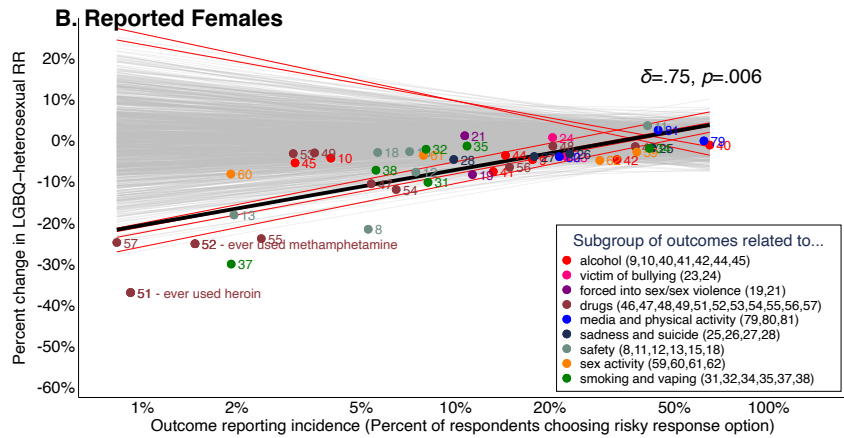
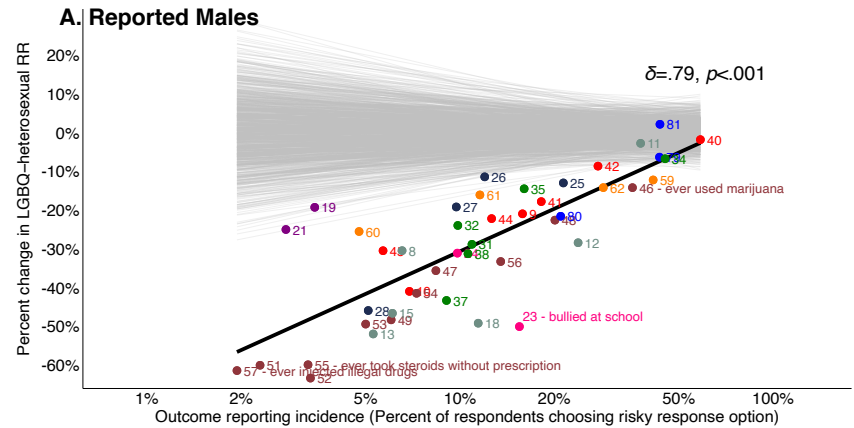
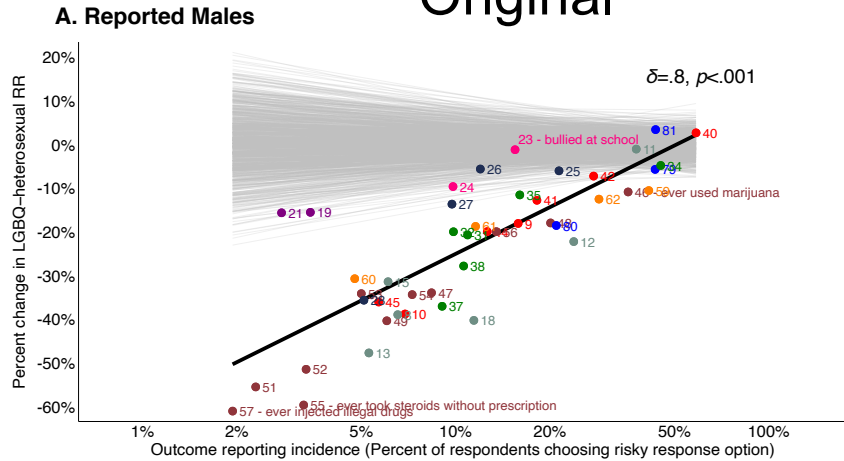
Are the patterns *robust* to screener alterations?

Broadly, *yes*. Some *differences* are specific to items included (as expected).



Original

+ likely invalid (bullied & heroin)

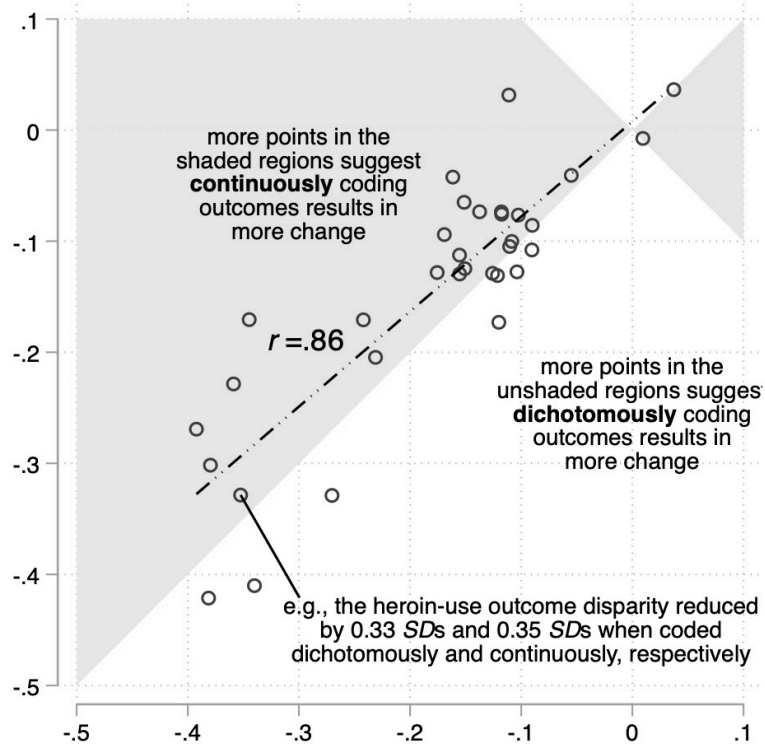


Does coding risk *dichotomously* (vs. *continuously*) diminish the effect of mischievous and invalid responses?

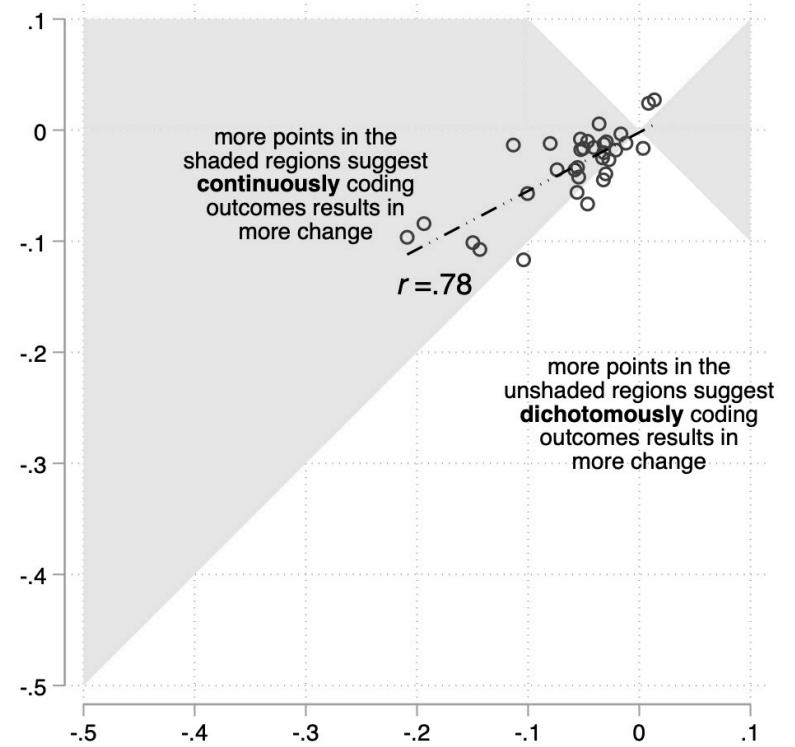


A. Reported Males

Standardized difference between models when coding outcomes dichotomously



B. Reported Females



Standardized difference between models when coding outcomes continuously

Implications

Don't trust teenagers??? 😊

Published findings (even with $p < .001$) may not be valid

Researchers should conduct robustness checks to assess invalid and mischievous response bias

- Estimate a **series** of disparities, not just a single “true” (or “best”) estimate, and assess robustness/stability

Robust research is needed for policymaking, and the explosion of LGBTQ research based on likely faulty data is not helping

Implications (continued)

Estimated disparities among males seem to be more affected than those among females

Not all outcomes equally affected—seems to depend on how extreme the most extreme item response-option is

- Helps explain why suicidal ideation is not affected much while suicide attempts is
- Relatedly, why suicide attempts behaves more like heroin use than suicidal ideation

Methods matter! Boosted regression approach (with data removal or reweighting) seems most efficient; regression has almost no effect

Final implications... for Practice

For males, the disparity estimates are quite sensitive...

- No validity-weighting: Male LGBTQ youth are at higher risk for ***almost all*** outcomes
- Validity-weighting: Male LGBTQ youth are at higher risk for a ***more limited set***, focused on bullying-victimization, sadness, suicidal ideation

Practitioners do not have unlimited resources

The validity-weighted results may help to focus those resources (and policies) on the more robust patterns

Questions?

joseph.cimpian@nyu.edu