

a little gender and identity study

blaise agüera y arcas

a weekend project

10/9/2016

Setup

65 question survey

Takes ~2.5 minutes to complete

Focusing on gender, sex, sexuality, identity and presentation

Goal: “data lab” to explore gendered identity and its correlates

Administered on Mechanical Turk

~3000 subjects at \$0.35 each*

Responses complete in a few hours

*Not done on Google's dime.

Questions

Body

Presentation

Attraction

Identity

Questions*

Body

Presentation

Attraction

Identity

*In the actual survey the questions are not grouped into sections. Also, lines between “nature”, “culture”, and “choice” are not always distinct, and in some cases contentious or ambiguous. This applies to “innate” orientation vs. behavior, race vs. ethnicity, and physical traits that can be surgically altered.

Questions

Body

2 numeric,
5 yes/no

Presentation

Attraction

Identity

Age

Height

Do you menstruate?

Have you ever menstruated?

Do you have a penis?

Do you have a vagina?

Have you ever been pregnant?

Questions

Body

Presentation

Attraction

Identity

41 yes/no

- Do you have long hair?
- Do you wear dresses / skirts?
- Do you wear pants?
- Do you ever wear a bra?
- Do you have long nails?
- Do you ever paint your nails?
- Do you have an ear piercing?
- Do you ever wear stockings?
- Do you ever wear high heels?
- Do you ever wear ties?
- Do you sometimes wear your hair in a ponytail?
- Do you like football?
- Do you attend a book club?
- Do you cook?
- Do you wash the dishes?
- Do you bake?
- Do you have a beard or moustache?
- Do you ever use makeup?
- Do you ever use colored lipstick?
- Do you get manicures?
- Do you get pedicures?
- Do you wax?
- Do you have callouses on your hands?
- Do you sometimes wear work boots?
- Do you ride a motorcycle?
- Do you play shooter video games?
- Do you ever wear boxer shorts?
- Do you ever wear panties?
- Do you shave your legs?
- Do you shave your armpits?
- Do you shave your face?
- Have you ever been in a fistfight?
- Do you use the men's bathroom?
- Do you use the women's bathroom?
- Do you do household repairs?
- Do you wear colorful clothes?
- On paper, do people assume from your name that you are female?
- On paper, do people assume from your name that you are male?
- Are you married?
- Do you own a gun?
- Do you hunt?

Questions

Body

Presentation

Attraction —————
4 yes/no

Identity

Are you sexually attracted to women?

Are you romantically attracted to women?

Are you sexually attracted to men?

Are you romantically attracted to men?

Questions

Body

Presentation

Attraction

Identity

13 yes/no

Do you identify as male?

Do you identify as female?

Do you identify as Black / African American?

Do you identify as Latino or Hispanic American?

Do you identify as South Asian / Indian American?

Do you identify as Native American?

Do you identify as White / European American?

Do you identify as Asian American?

Are you heterosexual or straight?*

Are you homosexual, gay or lesbian?*

Are you bisexual or pansexual?

Is the right pronoun for you “she”?

Is the right pronoun for you “he”?

Spoilers

There are no rules for identity.

We see the gender binary.

We see the gender “spectrum”.

We see that sex-gender-sexuality-orientation is in fact higher-dimensional.

Presentation predicts identity more strongly than body data does.

Body and behavior are strong predictors of gender identity, weaker predictors of other identities.

More people are intersex, trans, or gender nonconforming than you probably think.

Same-sex sexual attraction is more common than same-sex romantic attraction.

Same-sex attraction among women is extremely common.

Young people are increasingly nonconformant to either old identities or gender/sexuality binaries.

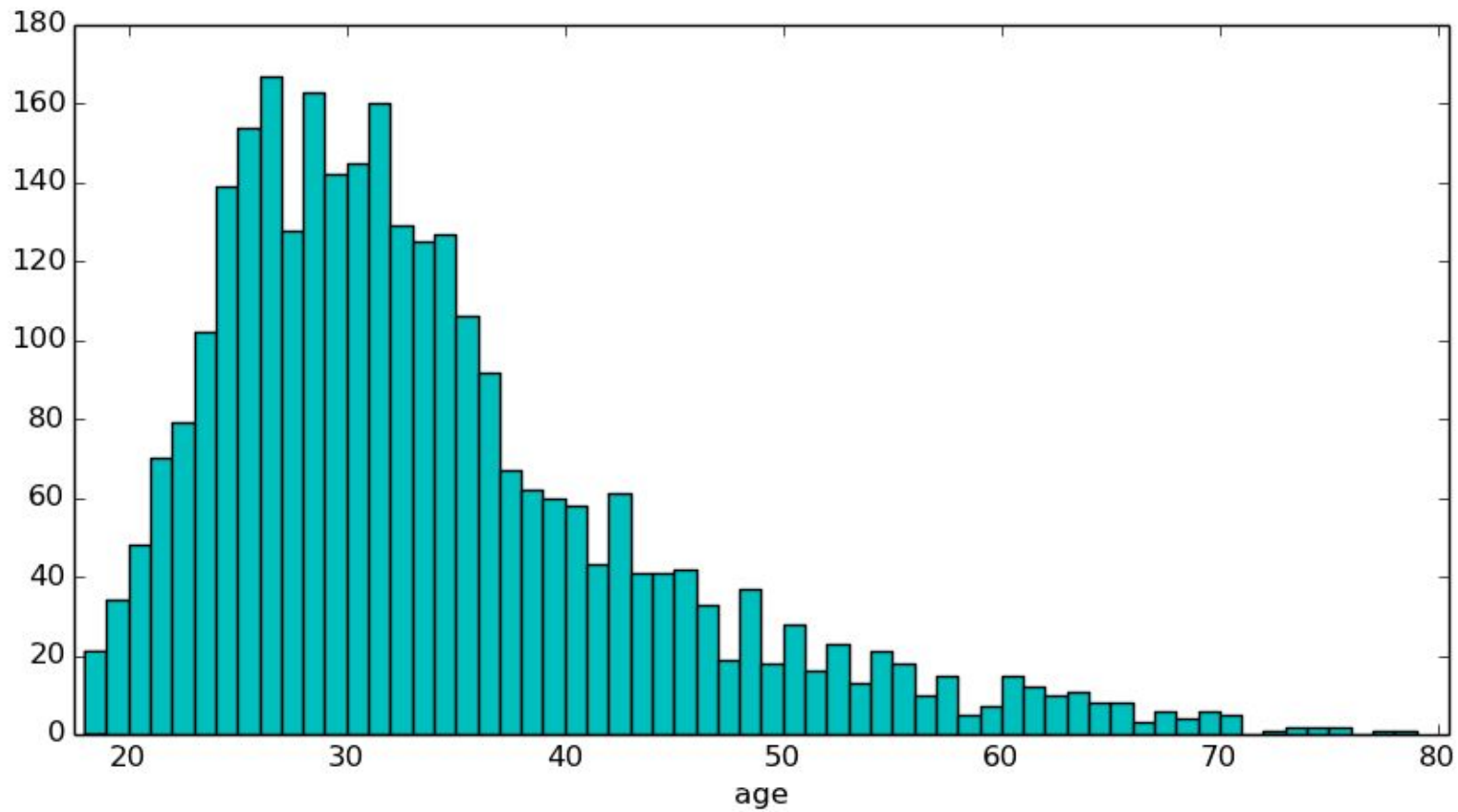
Attempts to validate identity with a body correlate will fail and do harm (cf. [sex testing in the Olympics](#)).

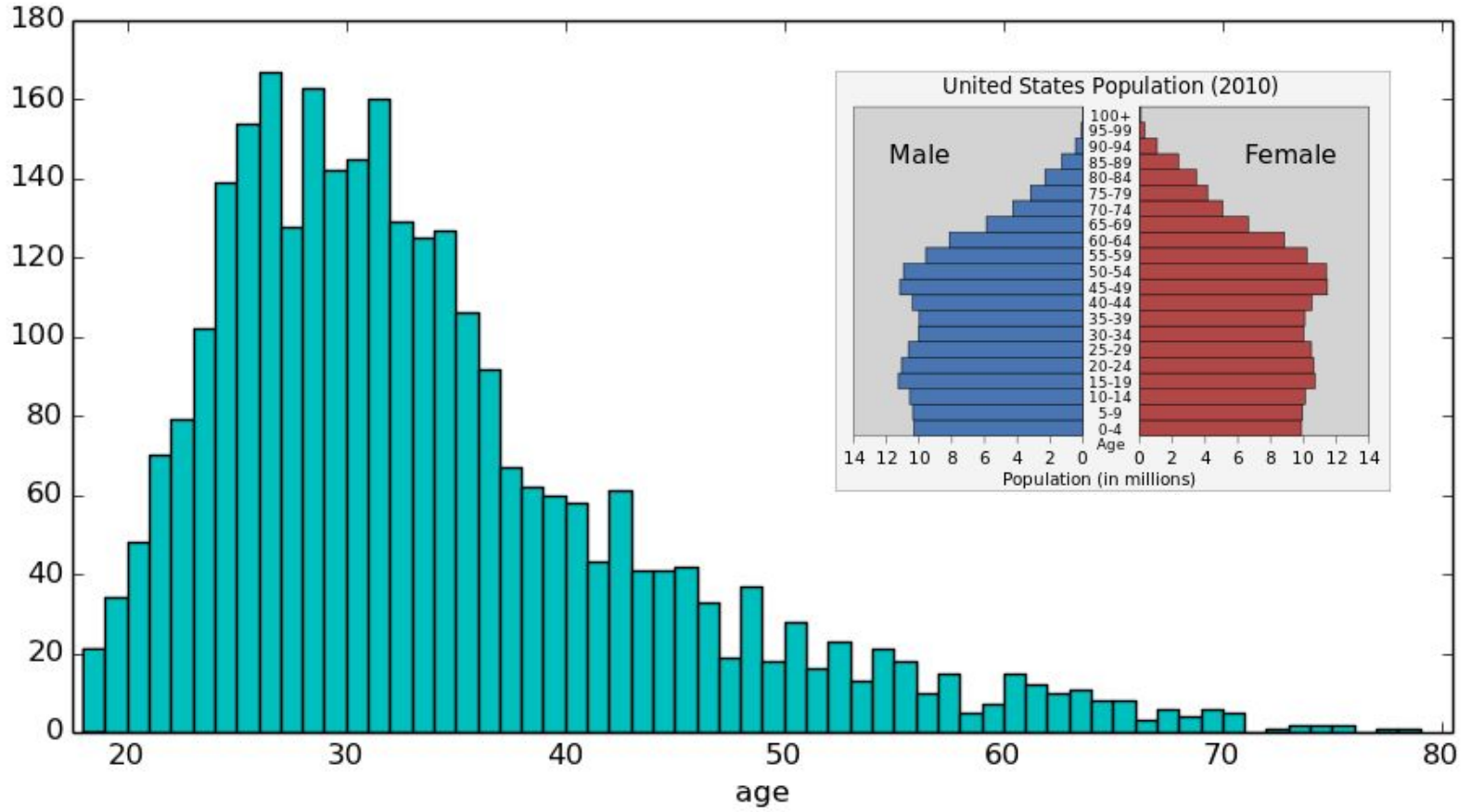
I. Who are these people?

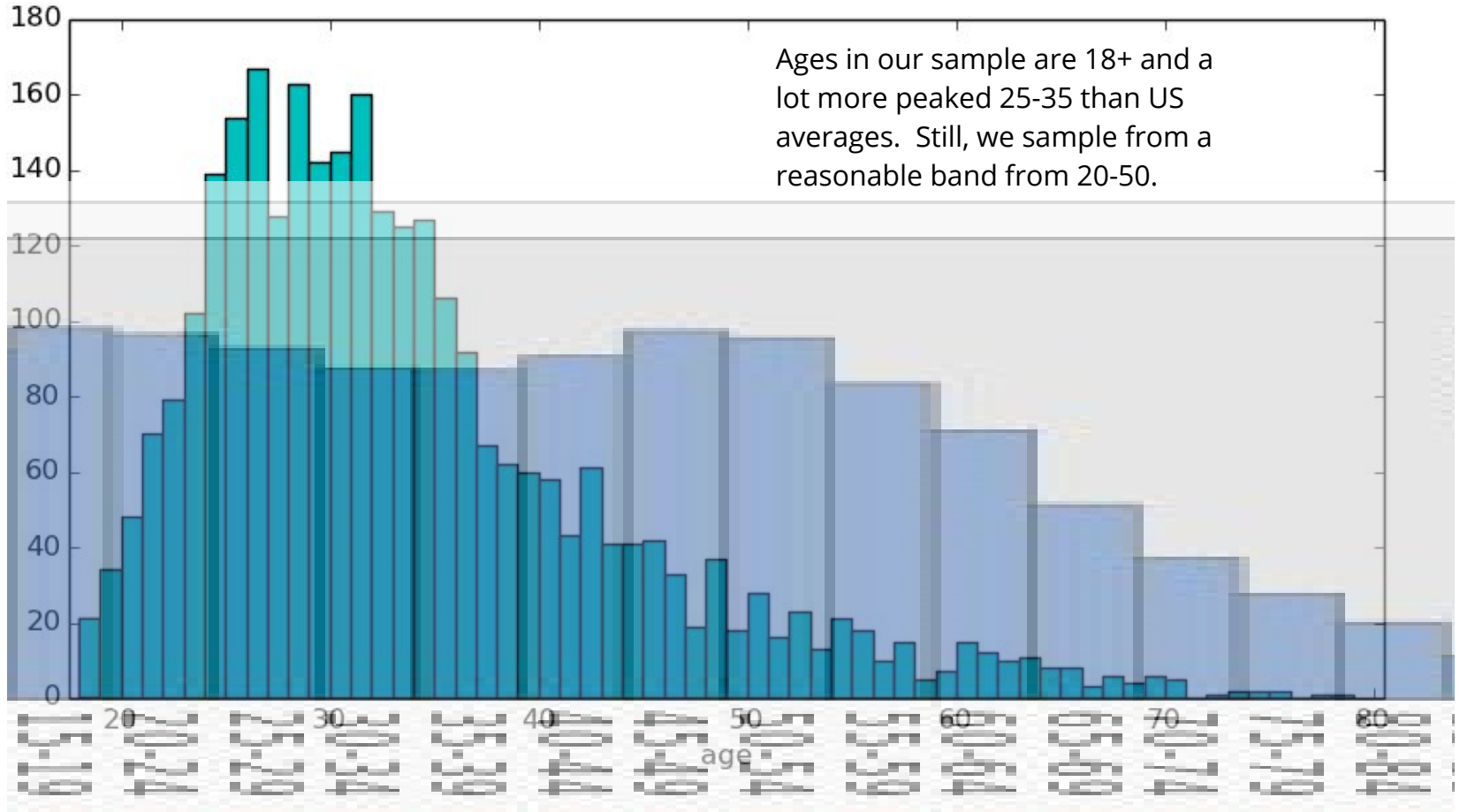
Basic findings on MTurk's utility and sampling biases available [here](#).

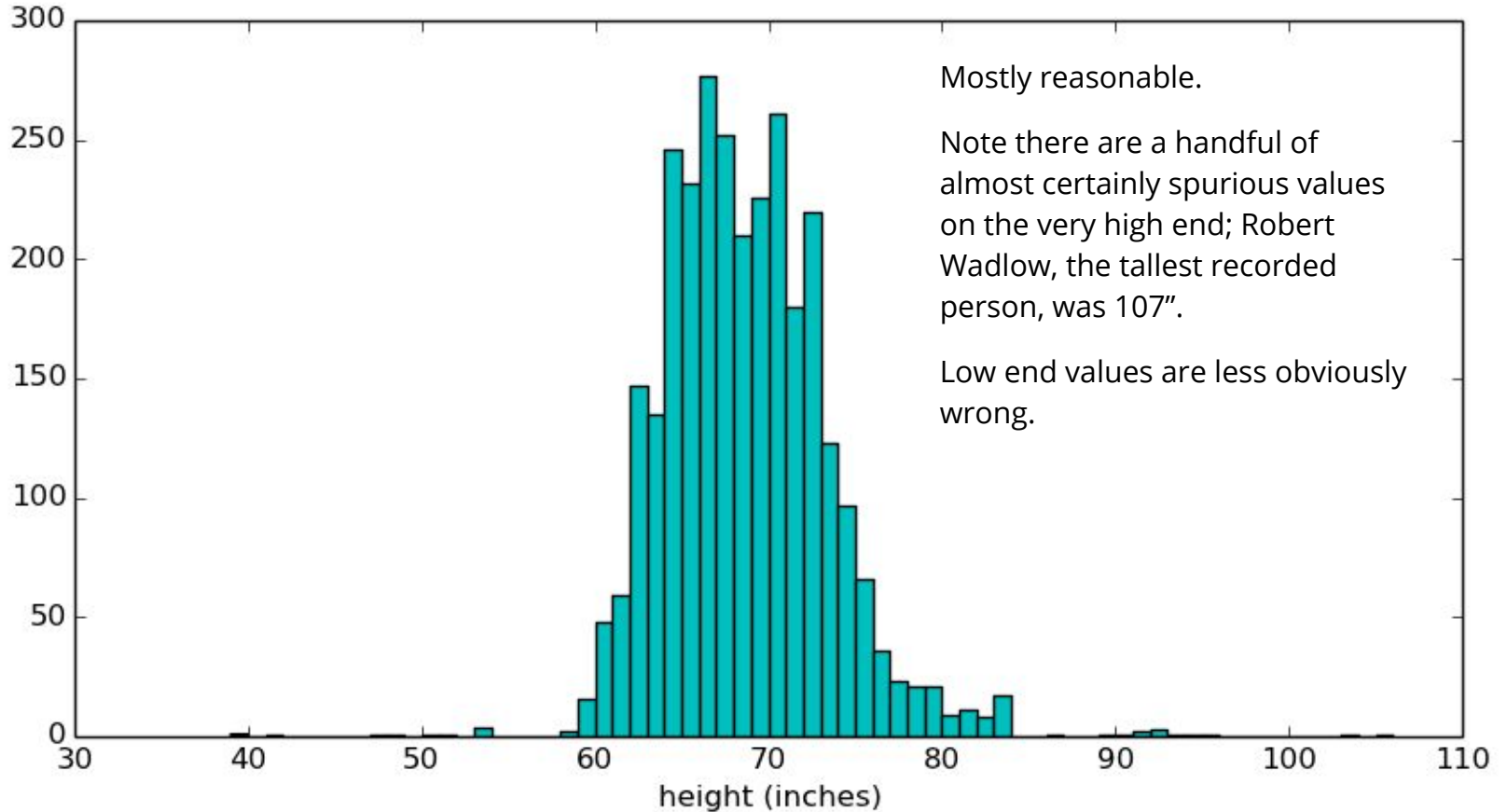
This study made no restrictions on workers and did no reweighting, stratification or other sampling tricks.

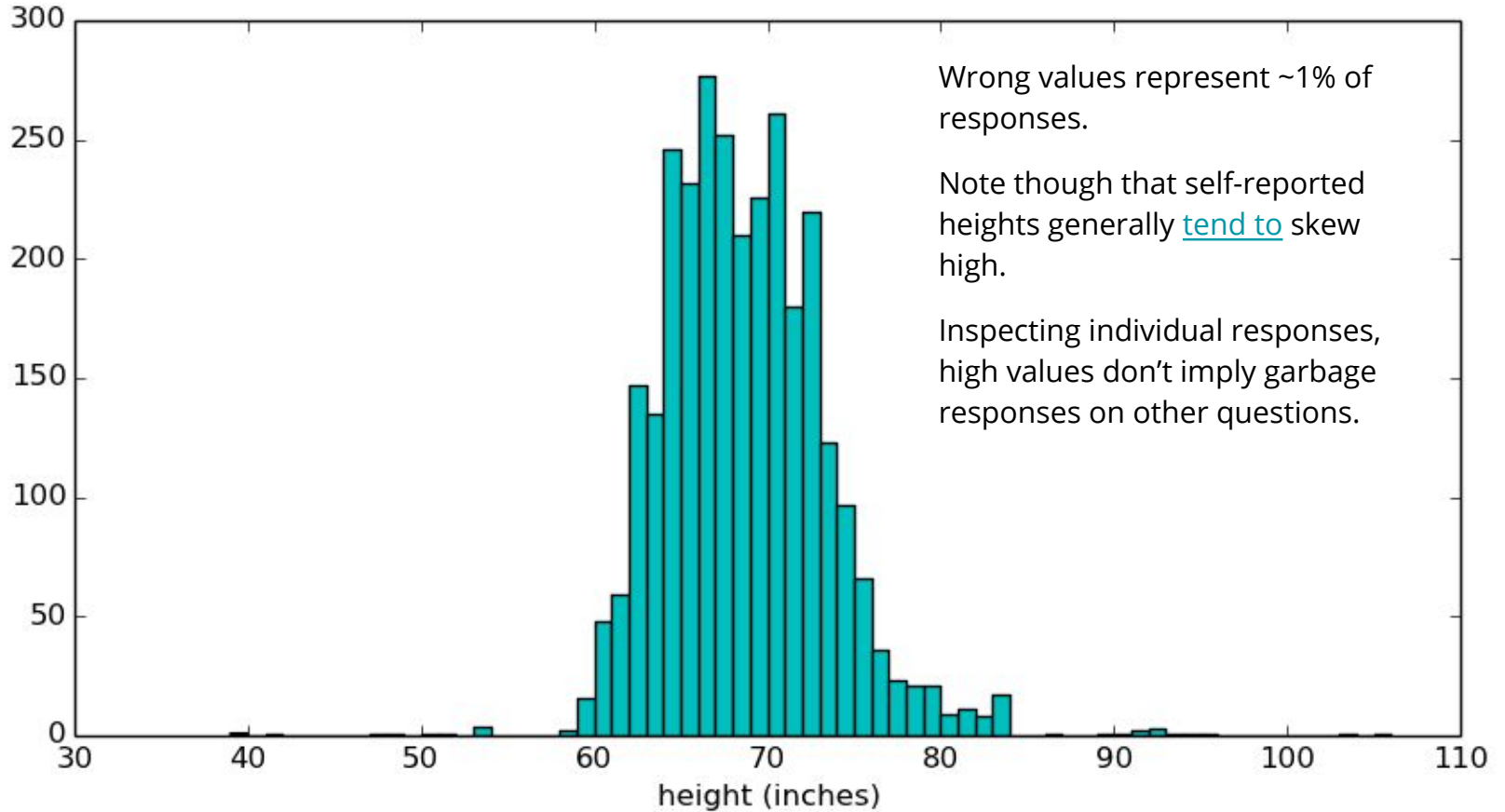
Caveat, but the results are still a lot better than a college campus survey.







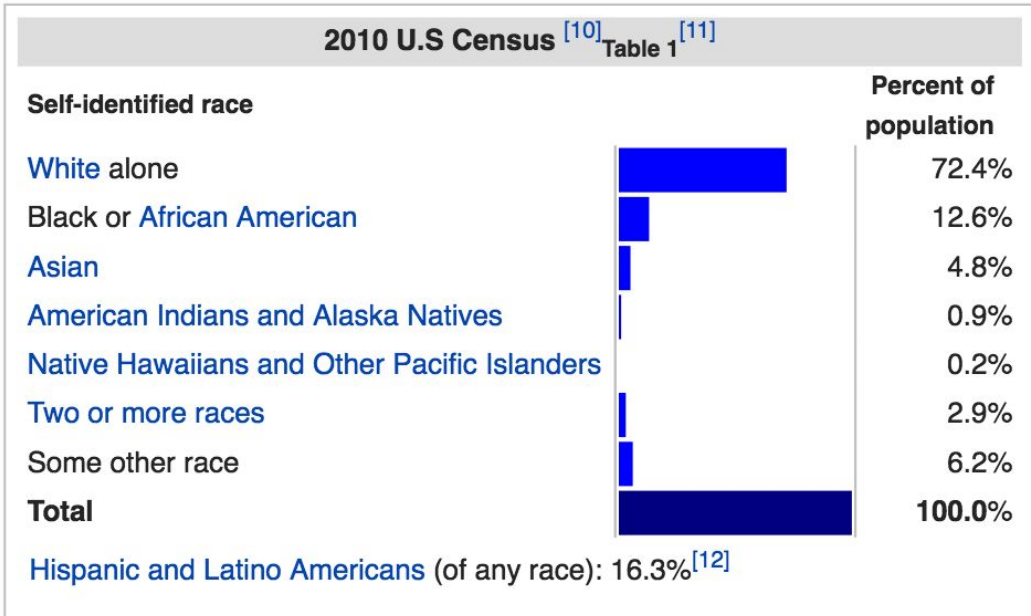




II. How do they identify?

1801 White (60.722%)
 337 Indian (11.362%)
 160 Black (5.394%)
 135 Asian (4.552%)
 132 Latino (4.450%)
 84 AsianIndian (2.832%)
 63 None (2.124%)
 62 LatinoWhite (2.090%)
 62 NativeWhite (2.090%)
 20 AsianWhite (0.674%)
 13 Native (0.438%)
 12 BlackWhite (0.405%)
 10 AsianBlackIndianLatinoNativeWhite (0.337%)
 8 BlackNative (0.270%)
 8 BlackNativeWhite (0.270%)
 7 LatinoNative (0.236%)
 4 AsianIndianWhite (0.135%)
 4 BlackLatino (0.135%)
 4 LatinoNativeWhite (0.135%)
 3 AsianBlackIndian (0.101%)
 3 AsianIndianNative (0.101%)
 3 AsianLatino (0.101%)
 3 IndianNative (0.101%)
 2 AsianBlack (0.067%)
 2 AsianBlackWhite (0.067%)
 2 AsianIndianLatino (0.067%)
 2 AsianIndianNativeWhite (0.067%)
 2 AsianLatinoNativeWhite (0.067%)
 2 AsianLatinoWhite (0.067%)
 2 AsianNative (0.067%)
 2 BlackLatinoNativeWhite (0.067%)
 2 BlackLatinoWhite (0.067%)
 2 IndianWhite (0.067%)
 1 AsianBlackIndianNative (0.034%)
 1 AsianBlackLatinoNative (0.034%)
 1 AsianBlackNativeWhite (0.034%)
 1 BlackIndian (0.034%)
 1 BlackIndianLatino (0.034%)
 1 BlackIndianLatinoNative (0.034%)
 1 BlackLatinoNative (0.034%)
 1 IndianLatinoNativeWhite (0.034%)

Relative to US Census, we have more from India (11%, which is counted in the Census's 4.8% Asian figure); half or fewer Black and Latino; and many more multiracial. In the end though our 61% White is not so far from the Census's 72%, and the major categories are all represented.



1230 identifying as female (41.47%)
1233 identifying only with "she" (41.57%)
1703 identifying as male (57.42%)
1681 identifying only with "he" (56.68%)
 16 identifying as both female and male (0.54%)
 21 identifying with both "she" and "he" (0.71%)
 17 identifying as neither female nor male (0.57%)
 31 identifying with neither "she" nor "he" (1.05%)
 52 identifying with both or neither of "she" or "he" (1.75%)

2558 identifying as hetero (86.24%)
 116 identifying as "homosexual, gay or lesbian" (3.91%)
 369 identifying as bisexual (12.44%)
 462 identifying as LGB (15.58%)
 120 identifying as 2 or more of hetero, homo, bi (4.05%)

Men are a bit overrepresented
(as usual, sigh).

We'll look shortly at how orientation
self-identification compares with
the literature.

1230 identifying as female (41.47%)
1233 identifying only with "she" (41.57%)
1703 identifying as male (57.42%)
1681 identifying only with "he" (56.68%)
 16 identifying as both female and male (0.54%)
 21 identifying with both "she" and "he" (0.71%)
 17 identifying as neither female nor male (0.57%)
 31 identifying with neither "she" nor "he" (1.05%)
 52 identifying with both or neither of "she" or "he" (1.75%)

2558 identifying as hetero (86.24%)
 116 identifying as "homosexual, gay or lesbian" (3.91%)
 369 identifying as bisexual (12.44%)
 462 identifying as LGB (15.58%)
 120 identifying as 2 or more of hetero, homo, bi (4.05%)

Men are a bit overrepresented
(as usual, sigh).

We'll look shortly at how orientation
self-identification compares with
the literature.

Note that nearly every possible
edge case exists in the survey data.

Hi Mindy,* just wanted to give you feedback
that I am asexual, therefore I marked "no" to
the questions about being straight, bisexual,
or the other options.
-Mandy Z

*I used the pseudonym Mindy Ferris to conduct the survey.

1230 identifying as female (41.47%)
1233 identifying only with "she" (41.57%)
1703 identifying as male (57.42%)
1681 identifying only with "he" (56.68%)
16 identifying as both female and male (0.54%)
21 identifying with both "she" and "he" (0.71%)
17 identifying as neither female nor male (0.57%)
31 identifying with neither "she" nor "he" (1.05%)
52 identifying with both or neither of "she" or "he" (1.75%)

2558 identifying as hetero (86.24%)
116 identifying as "homosexual, gay or lesbian" (3.91%)
369 identifying as bisexual (12.44%)
462 identifying as LGB (15.58%)
120 identifying as 2 or more of hetero, homo, bi (4.05%)

Naively these may seem like they ought to be perfectly correlated.

1230 identifying as female (41.47%)
1233 identifying only with "she" (41.57%)
1703 identifying as male (57.42%)
1681 identifying only with "he" (56.68%)

They are strongly correlated, but there are false negatives and false positives in the 1-3% range.

	Pearson's ϕ	False-positive	False-negative	Error-at-chance
identifies-female predicts "she"	0.972	1.28%	1.91%	57.72%
identifies-male predicts "he"	0.965	1.98%	1.00%	42.62%
identifies-only-female predicts "she"	0.970	0.81%	2.71%	57.72%
identifies-only-male predicts "he"	0.966	1.47%	1.41%	42.62%

1230 identifying as female (41.47%)
 1233 identifying only with "she" (41.57%)
 1703 identifying as male (57.42%)
 1681 identifying only with "he" (56.68%)

They are strongly correlated, but there are false negatives and false positives in the 1-3% range.

	Pearson's ϕ	False-positive	False-negative	Error-at-chance
identifies-female predicts "she"	0.972	1.28%	1.91%	57.72%
identifies-male predicts "he"	0.965	1.98%	1.00%	42.62%
identifies-only-female predicts "she"	0.970	0.81%	2.71%	57.72%
identifies-only-male predicts "he"	0.966	1.47%	1.41%	42.62%

[Pearson's \$\phi\$ coefficient](#) is a measure of correlation significance (analogous to [Pearson's r](#) but for categorical variables) ranging from -1 (perfectly anticorrelated) to +1 (perfectly correlated).

Chance is the likelihood of getting the predicted variable right by chance alone.

- 1230 identifying as female (41.47%)
- 1233 identifying only with "she" (41.57%)
- 1703 identifying as male (57.42%)
- 1681 identifying only with "he" (56.68%)
 - 16 identifying as both female and male (0.54%)
 - 21 identifying with both "she" and "he" (0.71%)
 - 17 identifying as neither female nor male (0.57%)
 - 31 identifying with neither "she" nor "he" (1.05%)
 - 52 identifying with both or neither of "she" or "he" (1.75%)

It is significant when people answer "yes" to both she and he, or to neither. These are correlated with gender-nxor, which we define here as answering yes to both or neither of identifying male or female.

	Pearson's ϕ	False-positive	False-negative	Error-at-chance
identifies-female predicts "she"	0.972	1.28%	1.91%	57.72%
identifies-male predicts "he"	0.965	1.98%	1.00%	42.62%
identifies-only-female predicts "she"	0.970	0.81%	2.71%	57.72%
identifies-only-male predicts "he"	0.966	1.47%	1.41%	42.62%
uses-"she"-and-"he" predicts gender-nxor	0.106	85.71%	90.91%	98.89%
uses-neither-"she"-nor-"he" predicts gender-nxor	0.179	80.65%	81.82%	98.89%
uses-both/neither-"she"/"he" predicts gender-nxor	0.206	82.69%	72.73%	98.89%

- 1230 identifying as female (41.47%)
- 1233 identifying only with "she" (41.57%)
- 1703 identifying as male (57.42%)
- 1681 identifying only with "he" (56.68%)
 - 16 identifying as both female and male (0.54%)
 - 21 identifying with both "she" and "he" (0.71%)
 - 17 identifying as neither female nor male (0.57%)
 - 31 identifying with neither "she" nor "he" (1.05%)
 - 52 identifying with both or neither of "she" or "he" (1.75%)

It is significant when people answer "yes" to both she and he, or to neither. These are correlated with gender-nxor, which we define here as answering yes to both or neither of identifying male or female.

	Pearson's ϕ	False-positive	False-negative	Error-at-chance
identifies-female predicts "she"	0.972	1.28%	1.91%	57.72%
identifies-male predicts "he"	0.965	1.98%	1.00%	42.62%
identifies-only-female predicts "she"	0.970	0.81%	2.71%	57.72%
identifies-only-male predicts "he"	0.966	1.47%	1.41%	42.62%
uses-"she"-and-"he" predicts gender-nxor	0.106	85.71%	90.91%	98.89%
uses-neither-"she"-nor-"he" predicts gender-nxor	0.179	80.65%	81.82%	98.89%
uses-both/neither-"she"/"he" predicts gender-nxor	0.206	82.69%	72.73%	98.89%

(gender-nxor is related to gender-fluid or agender, both identities we should in retrospect have asked about explicitly also).

III. Gender as an emergent binary, spectrum, and vector space

Questions

Body

Presentation

Attraction

Identity

Age

Height

Do you menstruate?

Have you ever menstruated?

Do you have a penis?

Do you have a vagina?

Have you ever been pregnant?

2 numeric,
5 yes/no

Questions

Body

Presentation

Attraction

Identity

2 numeric,
5 yes/no

Age

Height

Do you menstruate?

Have you ever menstruated?

Do you have a penis?

Do you have a vagina?

Have you ever been pregnant?

Let's do the simplest kind of learning on this data: linear analysis.

Questions

Body

Presentation

Attraction

Identity

Age

Height

Do you menstruate?

Have you ever menstruated?

Do you have a penis?

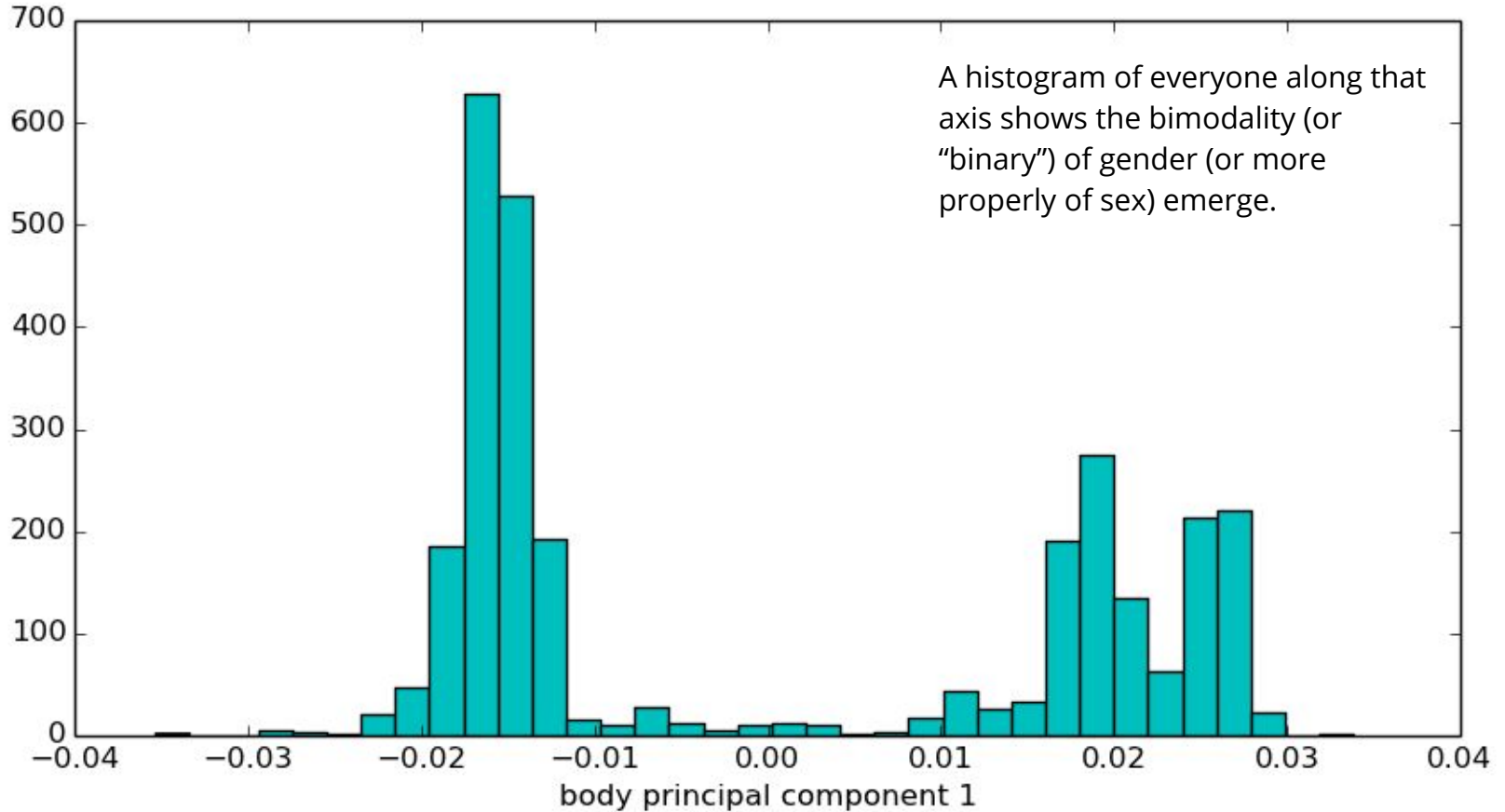
Do you have a vagina?

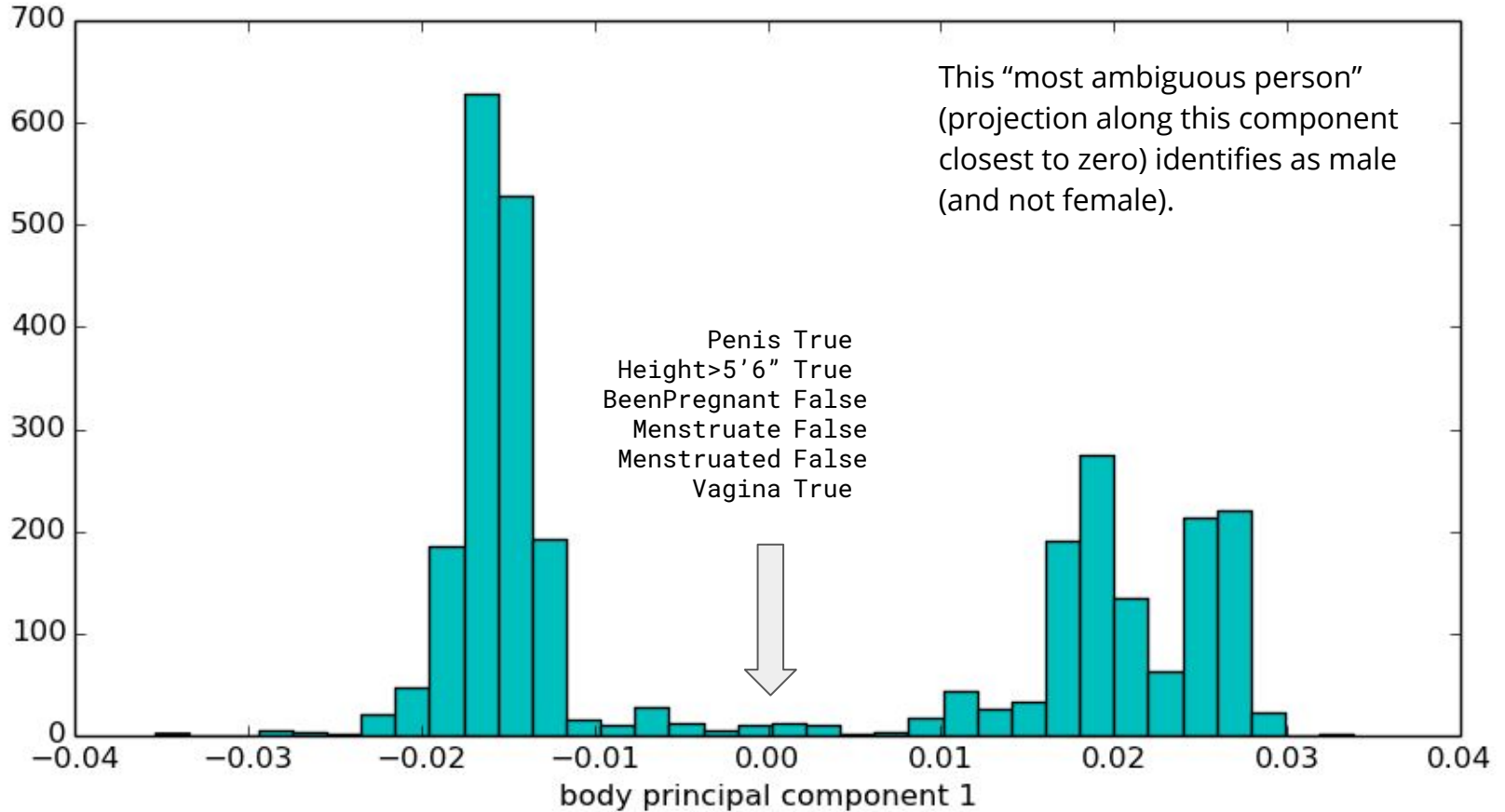
Have you ever been pregnant?

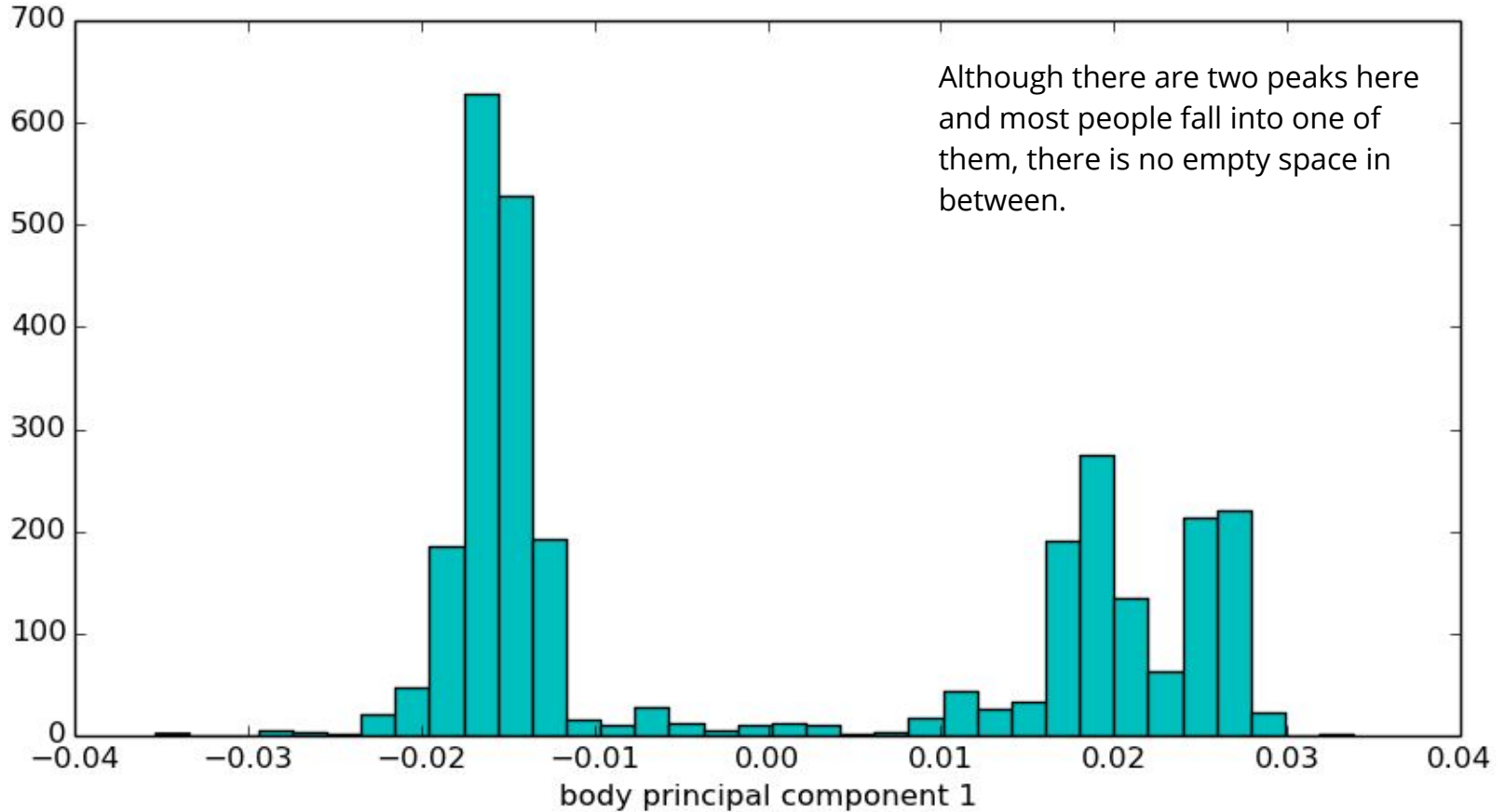
2 numeric,
5 yes/no

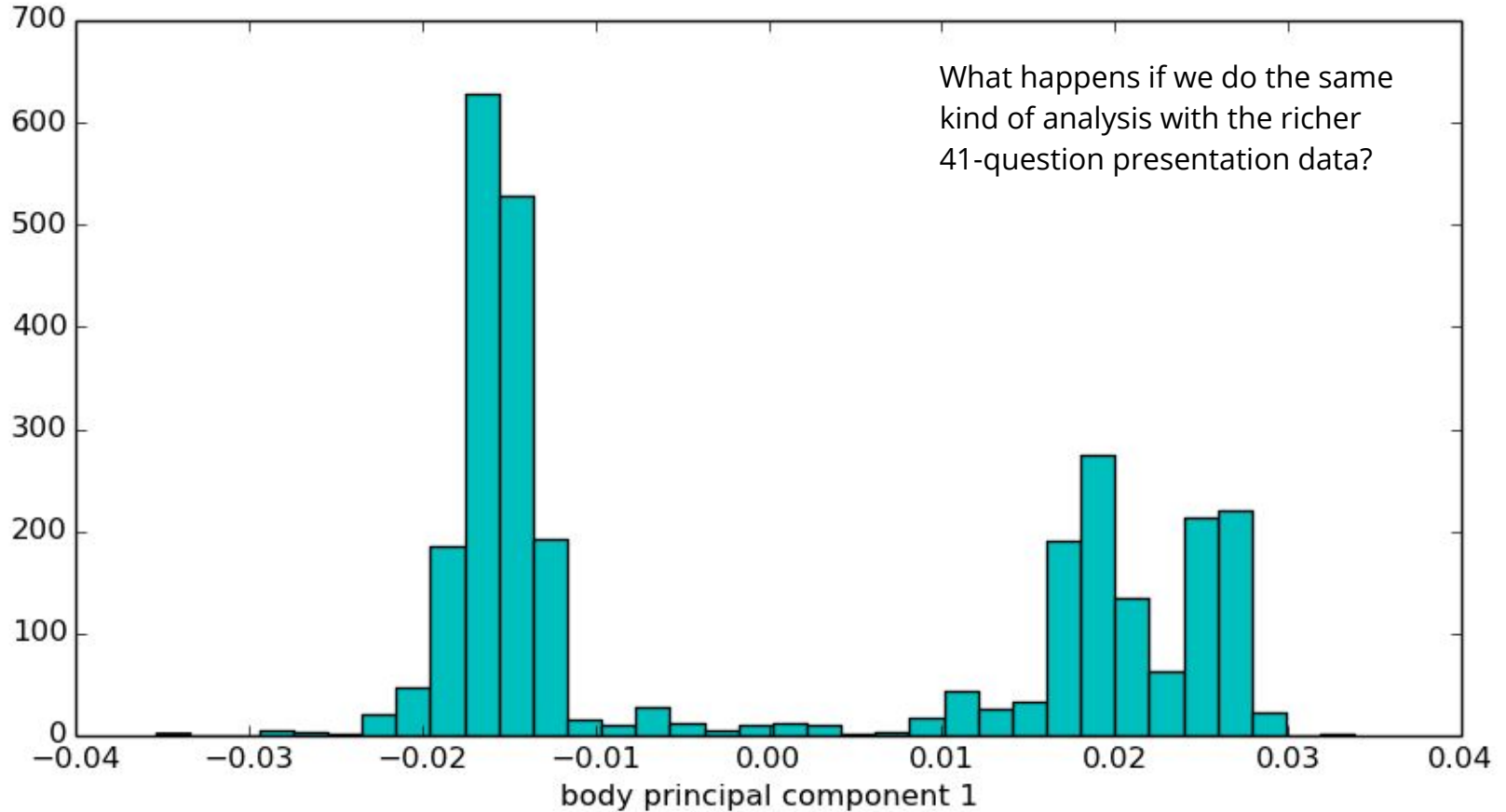
Taking the [Singular Value Decomposition](#) (SVD) we get a leading mode that looks like it pulls out gender as the dominant dimension:

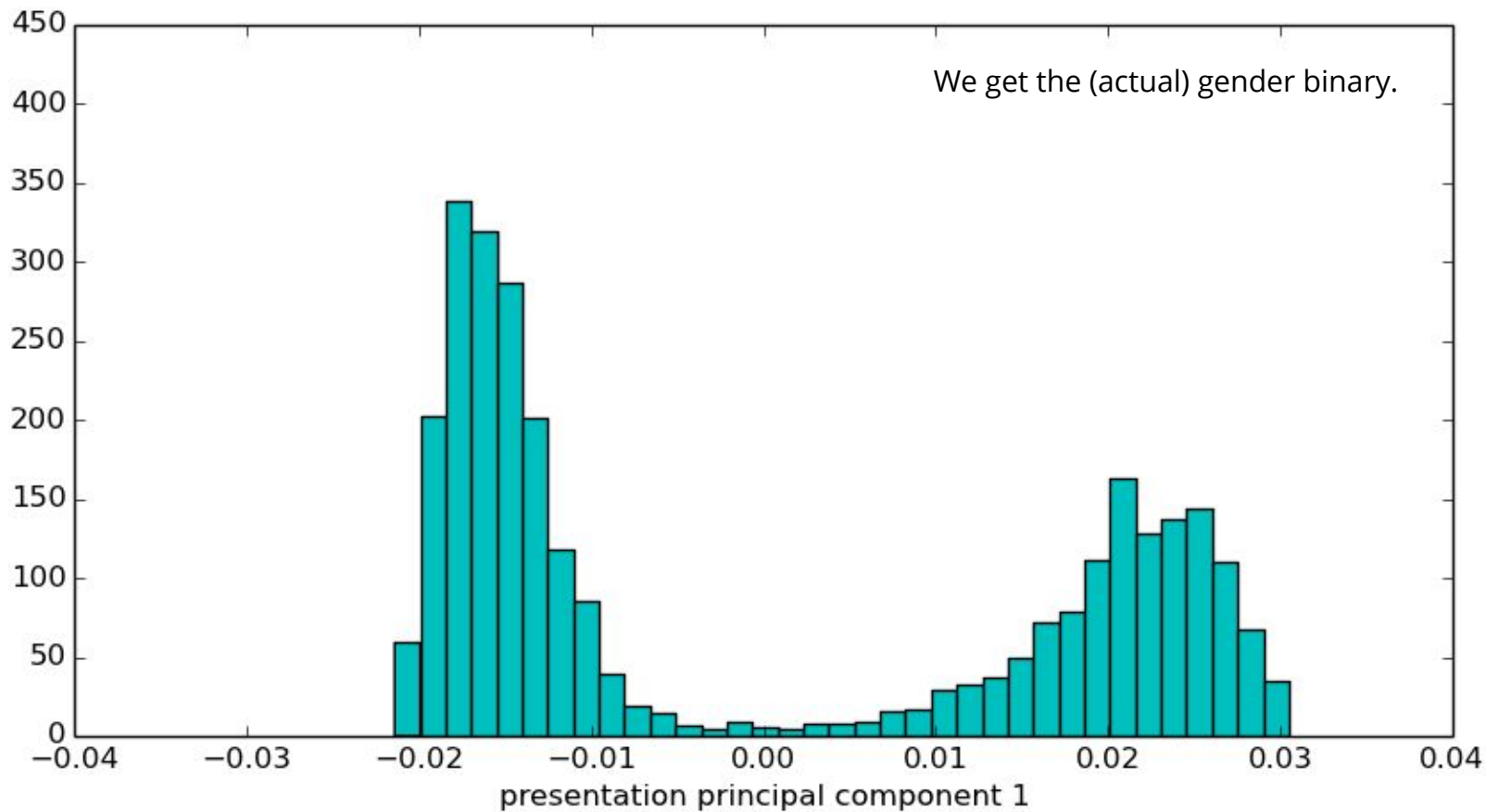
```
-0.463768 Penis  
-0.327008 HeightGt5ft6  
0.277065 BeenPregnant  
0.414318 Menstruate  
0.460653 Menstruated  
0.466218 Vagina
```

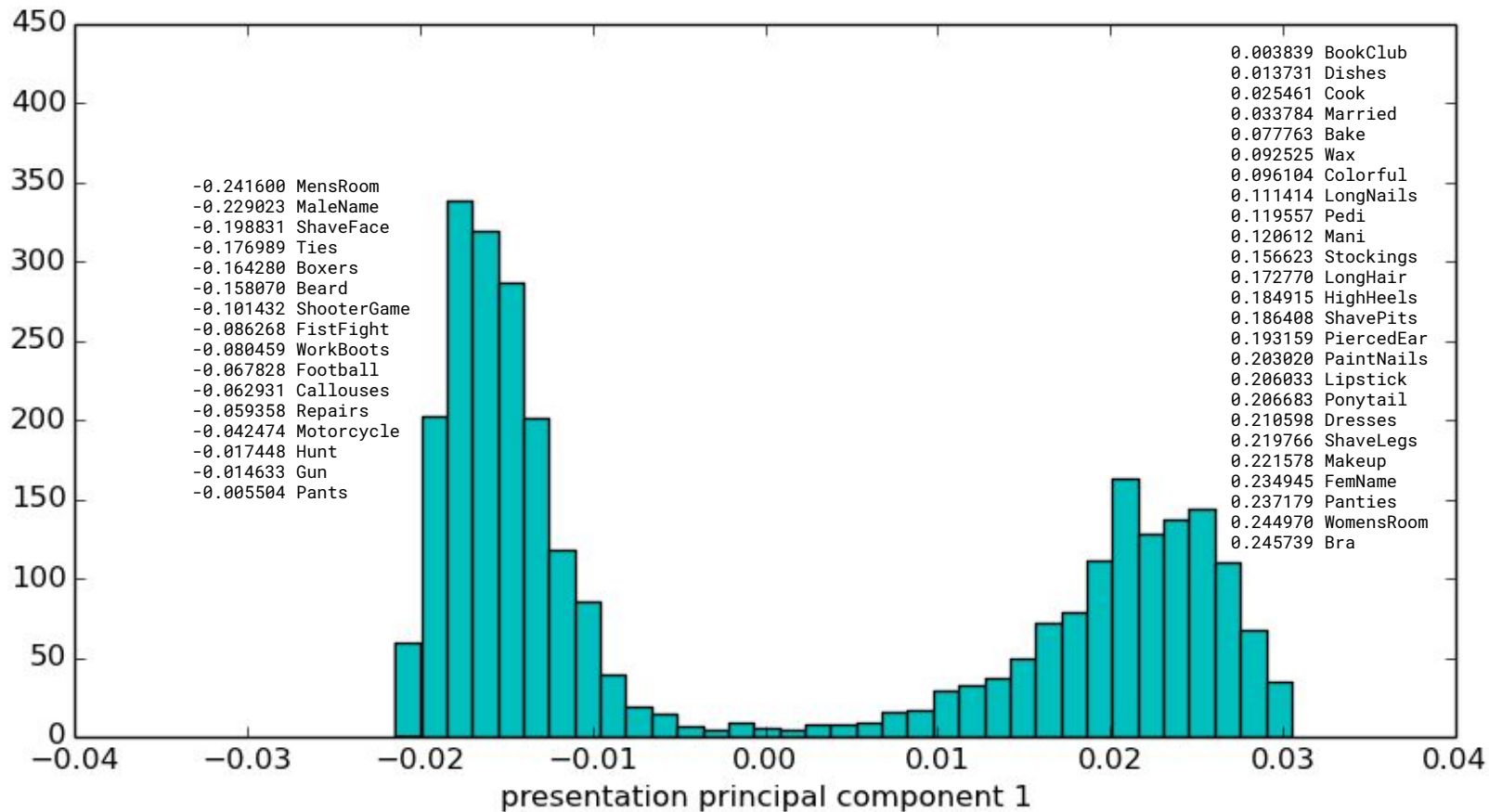


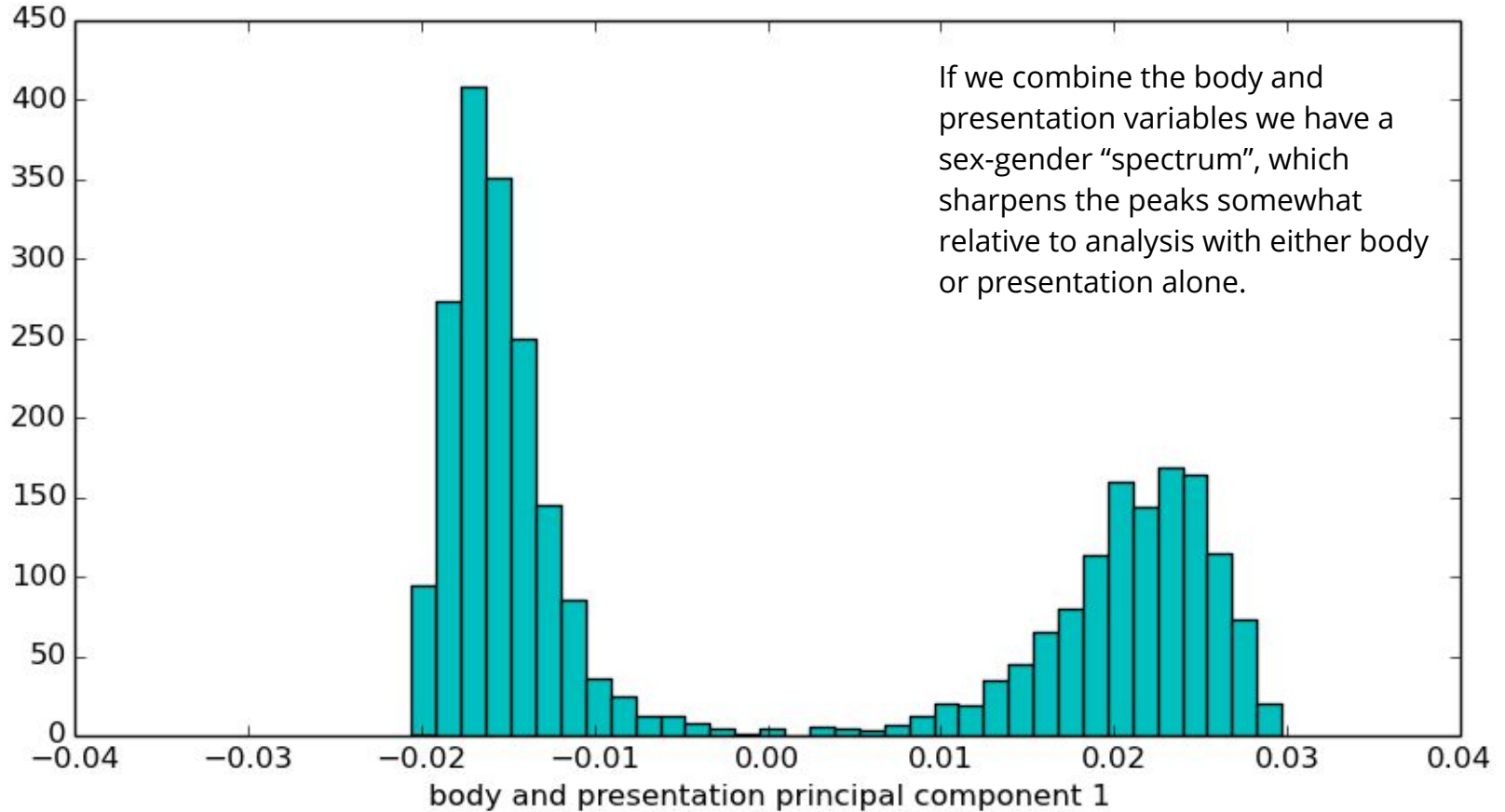




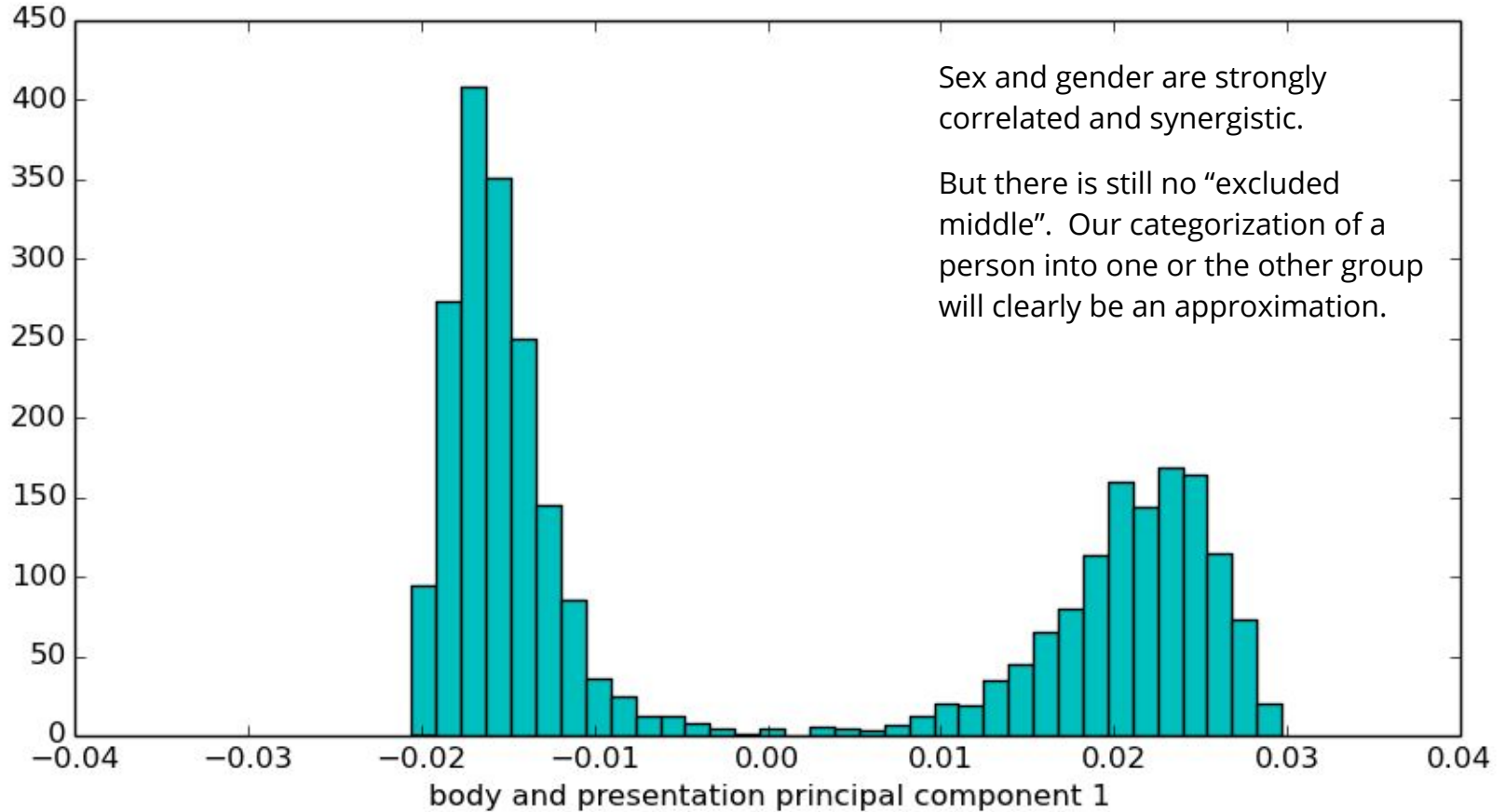






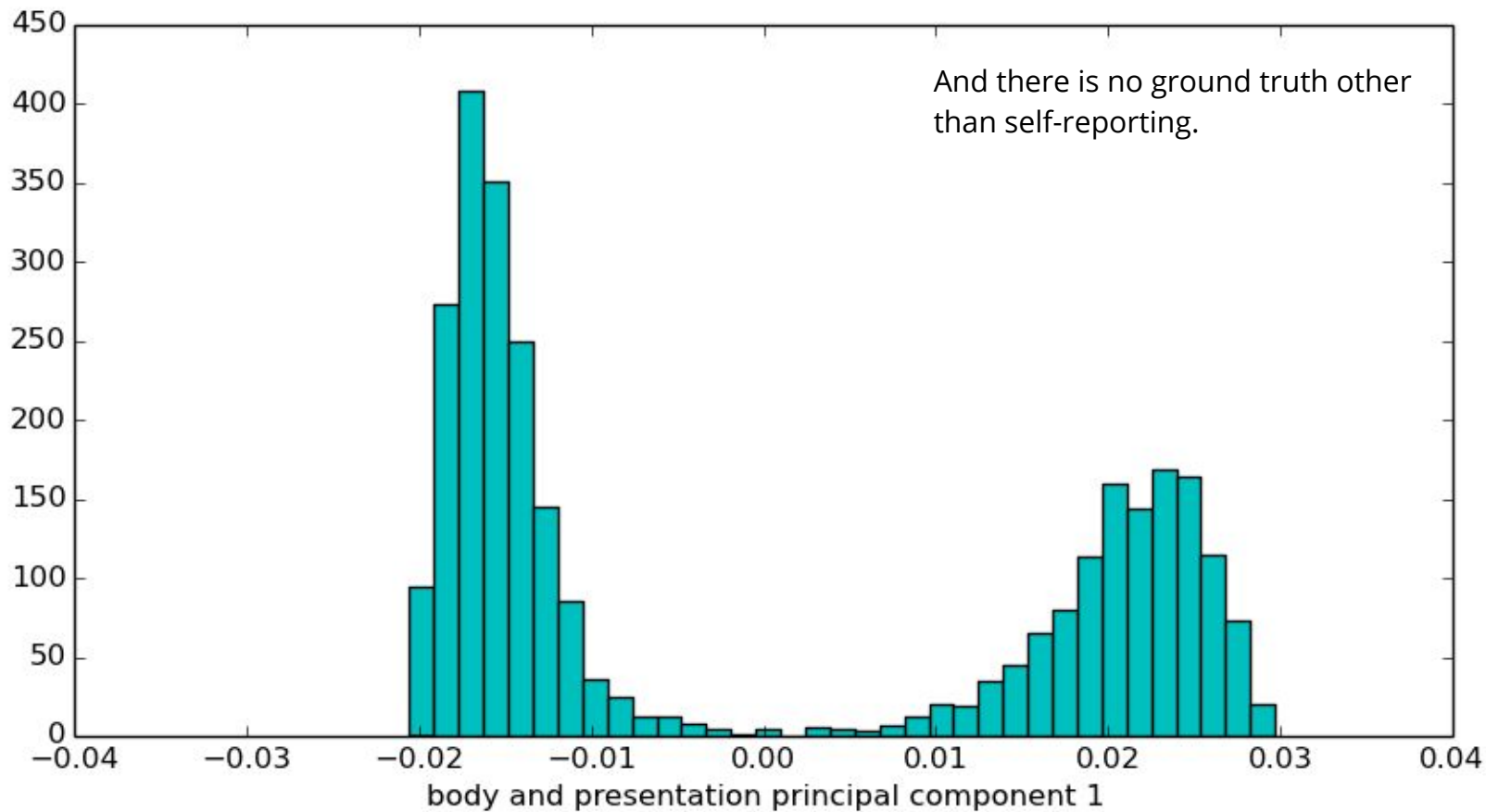


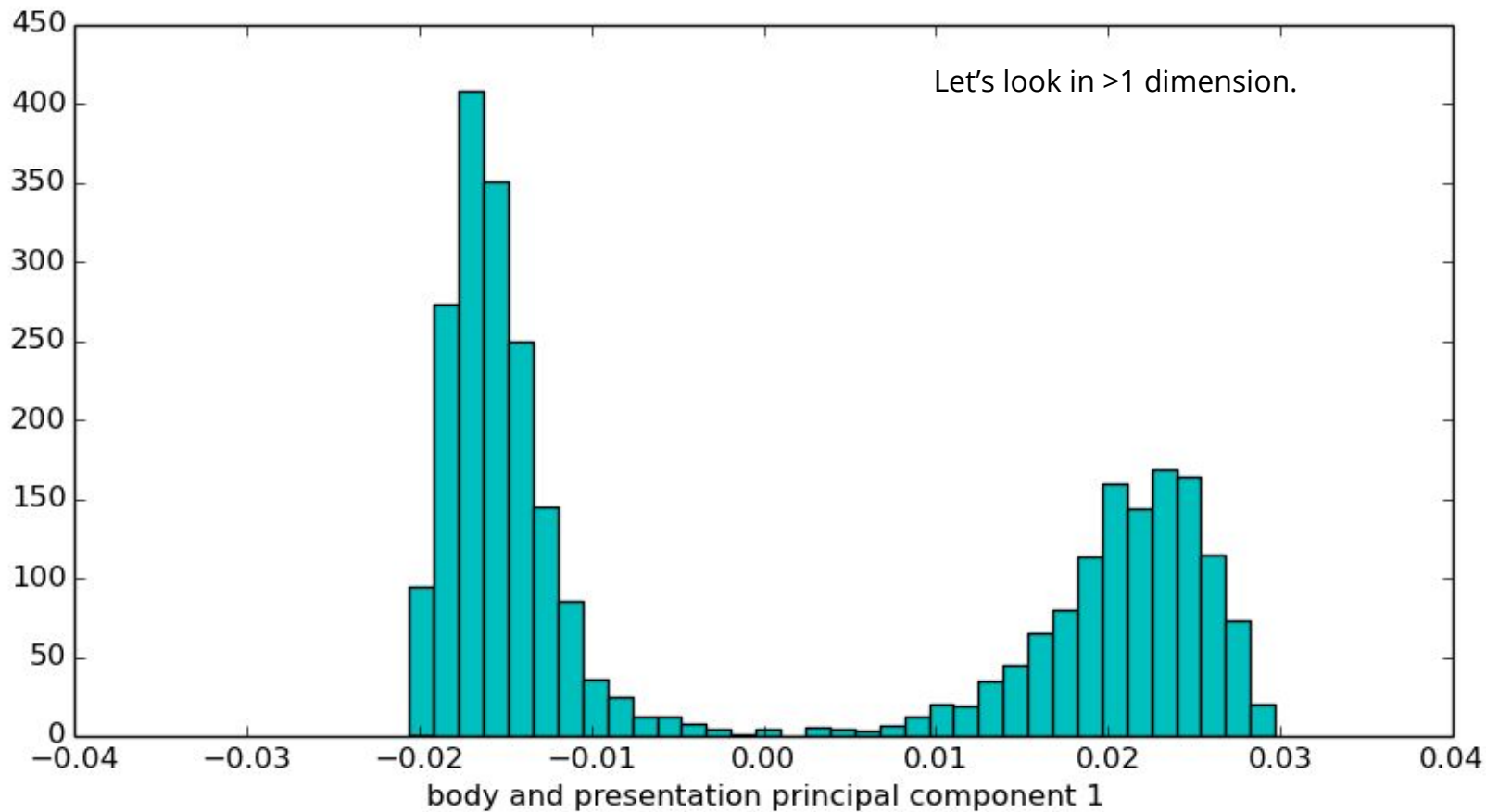
If we combine the body and presentation variables we have a sex-gender "spectrum", which sharpens the peaks somewhat relative to analysis with either body or presentation alone.



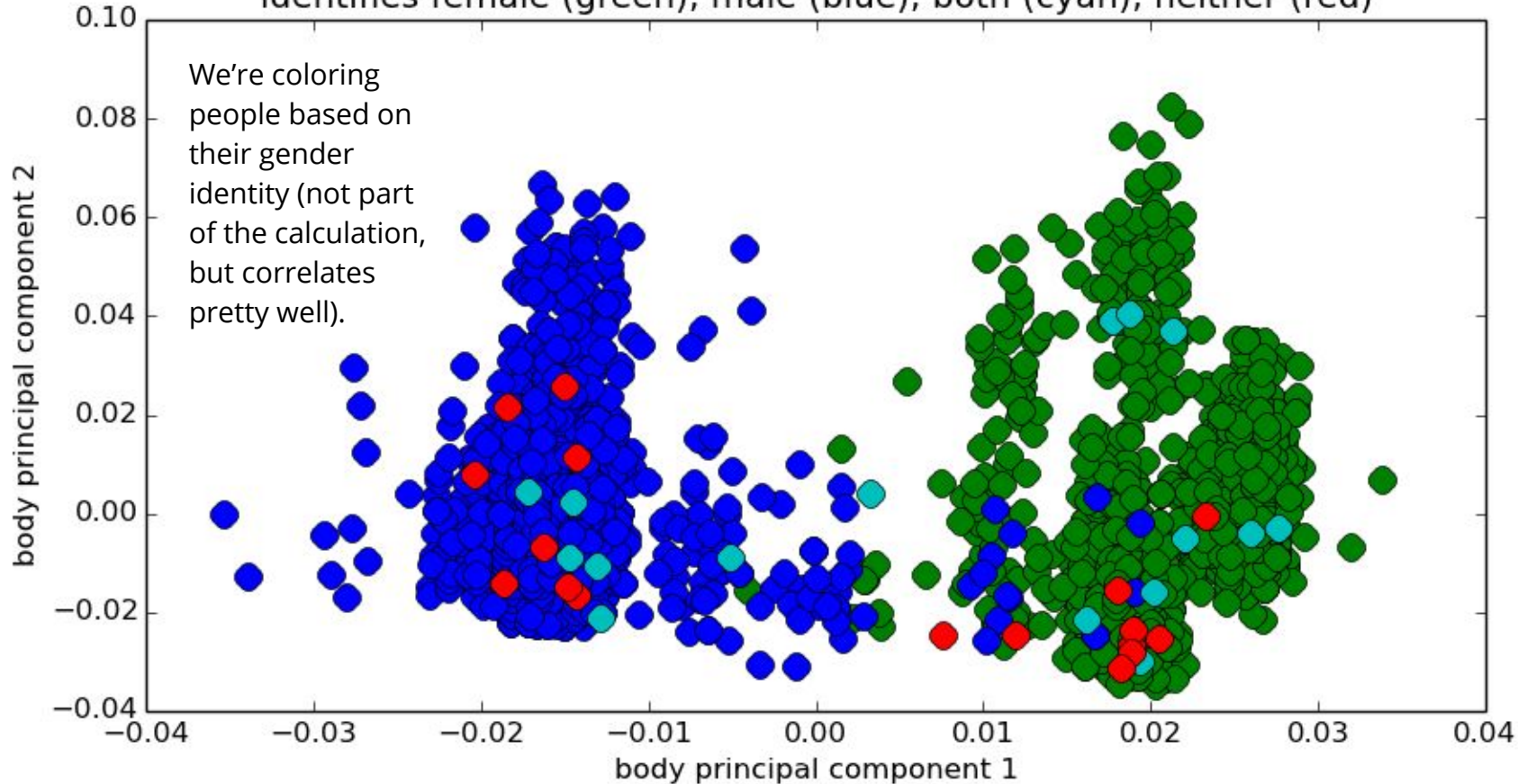
Sex and gender are strongly correlated and synergistic.

But there is still no “excluded middle”. Our categorization of a person into one or the other group will clearly be an approximation.

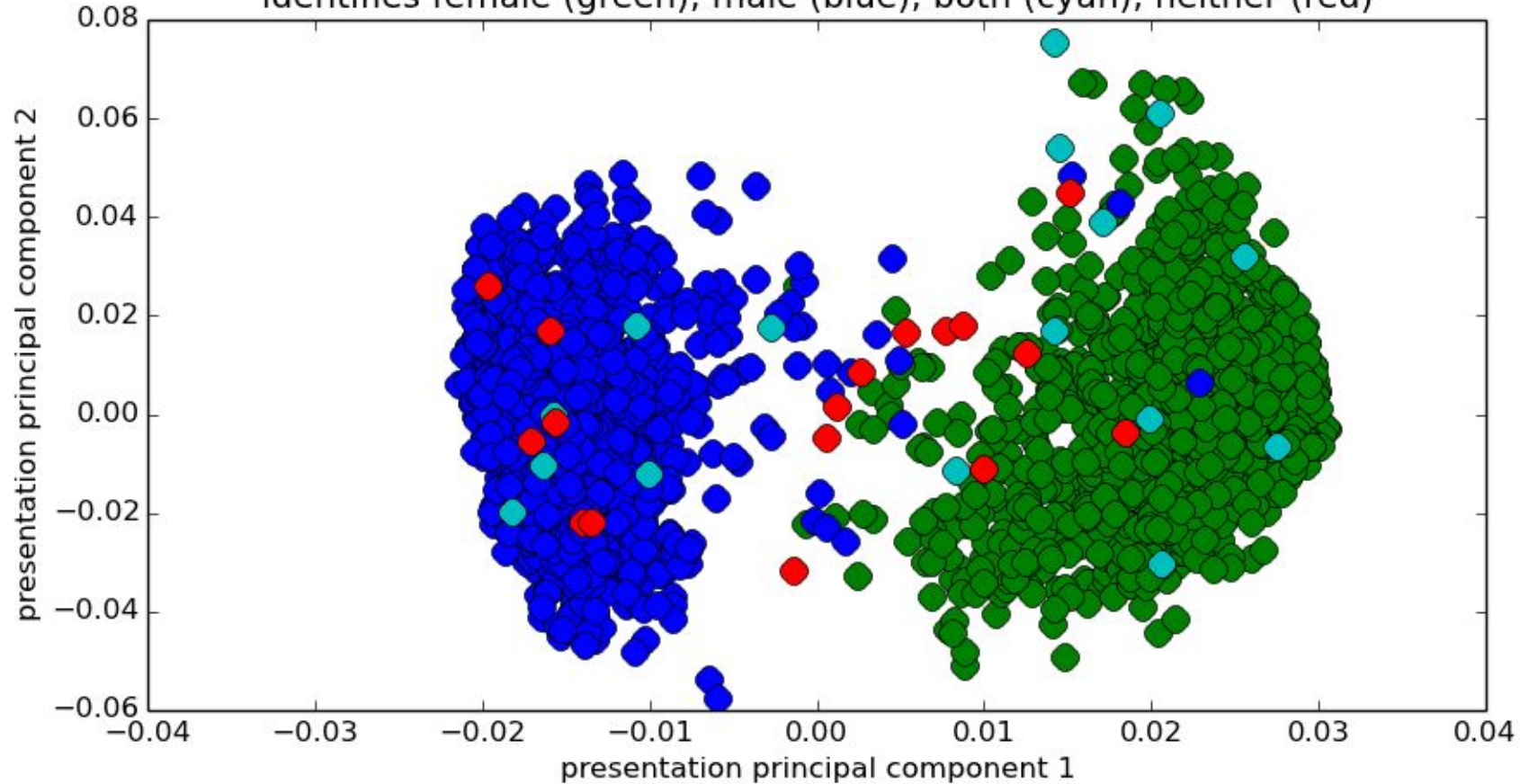


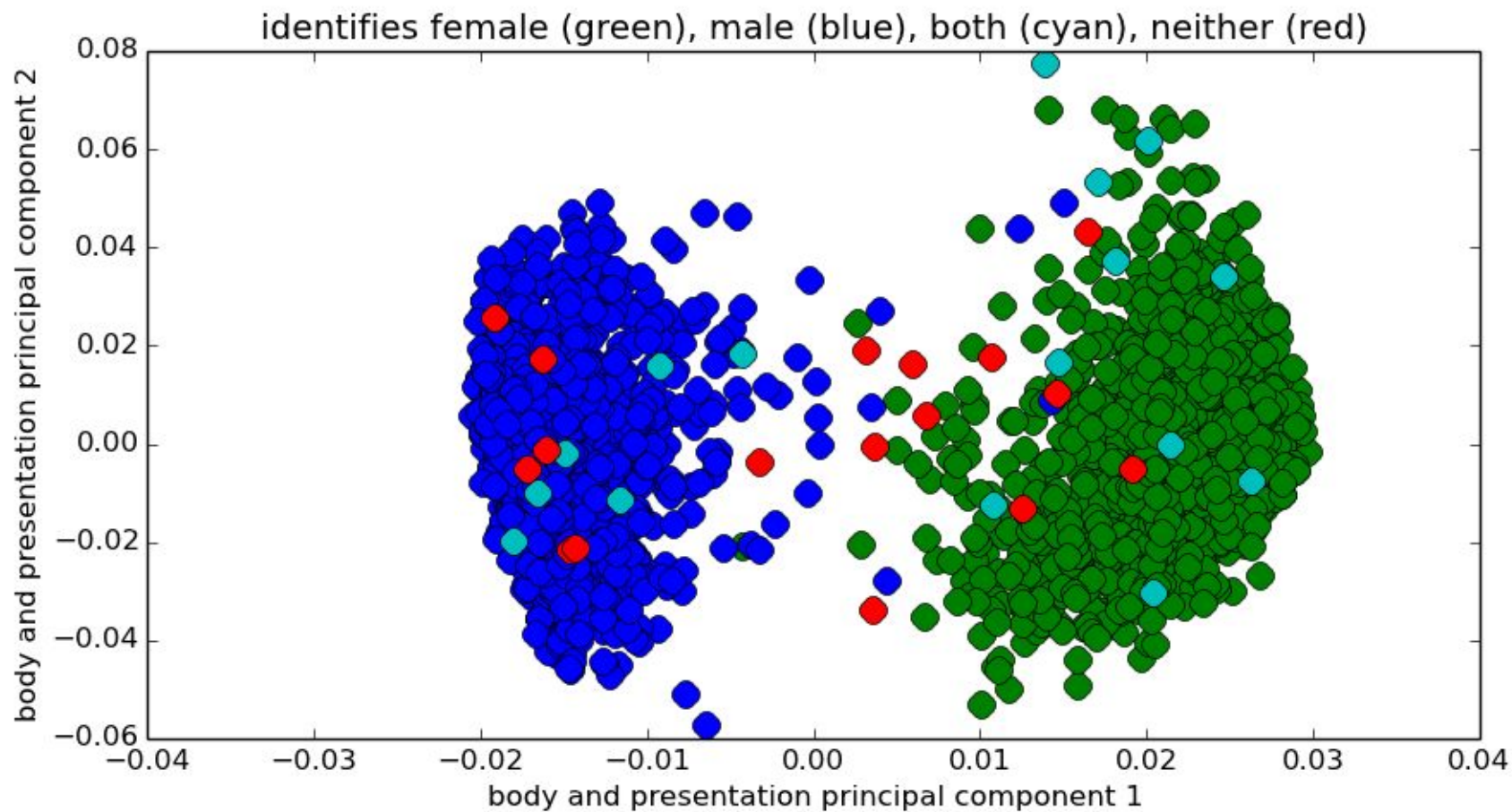


identifies female (green), male (blue), both (cyan), neither (red)

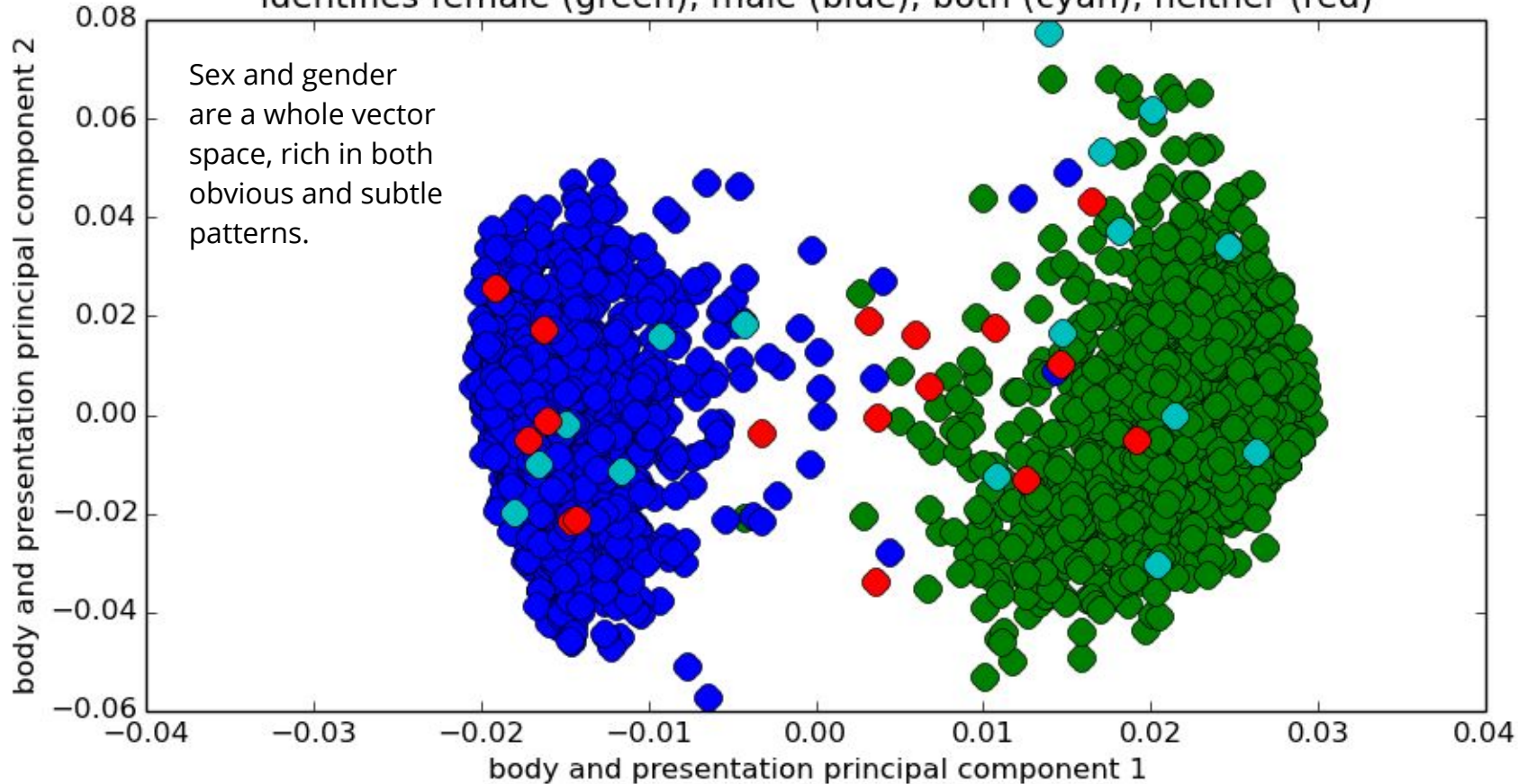


identifies female (green), male (blue), both (cyan), neither (red)

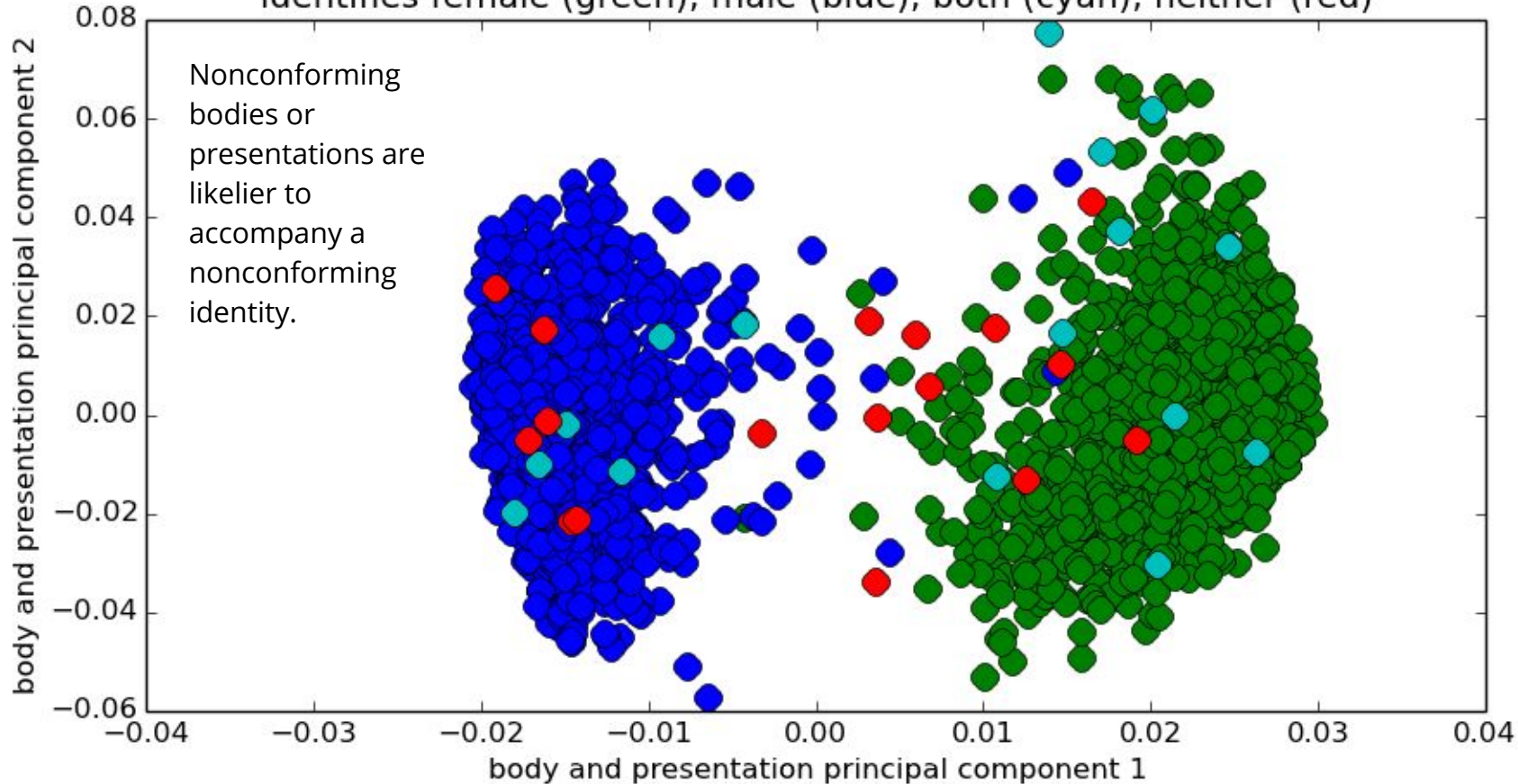




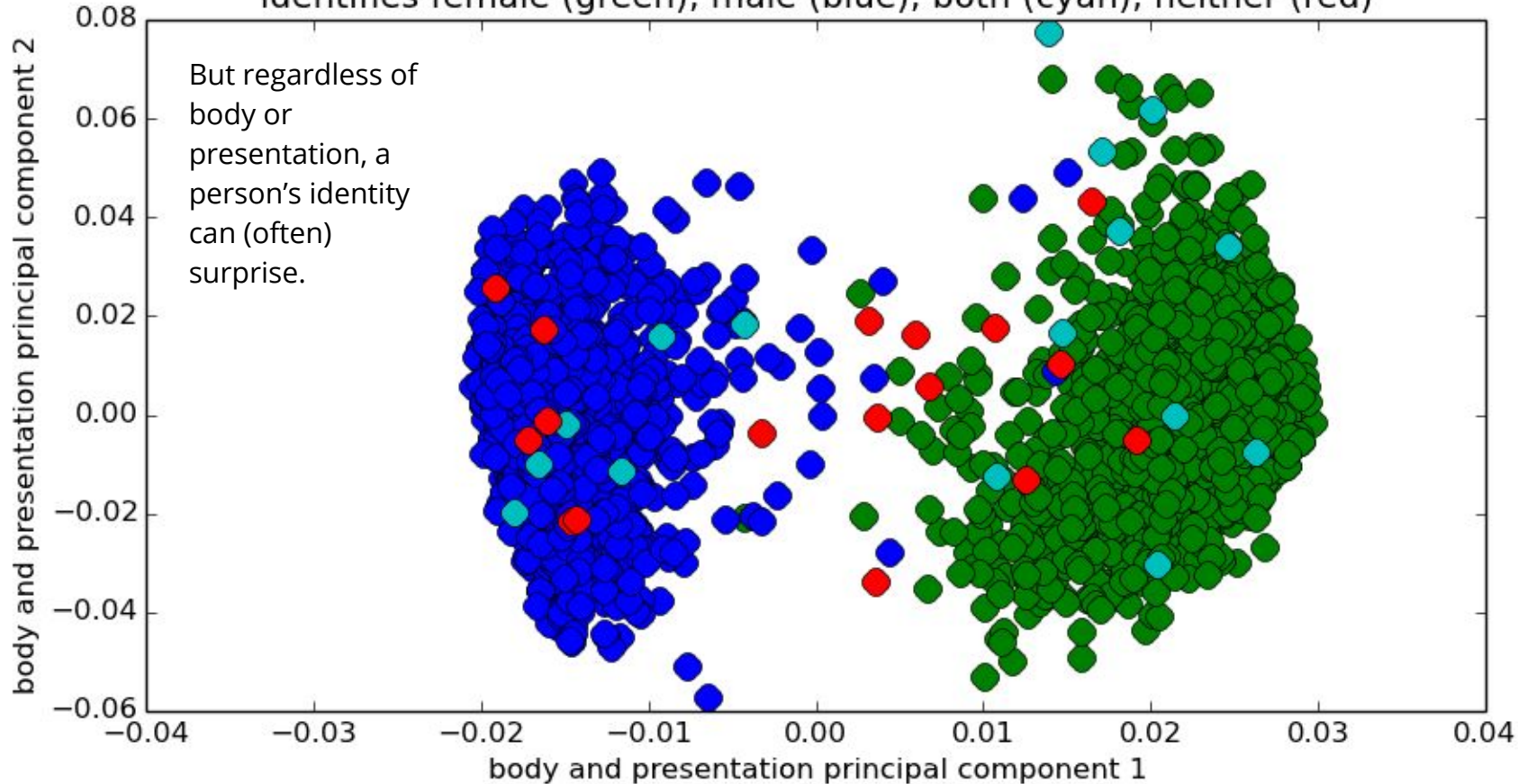
identifies female (green), male (blue), both (cyan), neither (red)



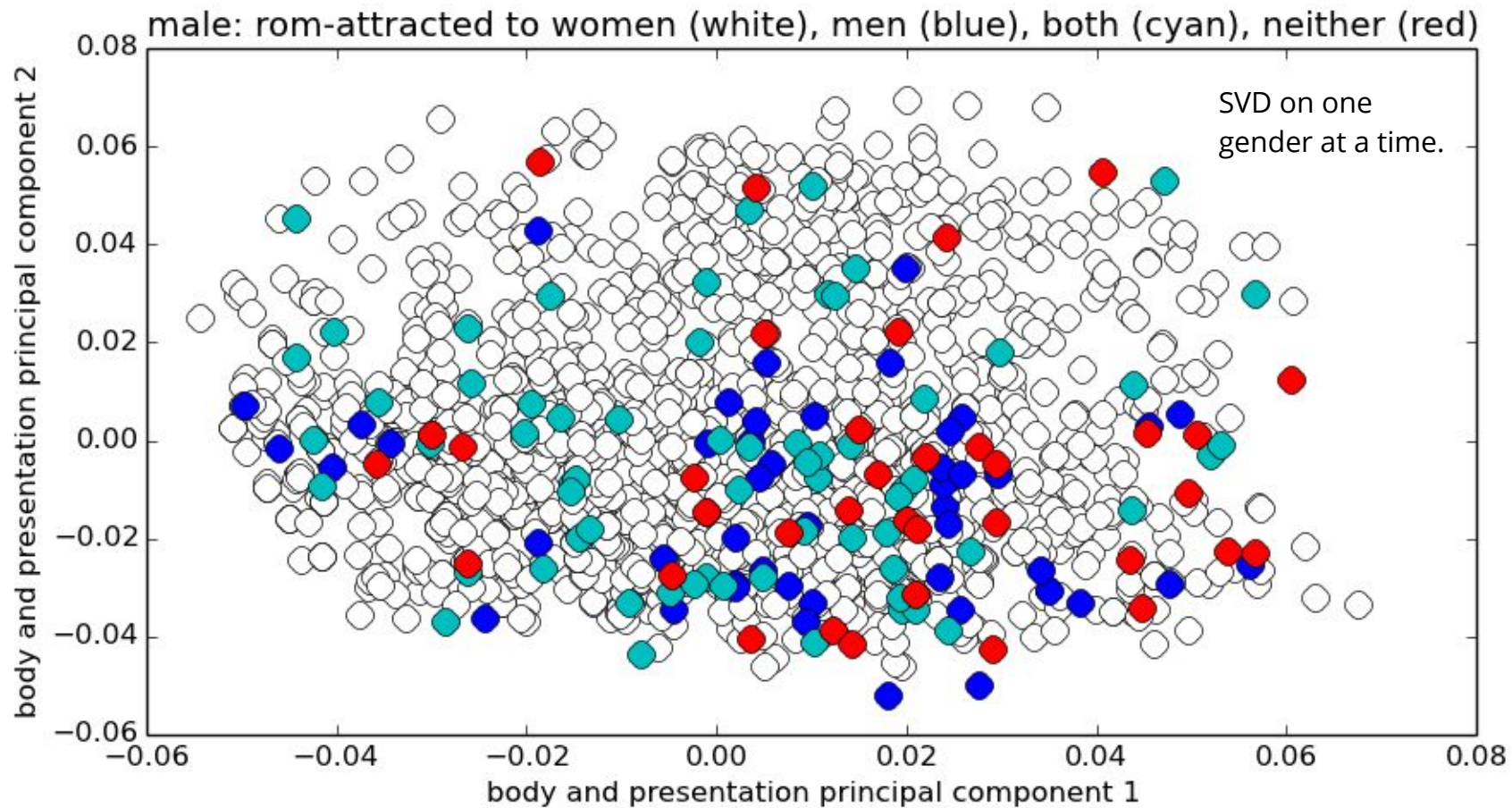
identifies female (green), male (blue), both (cyan), neither (red)

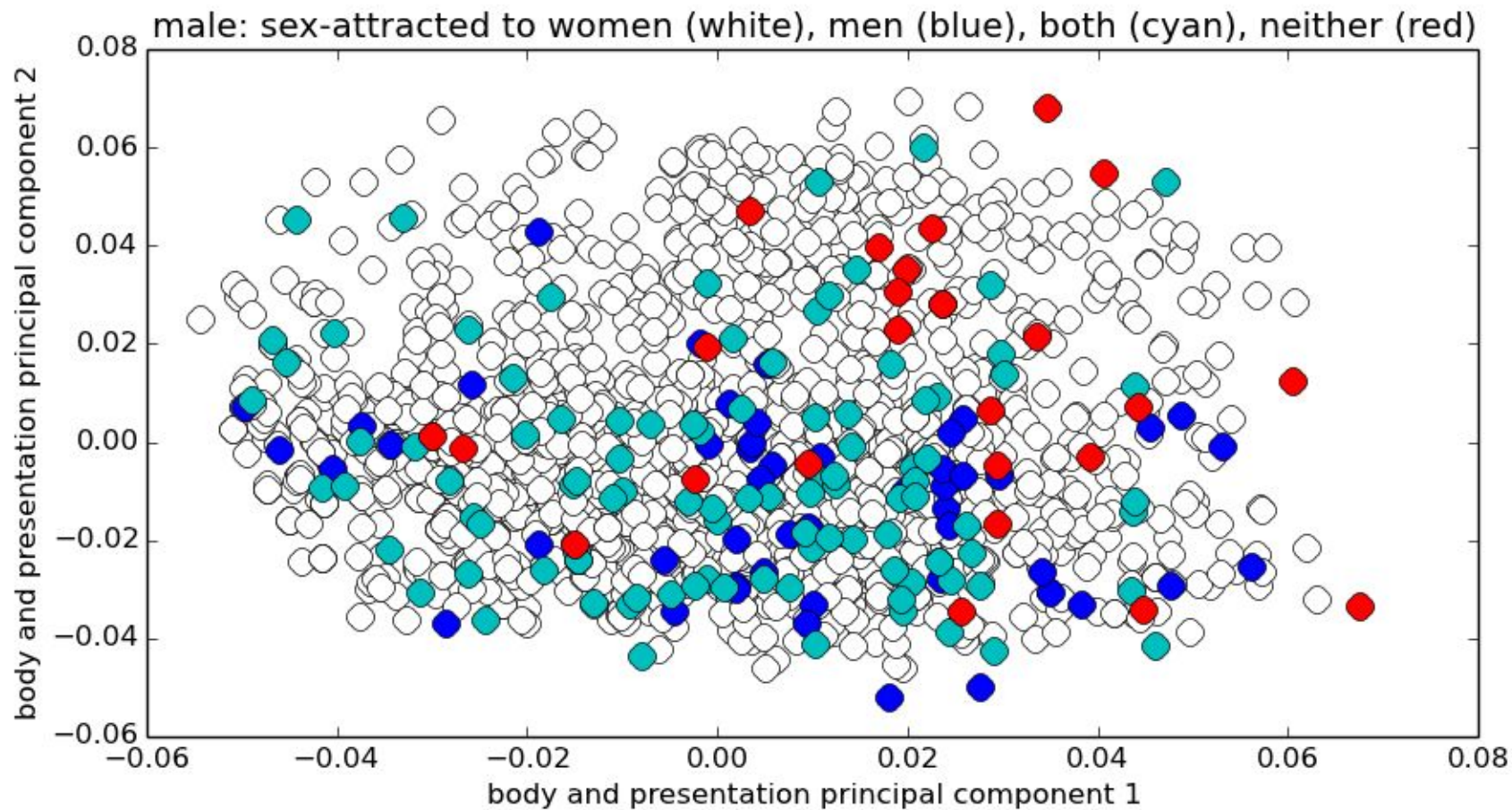


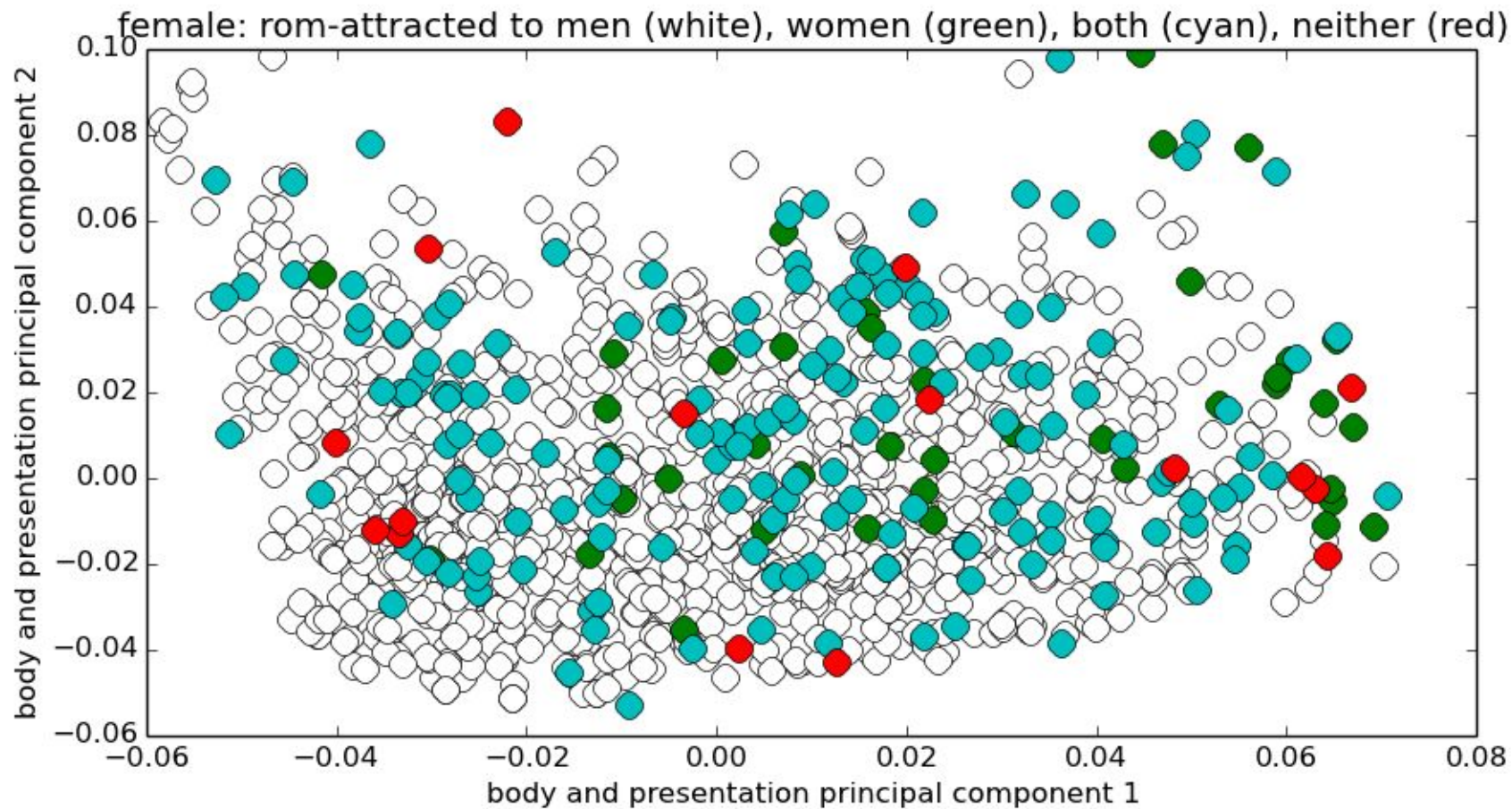
identifies female (green), male (blue), both (cyan), neither (red)

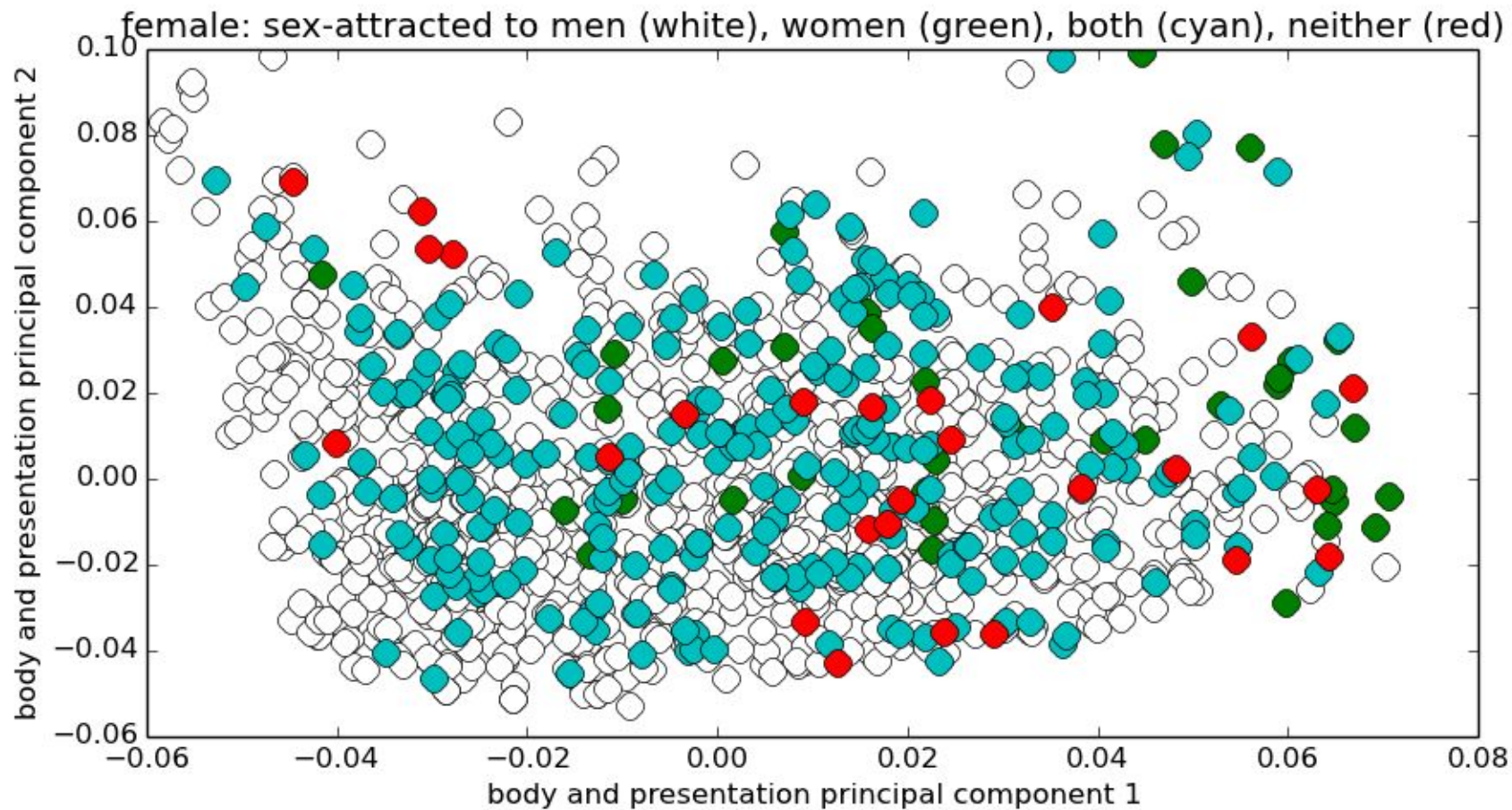


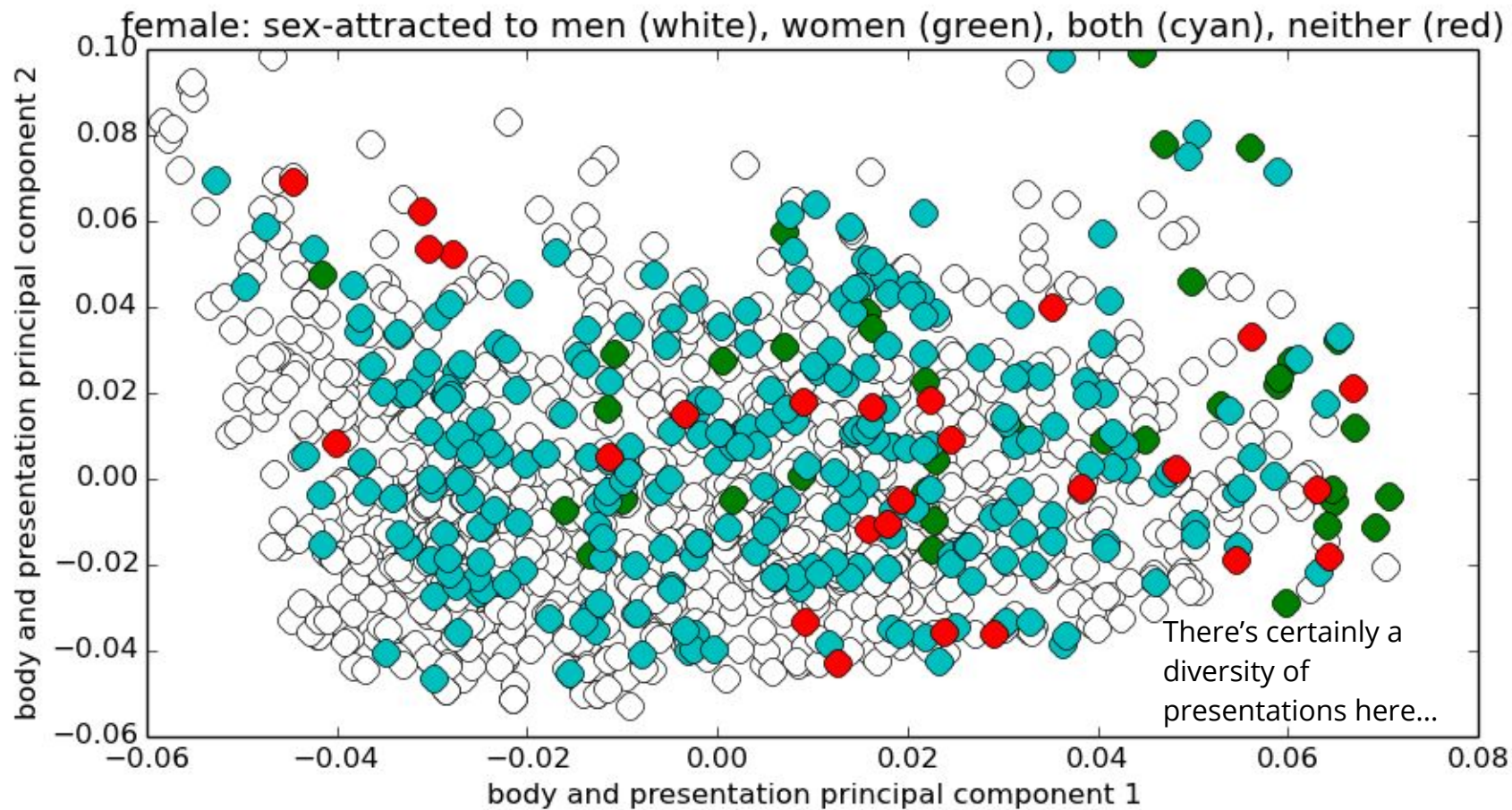
IV. Gaydar (and other kinds of *-radar) exist but are iff

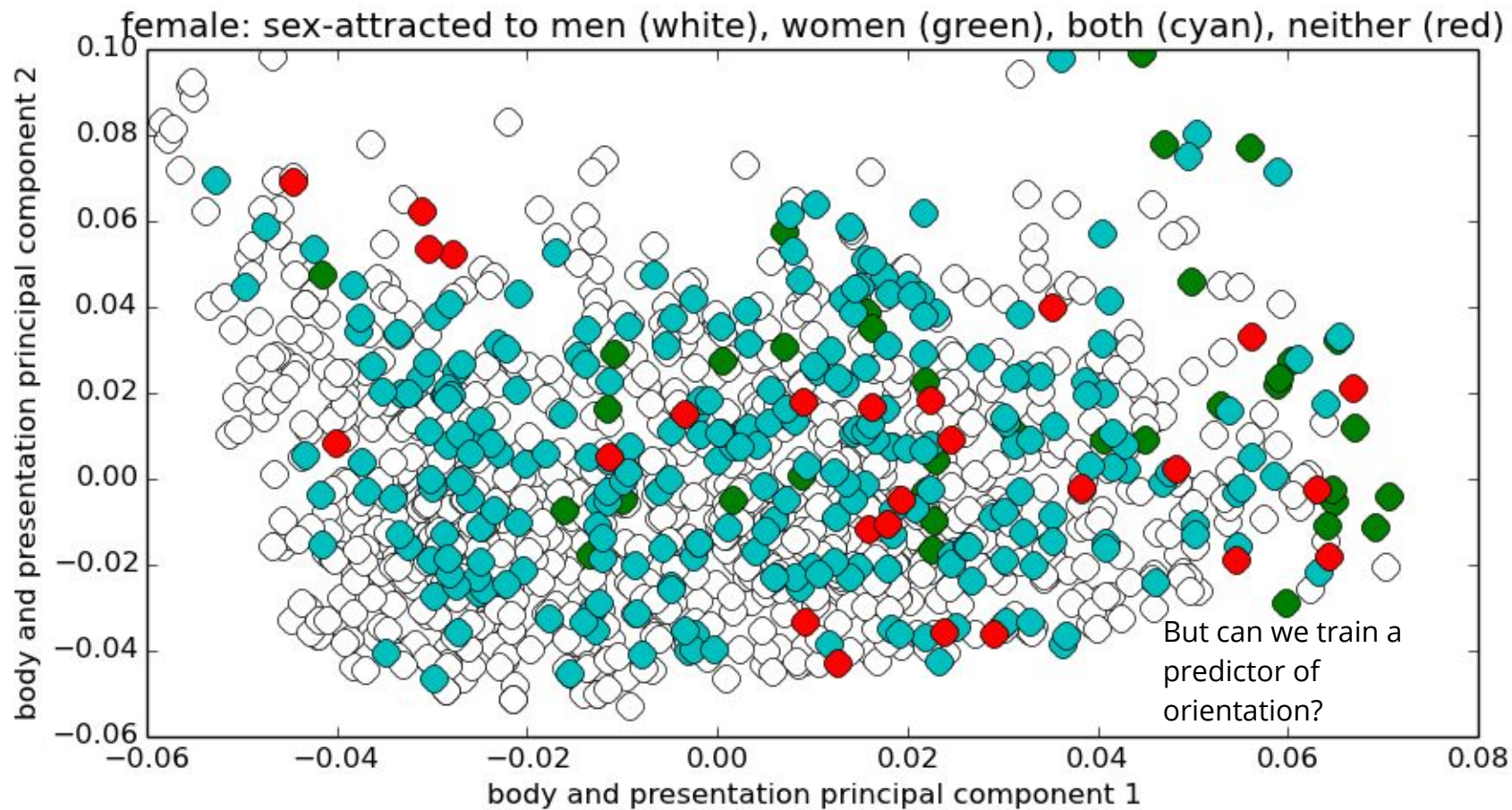












			Pearson's ϕ	Error-rate	Error-at-chance	
	body	predicts	identifies-only-male[all]	0.969 ± 0.008	1.33% ± 0.34%	42.57% ± 0.58%
	presentation	predicts	identifies-only-male[all]	0.981 ± 0.006	0.81% ± 0.25%	42.59% ± 0.72%
	body	predicts	identifies-only-female[all]	0.960 ± 0.011	2.32% ± 0.64%	58.60% ± 0.52%
	presentation	predicts	identifies-only-female[all]	0.971 ± 0.006	1.70% ± 0.38%	58.38% ± 0.54%
	body	predicts	same-sex-attracted[women]	0.081 ± 0.055	69.16% ± 9.59%	73.78% ± 0.93%
	presentation	predicts	same-sex-attracted[women]	0.188 ± 0.042	59.90% ± 3.31%	73.59% ± 0.79%
	body+presentation	predicts	same-sex-attracted[women]	0.224 ± 0.056	57.60% ± 4.51%	73.41% ± 0.65%
	body	predicts	same-sex-attracted[men]	0.033 ± 0.047	79.70% ± 15.09%	90.18% ± 0.55%
	presentation	predicts	same-sex-attracted[men]	0.279 ± 0.068	65.05% ± 6.68%	90.14% ± 0.54%
	body+presentation	predicts	same-sex-attracted[men]	0.293 ± 0.063	63.72% ± 6.19%	90.17% ± 0.49%
	presentation	predicts	only-same-sex-attracted[women]	0.264 ± 0.105	71.71% ± 10.56%	97.29% ± 0.30%
	presentation	predicts	only-same-sex-attracted[men]	0.072 ± 0.112	91.18% ± 11.59%	97.43% ± 0.25%
	presentation	predicts	identifies-lesbian[women]	0.262 ± 0.107	70.55% ± 10.52%	95.45% ± 0.37%
	presentation	predicts	identifies-gay[men]	0.136 ± 0.082	83.74% ± 8.31%	96.76% ± 0.25%
	presentation	predicts	identifies-LGB[women]	0.220 ± 0.045	61.31% ± 3.92%	79.01% ± 0.82%
	presentation	predicts	identifies-LGB[men]	0.273 ± 0.056	65.13% ± 4.96%	89.39% ± 0.42%
	presentation	predicts	only-opp-sex-attracted[women]	0.189 ± 0.048	22.66% ± 2.10%	26.69% ± 0.91%
	presentation	predicts	only-opp-sex-attracted[men]	0.286 ± 0.055	7.30% ± 0.74%	10.37% ± 0.43%
	presentation	predicts	intersex[all]	0.087 ± 0.096	90.14% ± 9.63%	98.56% ± 0.17%
	presentation	predicts	asexual[women]	0.060 ± 0.151	92.13% ± 14.24%	99.29% ± 0.16%

Each row represents an experiment in which a linear model is trained on a randomly chosen 70% of the data, and tested on the remaining 30%. This is repeated 40 times to get means and standard deviations of the measured quantities. Error-rate is the equal error rate, which we get by adjusting the decision threshold until false positives and false negatives are equal (giving us a single number).

			Pearson's ϕ	Error-rate	Error-at-chance
	body predicts	identifies-only-male[all]	0.969 ± 0.008	1.33% ± 0.34%	42.57% ± 0.58%
	presentation predicts	identifies-only-male[all]	0.981 ± 0.006	0.81% ± 0.25%	42.59% ± 0.72%
	body predicts	identifies-only-female[all]	0.960 ± 0.011	2.32% ± 0.64%	58.60% ± 0.52%
	presentation predicts	identifies-only-female[all]	0.971 ± 0.006	1.70% ± 0.38%	58.38% ± 0.54%
	body predicts	same-sex-attracted[women]	0.081 ± 0.055	69.16% ± 9.59%	73.78% ± 0.93%
	presentation predicts	same-sex-attracted[women]	0.188 ± 0.042	59.90% ± 3.31%	73.59% ± 0.79%
body+presentation	predicts	same-sex-attracted[women]	0.224 ± 0.056	57.60% ± 4.51%	73.41% ± 0.65%
	body predicts	same-sex-attracted[men]	0.033 ± 0.047	79.70% ± 15.09%	90.18% ± 0.55%
	presentation predicts	same-sex-attracted[men]	0.279 ± 0.068	65.05% ± 6.68%	90.14% ± 0.54%
body+presentation	predicts	same-sex-attracted[men]	0.293 ± 0.063	63.72% ± 6.19%	90.17% ± 0.49%
	presentation predicts	only-same-sex-attracted[women]	0.264 ± 0.105	71.71% ± 10.56%	97.29% ± 0.30%
	presentation predicts	only-same-sex-attracted[men]	0.072 ± 0.112	91.18% ± 11.59%	97.43% ± 0.25%
	presentation predicts	identifies-lesbian[women]	0.262 ± 0.107	70.55% ± 10.52%	95.45% ± 0.37%
	presentation predicts	identifies-gay[men]	0.136 ± 0.082	83.74% ± 8.31%	96.76% ± 0.25%
	presentation predicts	identifies-LGB[women]	0.220 ± 0.045	61.31% ± 3.92%	79.01% ± 0.82%
	presentation predicts	identifies-LGB[men]	0.273 ± 0.056	65.13% ± 4.96%	89.39% ± 0.42%
	presentation predicts	only-opp-sex-attracted[women]	0.189 ± 0.048	22.66% ± 2.10%	26.69% ± 0.91%
	presentation predicts	only-opp-sex-attracted[men]	0.286 ± 0.055	7.30% ± 0.74%	10.37% ± 0.43%
	presentation predicts	intersex[all]	0.087 ± 0.096	90.14% ± 9.63%	98.56% ± 0.17%
	presentation predicts	asexual[women]	0.060 ± 0.151	92.13% ± 14.24%	99.29% ± 0.16%

Bold items are powerfully correlated enough to explain why we are usually comfortable assuming a pronoun without asking when we meet someone— though there are individuals in the middle who are ambiguous, and the data suggest that even when presentation seems unambiguous we will on (rare) occasion get it wrong.

			Pearson's ϕ	Error-rate	Error-at-chance
	body predicts	identifies-only-male[all]	0.969 ± 0.008	1.33% ± 0.34%	42.57% ± 0.58%
	presentation predicts	identifies-only-male[all]	0.981 ± 0.006	0.81% ± 0.25%	42.59% ± 0.72%
	body predicts	identifies-only-female[all]	0.960 ± 0.011	2.32% ± 0.64%	58.60% ± 0.52%
	presentation predicts	identifies-only-female[all]	0.971 ± 0.006	1.70% ± 0.38%	58.38% ± 0.54%
	body predicts	same-sex-attracted[women]	0.081 ± 0.055	69.16% ± 9.59%	73.78% ± 0.93%
	presentation predicts	same-sex-attracted[women]	0.188 ± 0.042	59.90% ± 3.31%	73.59% ± 0.79%
body+presentation	predicts	same-sex-attracted[women]	0.224 ± 0.056	57.60% ± 4.51%	73.41% ± 0.65%
	body predicts	same-sex-attracted[men]	0.033 ± 0.047	79.70% ± 15.09%	90.18% ± 0.55%
	presentation predicts	same-sex-attracted[men]	0.279 ± 0.068	65.05% ± 6.68%	90.14% ± 0.54%
body+presentation	predicts	same-sex-attracted[men]	0.293 ± 0.063	63.72% ± 6.19%	90.17% ± 0.49%
	presentation predicts	only-same-sex-attracted[women]	0.264 ± 0.105	71.71% ± 10.56%	97.29% ± 0.30%
	presentation predicts	only-same-sex-attracted[men]	0.072 ± 0.112	91.18% ± 11.59%	97.43% ± 0.25%
	presentation predicts	identifies-lesbian[women]	0.262 ± 0.107	70.55% ± 10.52%	95.45% ± 0.37%
	presentation predicts	identifies-gay[men]	0.136 ± 0.082	83.74% ± 8.31%	96.76% ± 0.25%
	presentation predicts	identifies-LGB[women]	0.220 ± 0.045	61.31% ± 3.92%	79.01% ± 0.82%
	presentation predicts	identifies-LGB[men]	0.273 ± 0.056	65.13% ± 4.96%	89.39% ± 0.42%
	presentation predicts	only-opp-sex-attracted[women]	0.189 ± 0.048	22.66% ± 2.10%	26.69% ± 0.91%
	presentation predicts	only-opp-sex-attracted[men]	0.286 ± 0.055	7.30% ± 0.74%	10.37% ± 0.43%
	presentation predicts	intersex[all]	0.087 ± 0.096	90.14% ± 9.63%	98.56% ± 0.17%
	presentation predicts	asexual[women]	0.060 ± 0.151	92.13% ± 14.24%	99.29% ± 0.16%

Non-bold items in black show clear correlation, and represent categories most of us are familiar with. However the numbers show us that while there may be a category stereotype, using it to pattern-match will get us in trouble, yielding many false positives and false negatives.

			Pearson's ϕ	Error-rate	Error-at-chance
	body predicts	identifies-only-male[all]	0.969 ± 0.008	1.33% ± 0.34%	42.57% ± 0.58%
	presentation predicts	identifies-only-male[all]	0.981 ± 0.006	0.81% ± 0.25%	42.59% ± 0.72%
	body predicts	identifies-only-female[all]	0.960 ± 0.011	2.32% ± 0.64%	58.60% ± 0.52%
	presentation predicts	identifies-only-female[all]	0.971 ± 0.006	1.70% ± 0.38%	58.38% ± 0.54%
	body predicts	same-sex-attracted[women]	0.081 ± 0.055	69.16% ± 9.59%	73.78% ± 0.93%
	presentation predicts	same-sex-attracted[women]	0.188 ± 0.042	59.90% ± 3.31%	73.59% ± 0.79%
body+presentation	predicts	same-sex-attracted[women]	0.224 ± 0.056	57.60% ± 4.51%	73.41% ± 0.65%
	body predicts	same-sex-attracted[men]	0.033 ± 0.047	79.70% ± 15.09%	90.18% ± 0.55%
	presentation predicts	same-sex-attracted[men]	0.279 ± 0.068	65.05% ± 6.68%	90.14% ± 0.54%
body+presentation	predicts	same-sex-attracted[men]	0.293 ± 0.063	63.72% ± 6.19%	90.17% ± 0.49%
	presentation predicts	only-same-sex-attracted[women]	0.264 ± 0.105	71.71% ± 10.56%	97.29% ± 0.30%
	presentation predicts	only-same-sex-attracted[men]	0.072 ± 0.112	91.18% ± 11.59%	97.43% ± 0.25%
	presentation predicts	identifies-lesbian[women]	0.262 ± 0.107	70.55% ± 10.52%	95.45% ± 0.37%
	presentation predicts	identifies-gay[men]	0.136 ± 0.082	83.74% ± 8.31%	96.76% ± 0.25%
	presentation predicts	identifies-LGB[women]	0.220 ± 0.045	61.31% ± 3.92%	79.01% ± 0.82%
	presentation predicts	identifies-LGB[men]	0.273 ± 0.056	65.13% ± 4.96%	89.39% ± 0.42%
	presentation predicts	only-opp-sex-attracted[women]	0.189 ± 0.048	22.66% ± 2.10%	26.69% ± 0.91%
	presentation predicts	only-opp-sex-attracted[men]	0.286 ± 0.055	7.30% ± 0.74%	10.37% ± 0.43%
	presentation predicts	intersex[all]	0.087 ± 0.096	90.14% ± 9.63%	98.56% ± 0.17%
	presentation predicts	asexual[women]	0.060 ± 0.151	92.13% ± 14.24%	99.29% ± 0.16%

Yellow items are likely real correlations, but are weak.

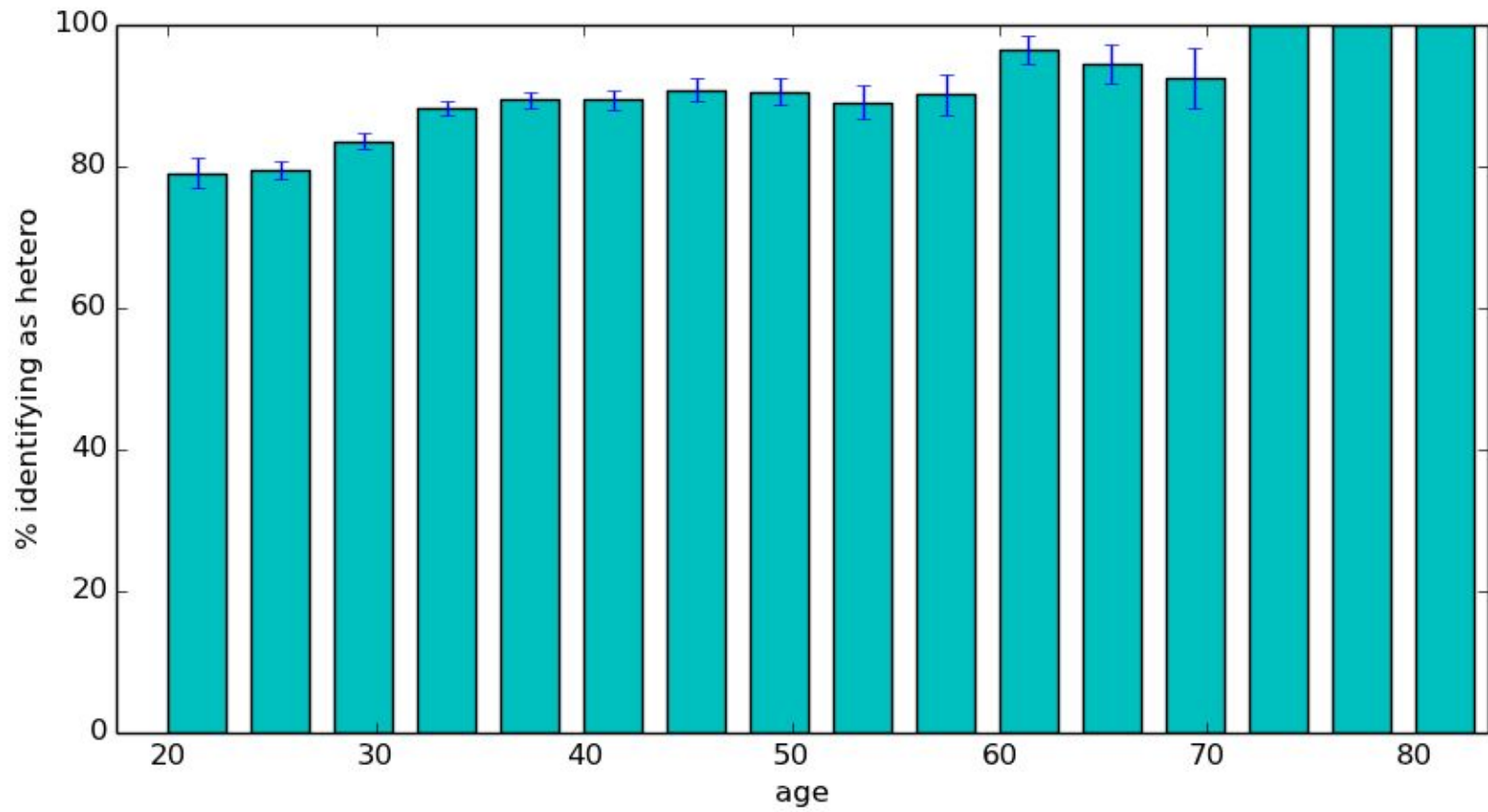
			Pearson's ϕ	Error-rate	Error-at-chance
	body predicts	identifies-only-male[all]	0.969 ± 0.008	1.33% ± 0.34%	42.57% ± 0.58%
	presentation predicts	identifies-only-male[all]	0.981 ± 0.006	0.81% ± 0.25%	42.59% ± 0.72%
	body predicts	identifies-only-female[all]	0.960 ± 0.011	2.32% ± 0.64%	58.60% ± 0.52%
	presentation predicts	identifies-only-female[all]	0.971 ± 0.006	1.70% ± 0.38%	58.38% ± 0.54%
	body predicts	same-sex-attracted[women]	0.081 ± 0.055	69.16% ± 9.59%	73.78% ± 0.93%
	presentation predicts	same-sex-attracted[women]	0.188 ± 0.042	59.90% ± 3.31%	73.59% ± 0.79%
body+presentation	predicts	same-sex-attracted[women]	0.224 ± 0.056	57.60% ± 4.51%	73.41% ± 0.65%
	body predicts	same-sex-attracted[men]	0.033 ± 0.047	79.70% ± 15.09%	90.18% ± 0.55%
	presentation predicts	same-sex-attracted[men]	0.279 ± 0.068	65.05% ± 6.68%	90.14% ± 0.54%
body+presentation	predicts	same-sex-attracted[men]	0.293 ± 0.063	63.72% ± 6.19%	90.17% ± 0.49%
	presentation predicts	only-same-sex-attracted[women]	0.264 ± 0.105	71.71% ± 10.56%	97.29% ± 0.30%
	presentation predicts	only-same-sex-attracted[men]	0.072 ± 0.112	91.18% ± 11.59%	97.43% ± 0.25%
	presentation predicts	identifies-lesbian[women]	0.262 ± 0.107	70.55% ± 10.52%	95.45% ± 0.37%
	presentation predicts	identifies-gay[men]	0.136 ± 0.082	83.74% ± 8.31%	96.76% ± 0.25%
	presentation predicts	identifies-LGB[women]	0.220 ± 0.045	61.31% ± 3.92%	79.01% ± 0.82%
	presentation predicts	identifies-LGB[men]	0.273 ± 0.056	65.13% ± 4.96%	89.39% ± 0.42%
	presentation predicts	only-opp-sex-attracted[women]	0.189 ± 0.048	22.66% ± 2.10%	26.69% ± 0.91%
	presentation predicts	only-opp-sex-attracted[men]	0.286 ± 0.055	7.30% ± 0.74%	10.37% ± 0.43%
	presentation predicts	intersex[all]	0.087 ± 0.096	90.14% ± 9.63%	98.56% ± 0.17%
	presentation predicts	asexual[women]	0.060 ± 0.151	92.13% ± 14.24%	99.29% ± 0.16%

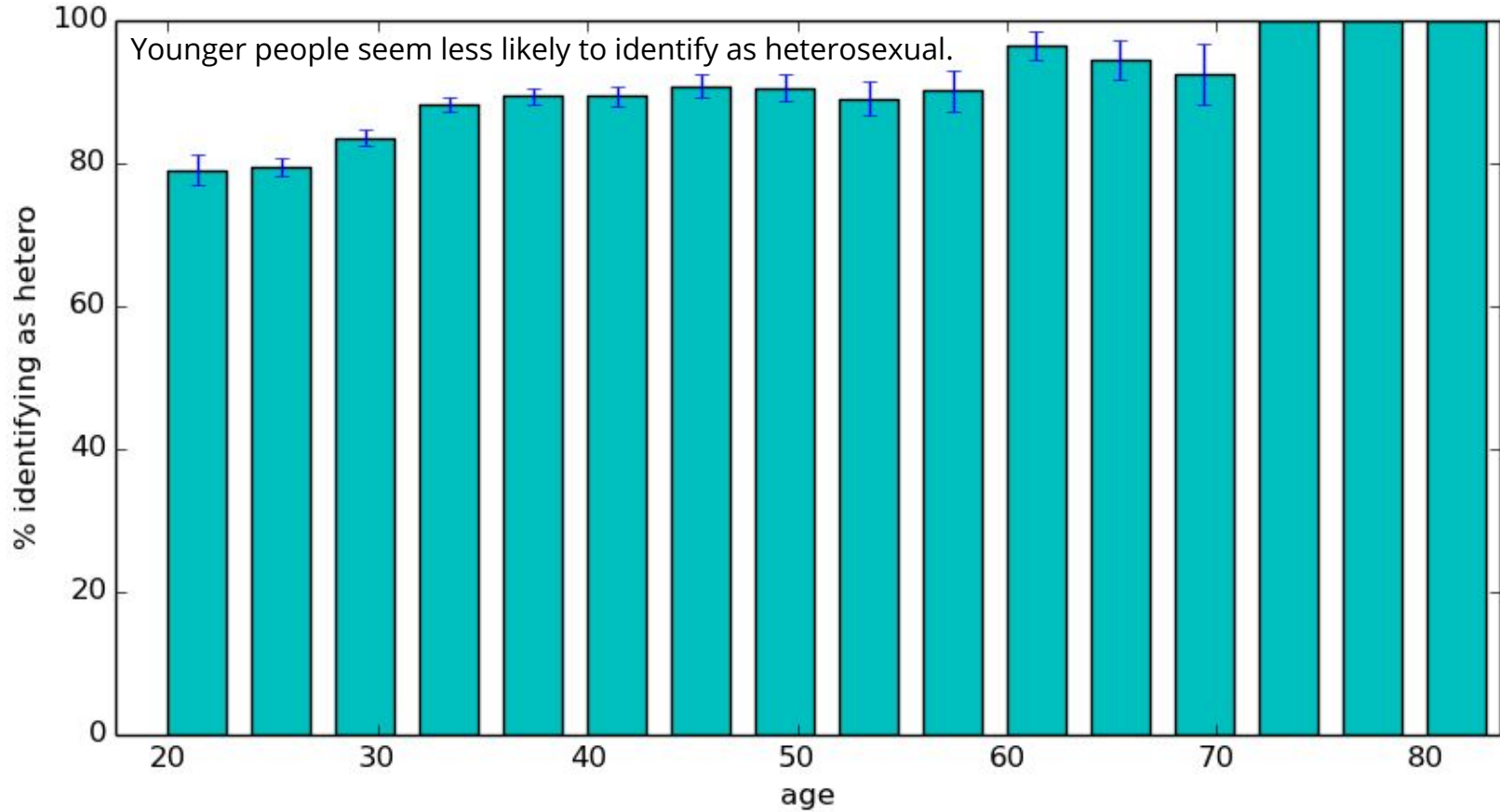
Red items are essentially at chance.

			Pearson's ϕ	Error-rate	Error-at-chance
	body predicts	identifies-only-male[all]	0.969 ± 0.008	1.33% ± 0.34%	42.57% ± 0.58%
	presentation predicts	identifies-only-male[all]	0.981 ± 0.006	0.81% ± 0.25%	42.59% ± 0.72%
	body predicts	identifies-only-female[all]	0.960 ± 0.011	2.32% ± 0.64%	58.60% ± 0.52%
	presentation predicts	identifies-only-female[all]	0.971 ± 0.006	1.70% ± 0.38%	58.38% ± 0.54%
	body predicts	same-sex-attracted[women]	0.081 ± 0.055	69.16% ± 9.59%	73.78% ± 0.93%
	presentation predicts	same-sex-attracted[women]	0.188 ± 0.042	59.90% ± 3.31%	73.59% ± 0.79%
body+presentation	predicts	same-sex-attracted[women]	0.224 ± 0.056	57.60% ± 4.51%	73.41% ± 0.65%
	body predicts	same-sex-attracted[men]	0.033 ± 0.047	79.70% ± 15.09%	90.18% ± 0.55%
	presentation predicts	same-sex-attracted[men]	0.279 ± 0.068	65.05% ± 6.68%	90.14% ± 0.54%
body+presentation	predicts	same-sex-attracted[men]	0.293 ± 0.063	63.72% ± 6.19%	90.17% ± 0.49%
	presentation predicts	only-same-sex-attracted[women]	0.264 ± 0.105	71.71% ± 10.56%	97.29% ± 0.30%
	presentation predicts	only-same-sex-attracted[men]	0.072 ± 0.112	91.18% ± 11.59%	97.43% ± 0.25%
	presentation predicts	identifies-lesbian[women]	0.262 ± 0.107	70.55% ± 10.52%	95.45% ± 0.37%
	presentation predicts	identifies-gay[men]	0.136 ± 0.082	83.74% ± 8.31%	96.76% ± 0.25%
	presentation predicts	identifies-LGB[women]	0.220 ± 0.045	61.31% ± 3.92%	79.01% ± 0.82%
	presentation predicts	identifies-LGB[men]	0.273 ± 0.056	65.13% ± 4.96%	89.39% ± 0.42%
	presentation predicts	only-opp-sex-attracted[women]	0.189 ± 0.048	22.66% ± 2.10%	26.69% ± 0.91%
	presentation predicts	only-opp-sex-attracted[men]	0.286 ± 0.055	7.30% ± 0.74%	10.37% ± 0.43%
	presentation predicts	intersex[all]	0.087 ± 0.096	90.14% ± 9.63%	98.56% ± 0.17%
	presentation predicts	asexual[women]	0.060 ± 0.151	92.13% ± 14.24%	99.29% ± 0.16%

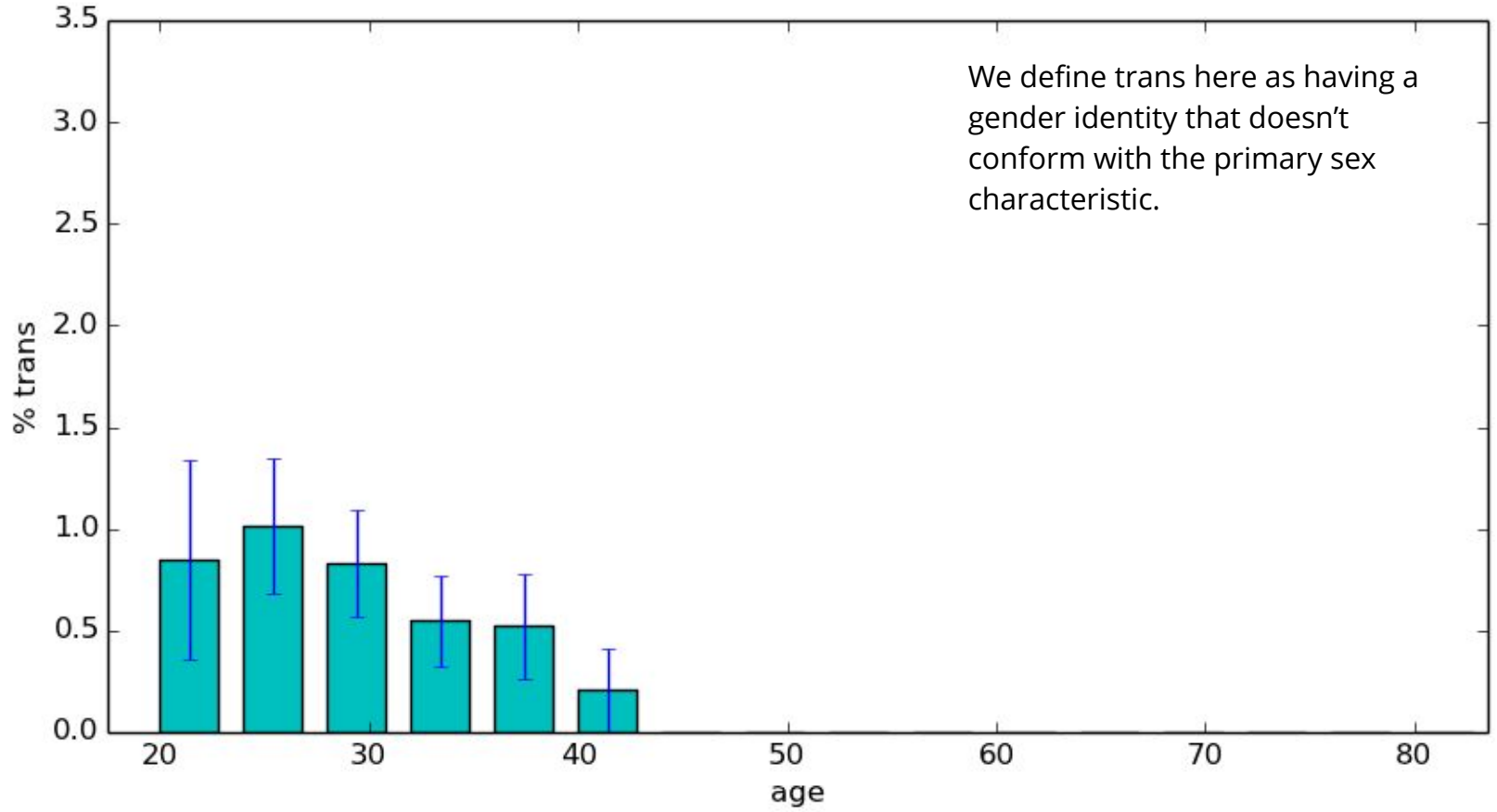
More questions, cleaner data, or a more powerful nonlinear model could bring some reds into yellow territory, and could bring some yellows into black territory, but are unlikely to generate more **boldfaced black** categories based on presentation.

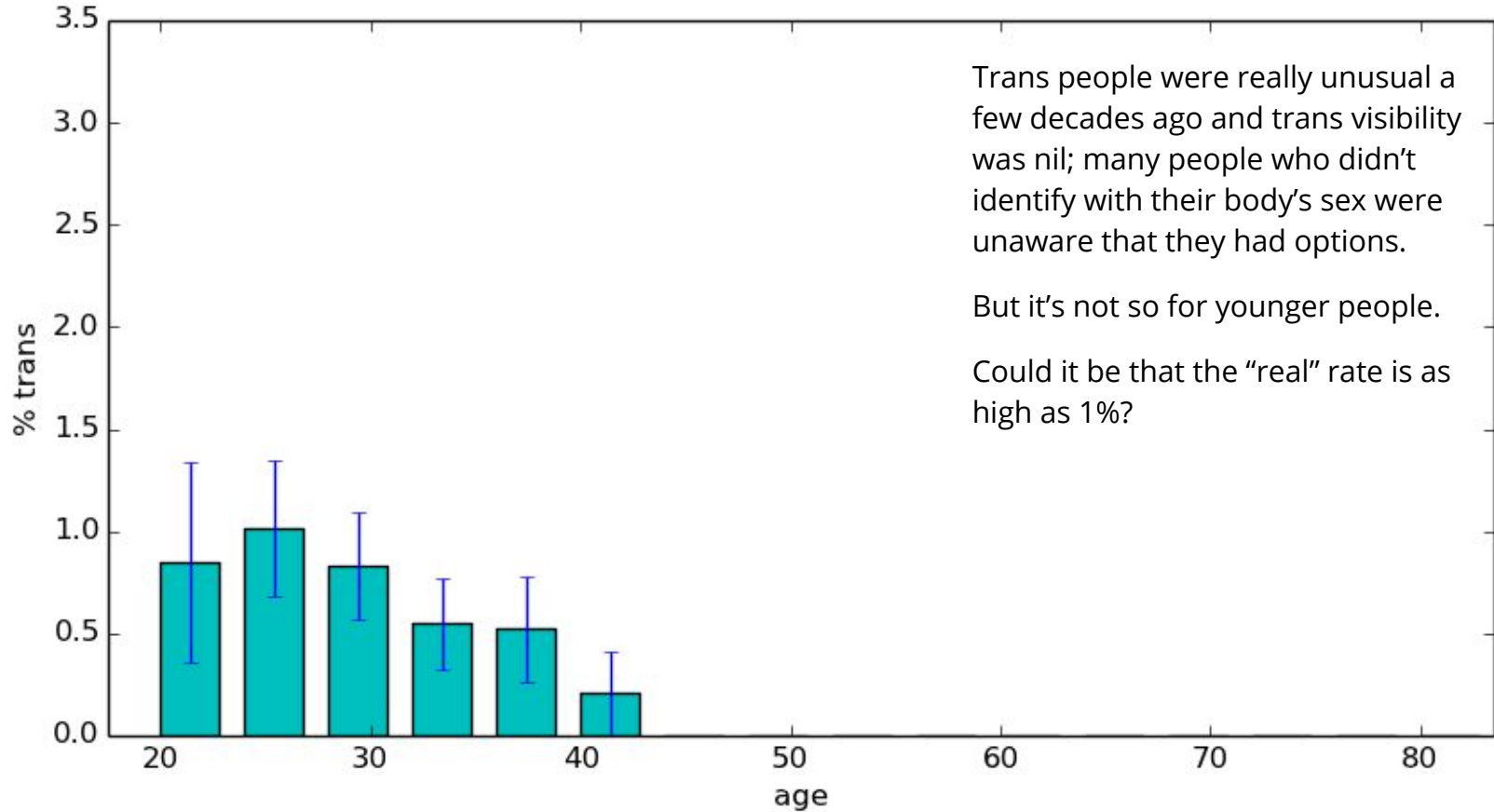
V. The kids are queer

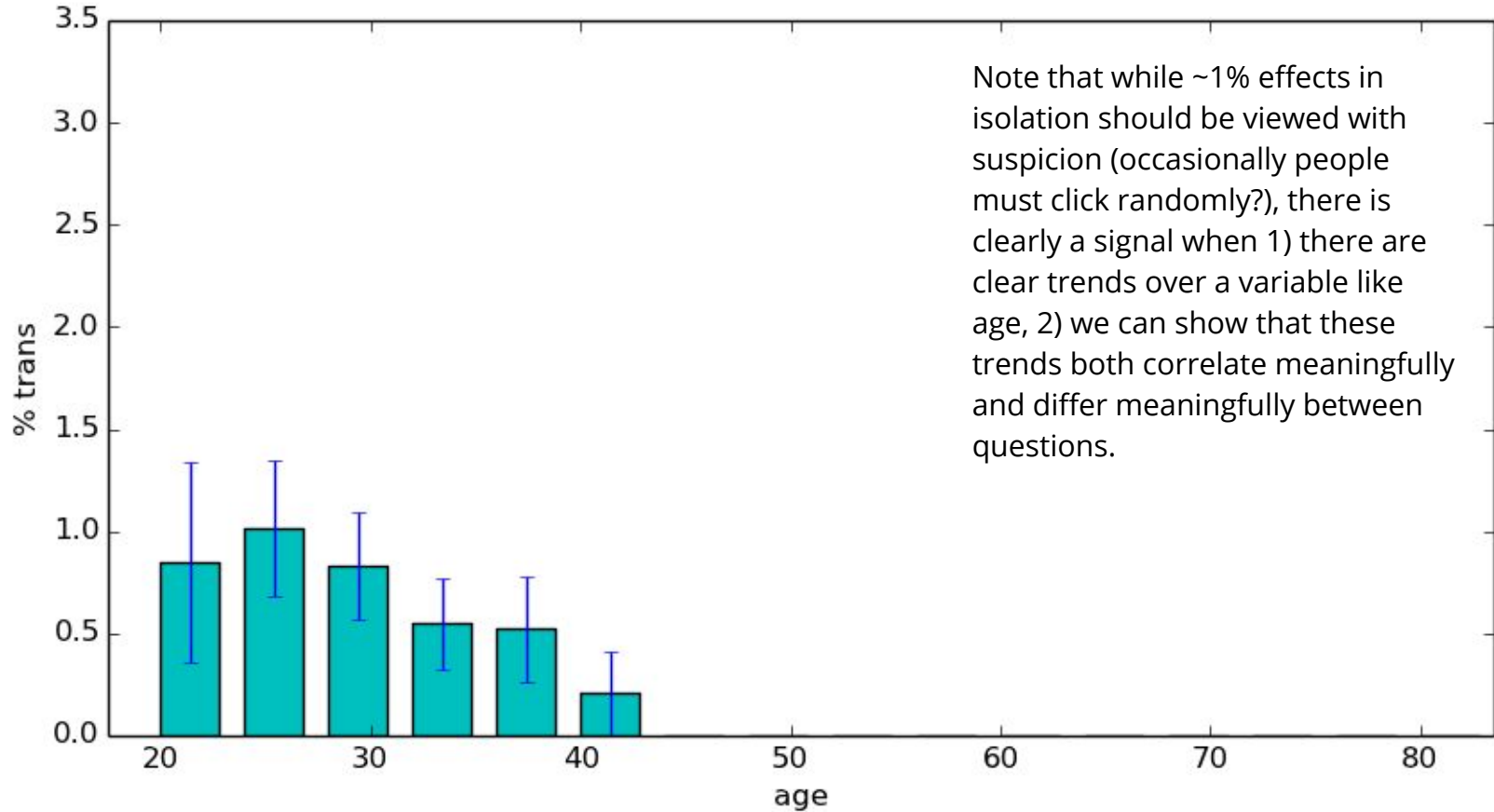




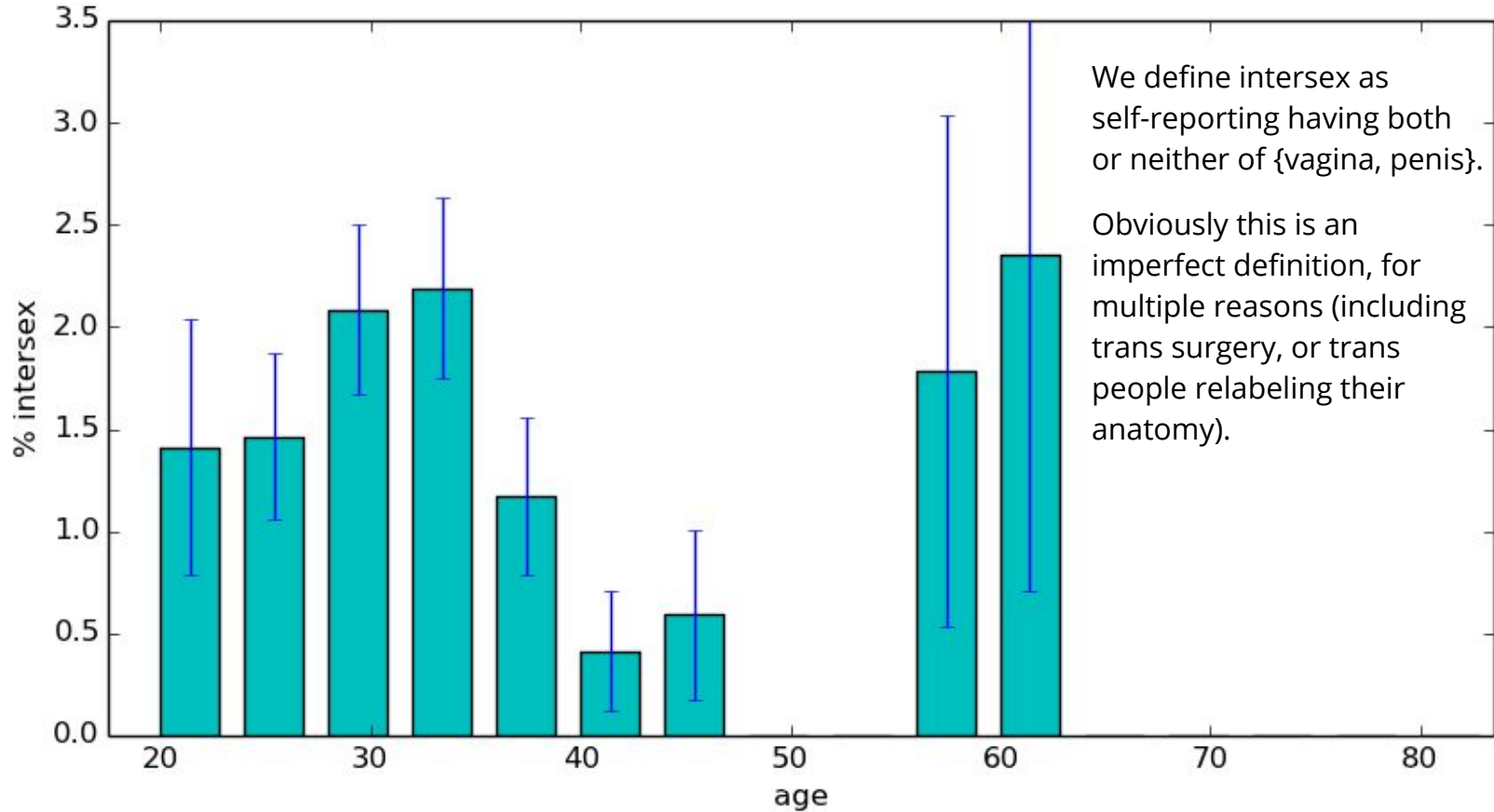
We define trans here as having a gender identity that doesn't conform with the primary sex characteristic.

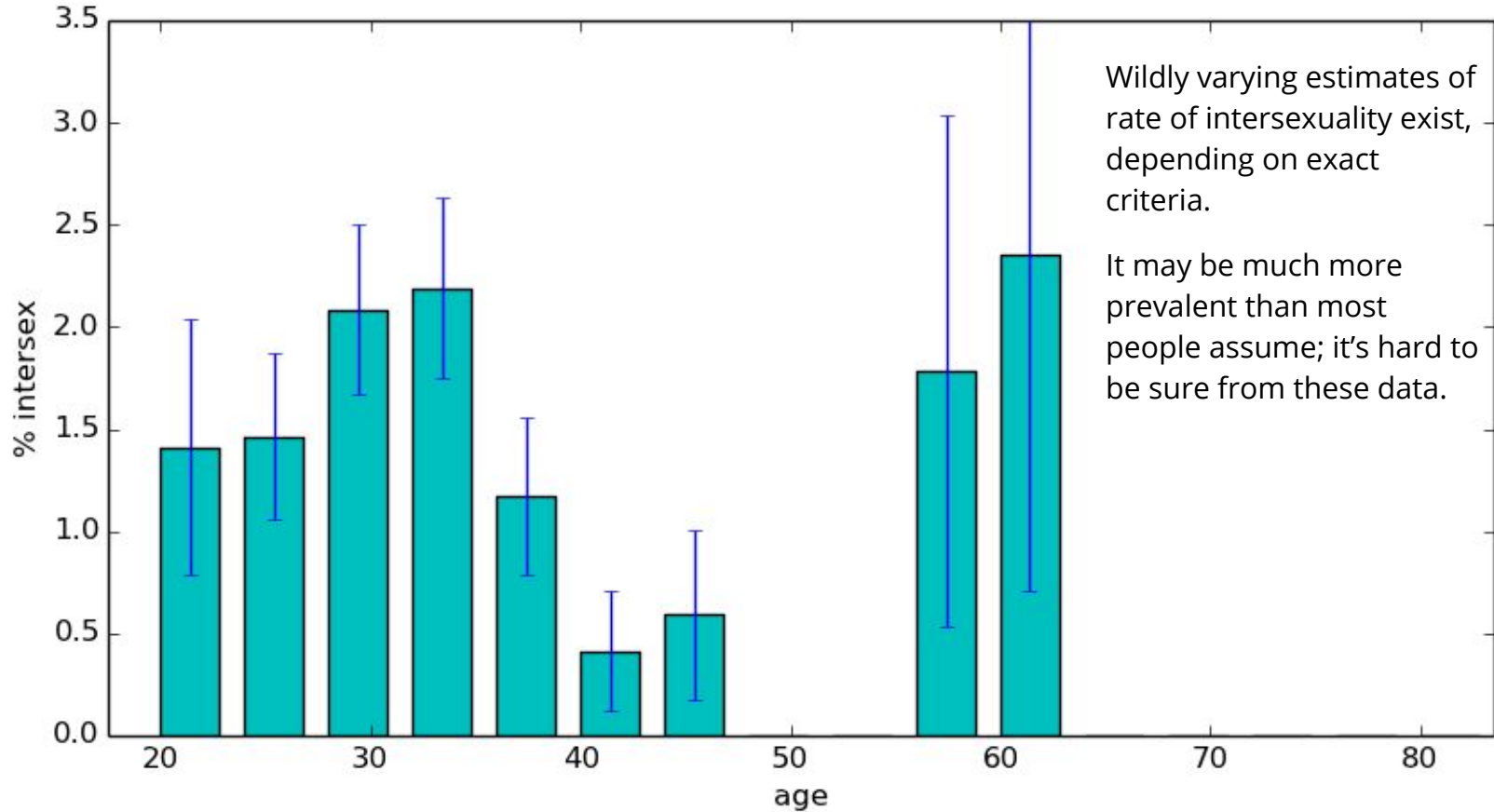


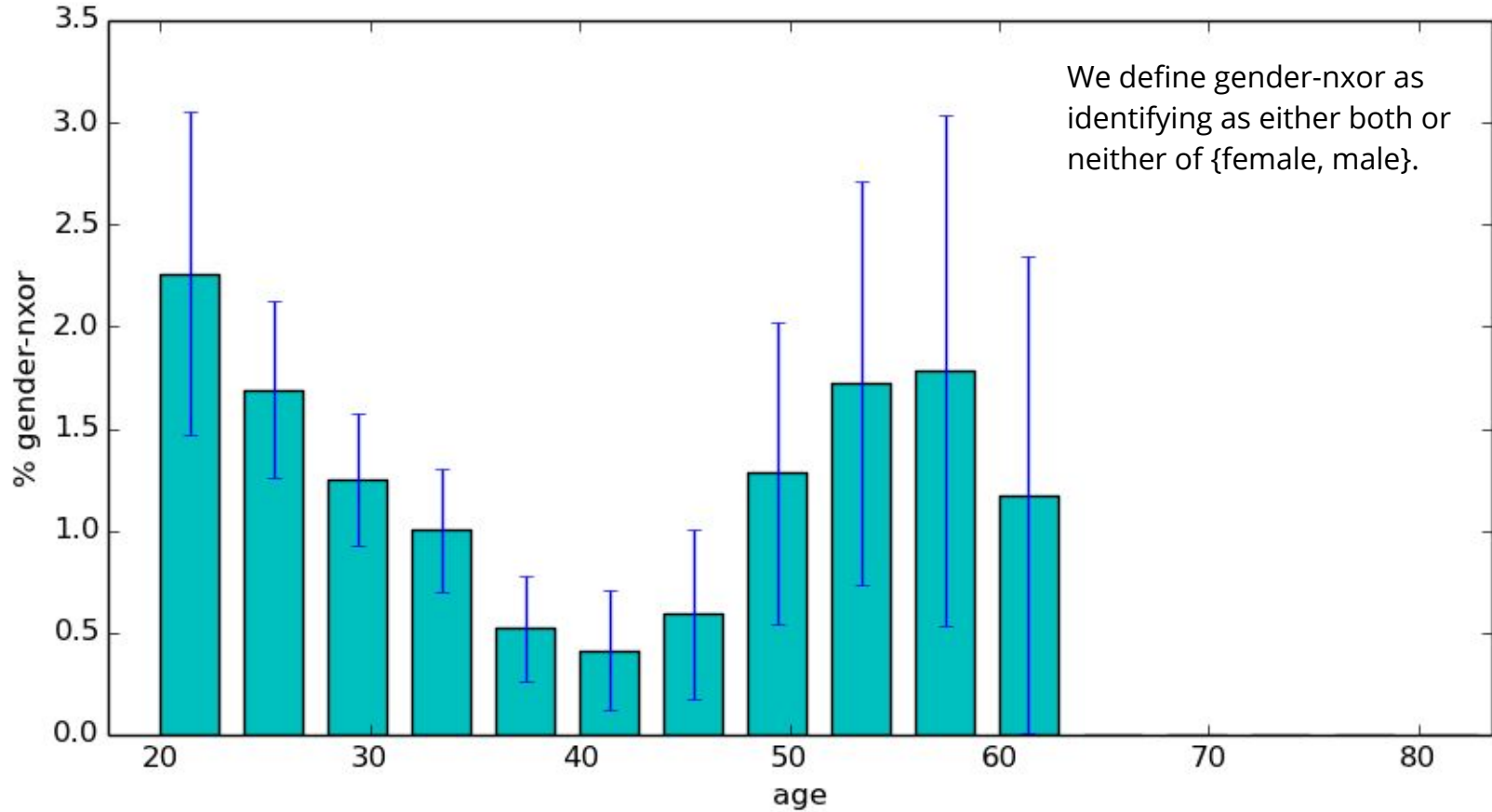


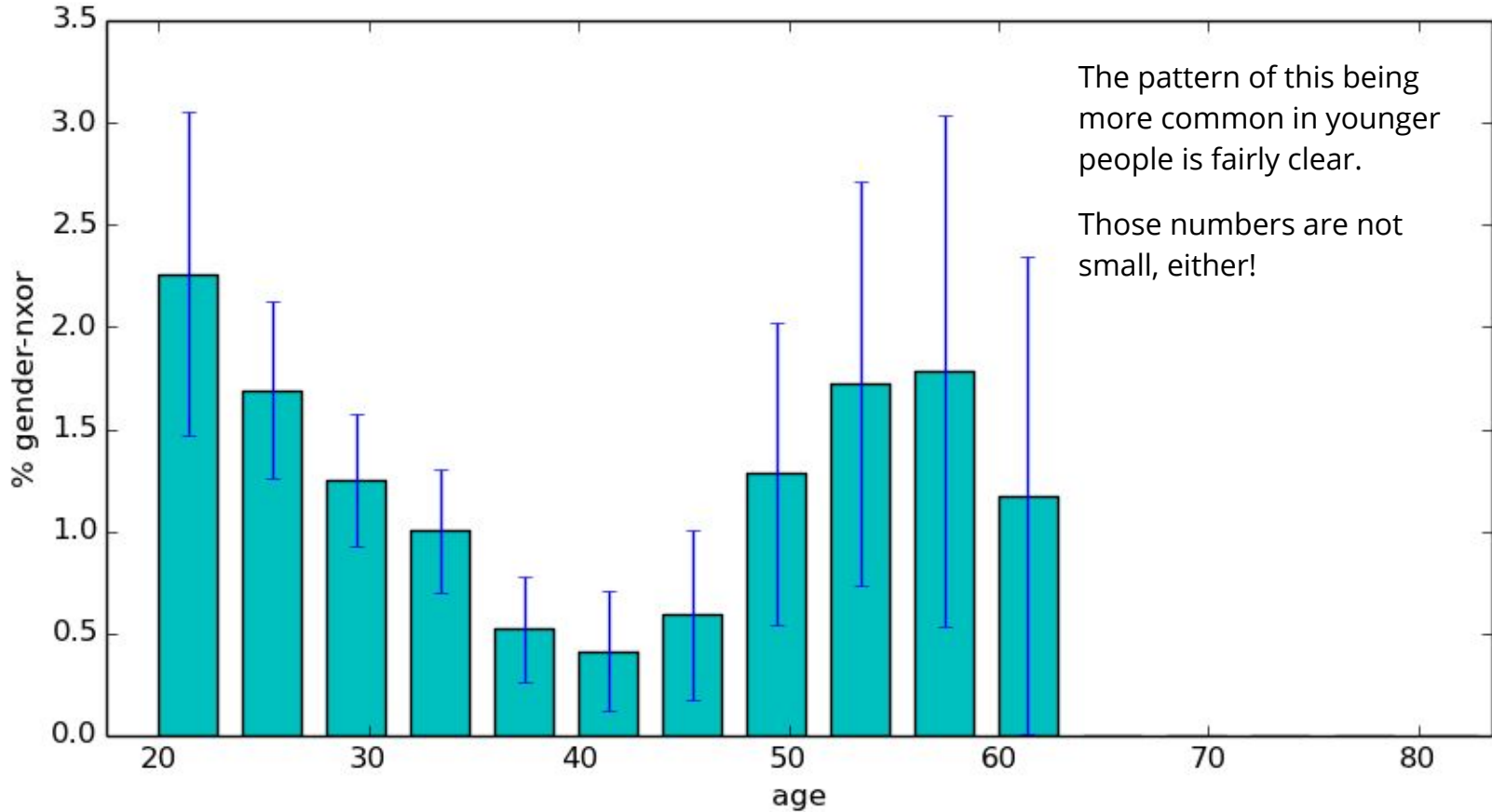


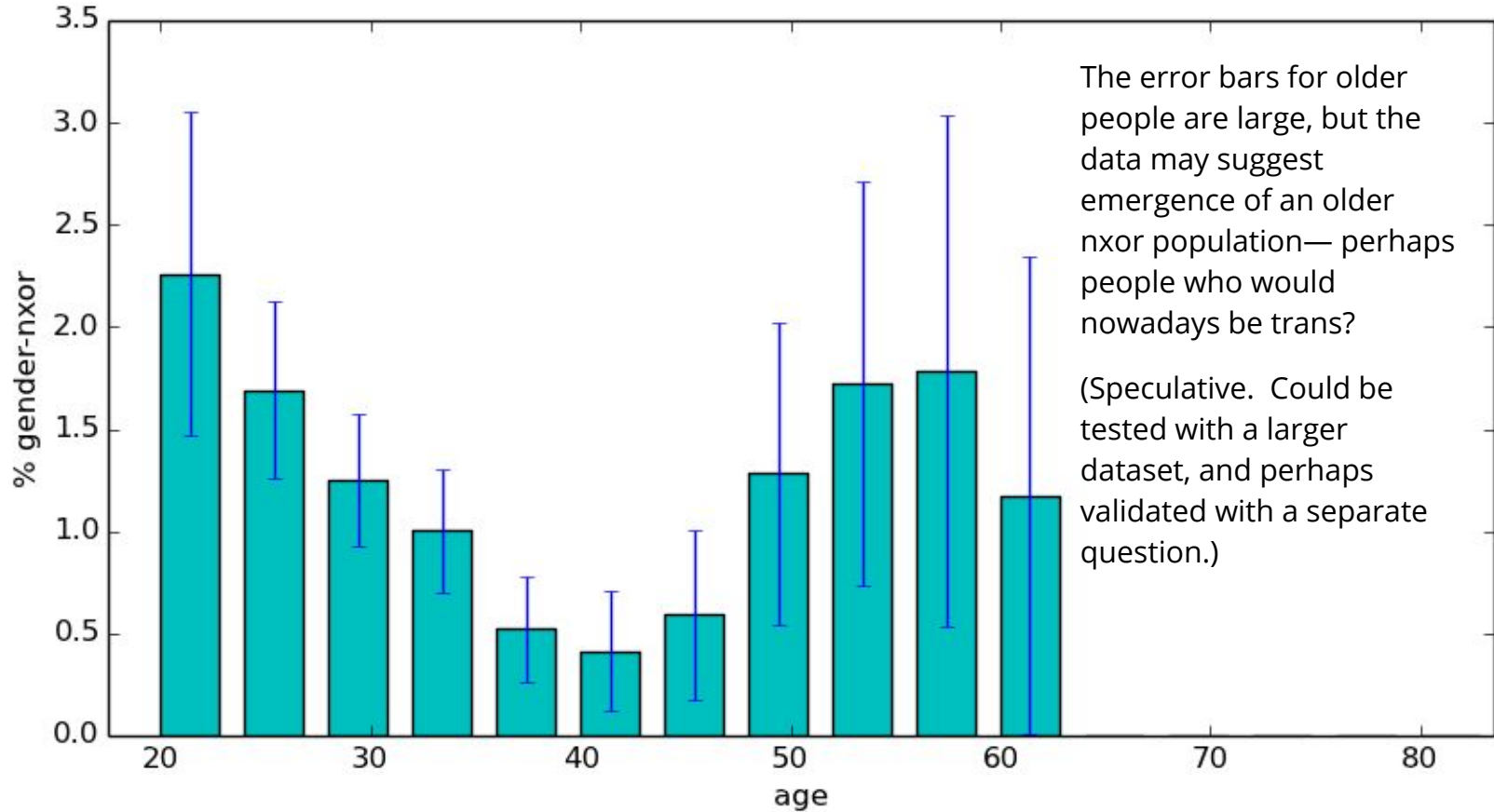
Note that while ~1% effects in isolation should be viewed with suspicion (occasionally people must click randomly?), there is clearly a signal when 1) there are clear trends over a variable like age, 2) we can show that these trends both correlate meaningfully and differ meaningfully between questions.

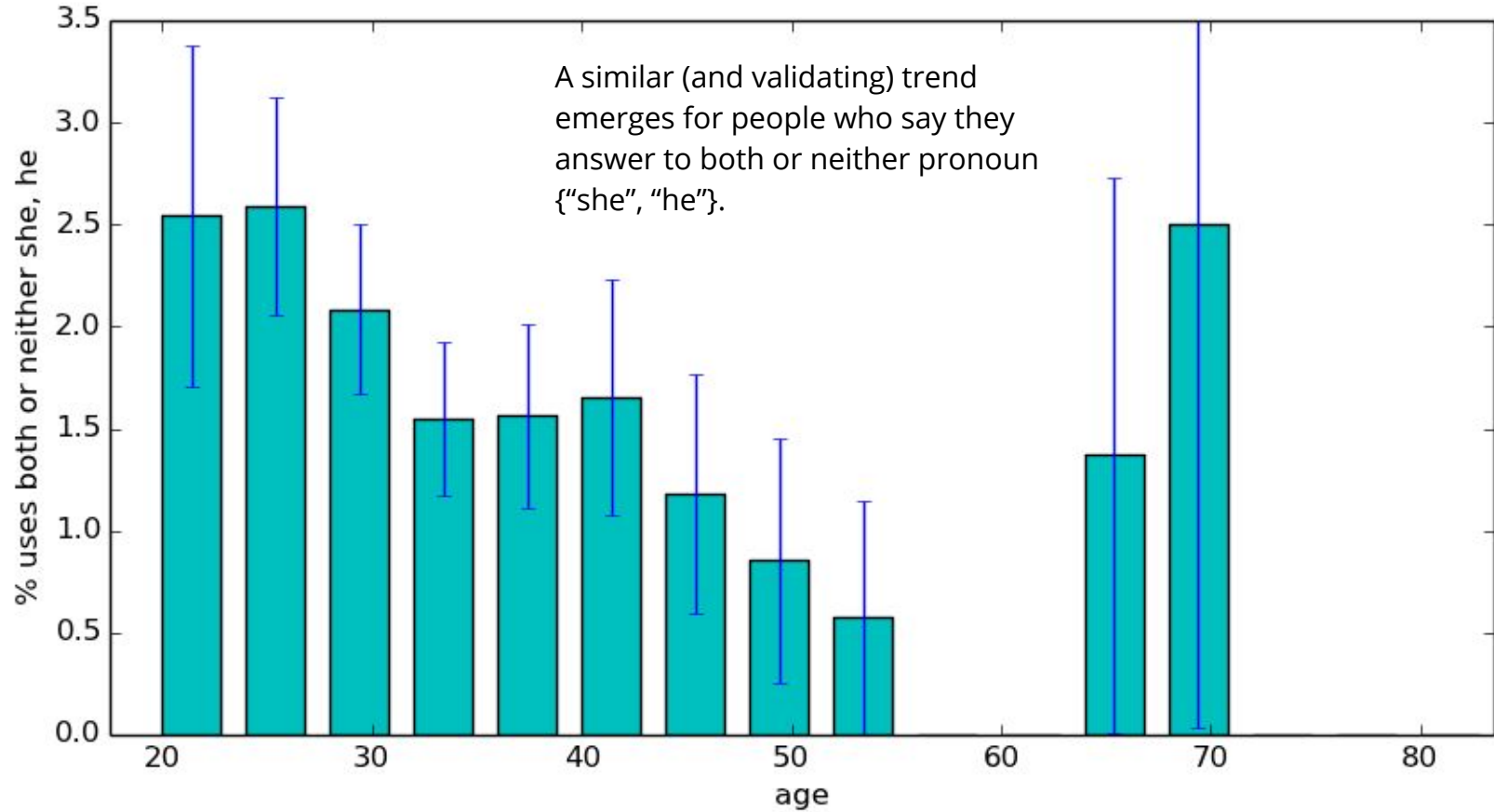


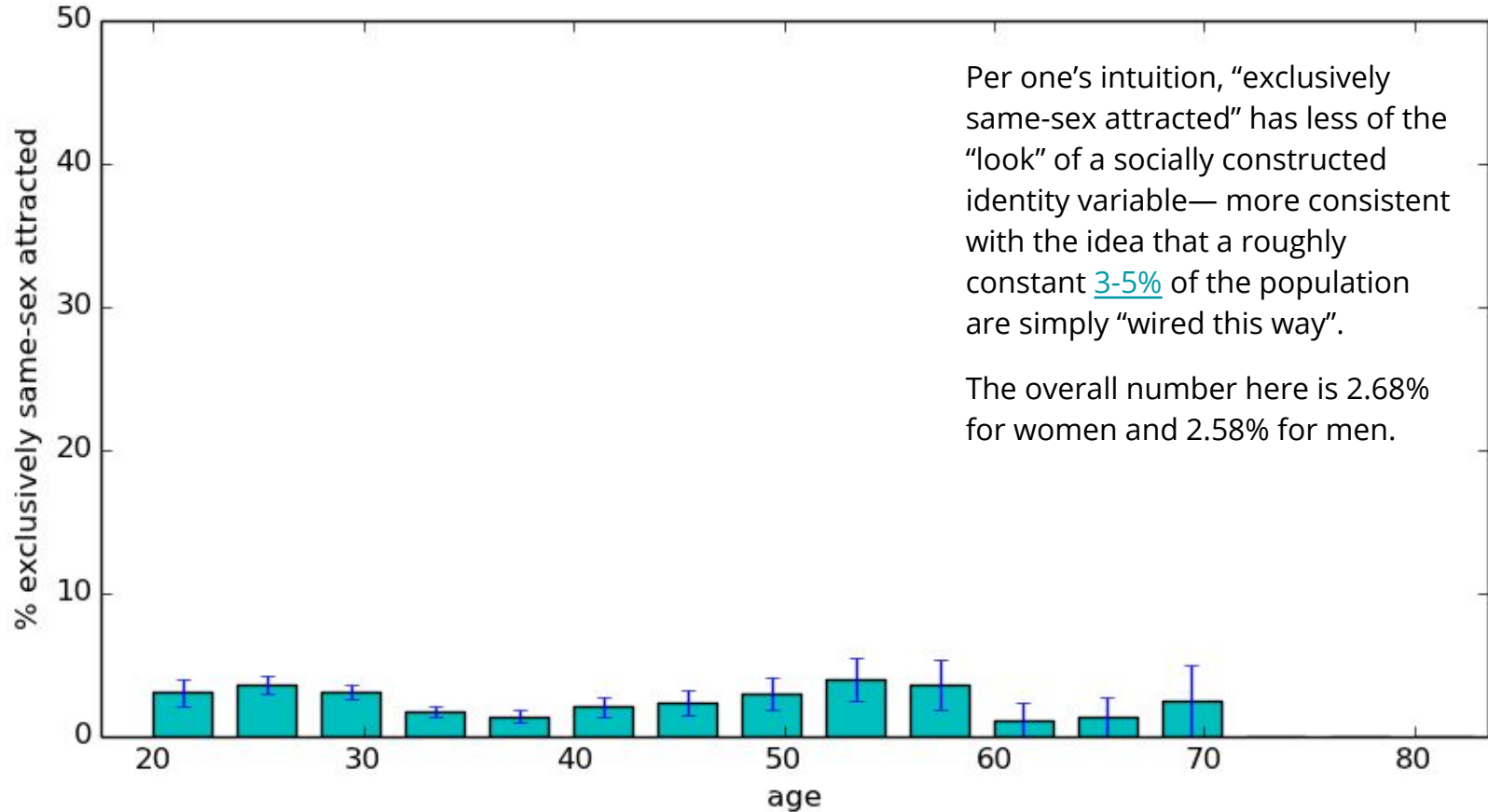


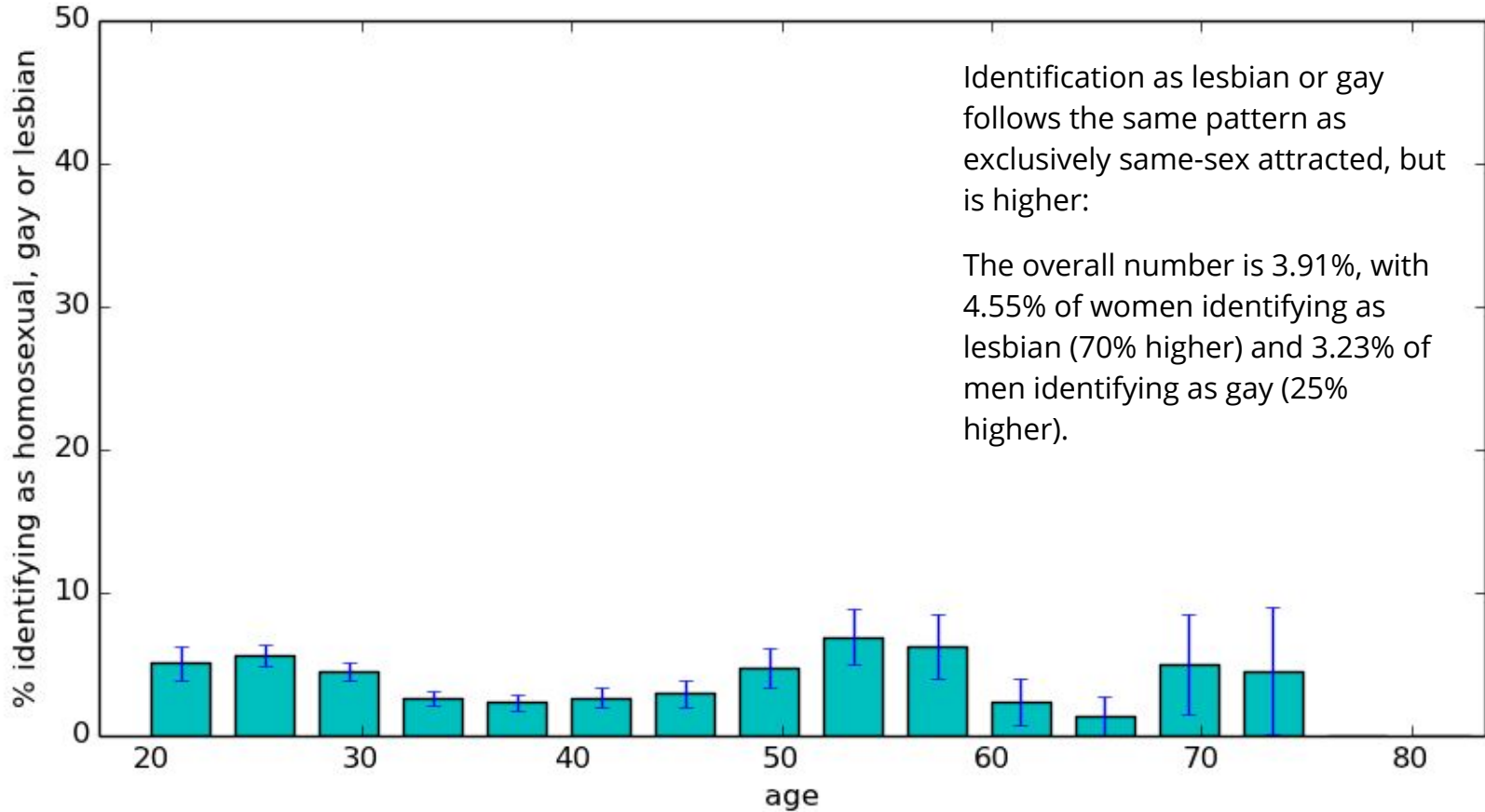


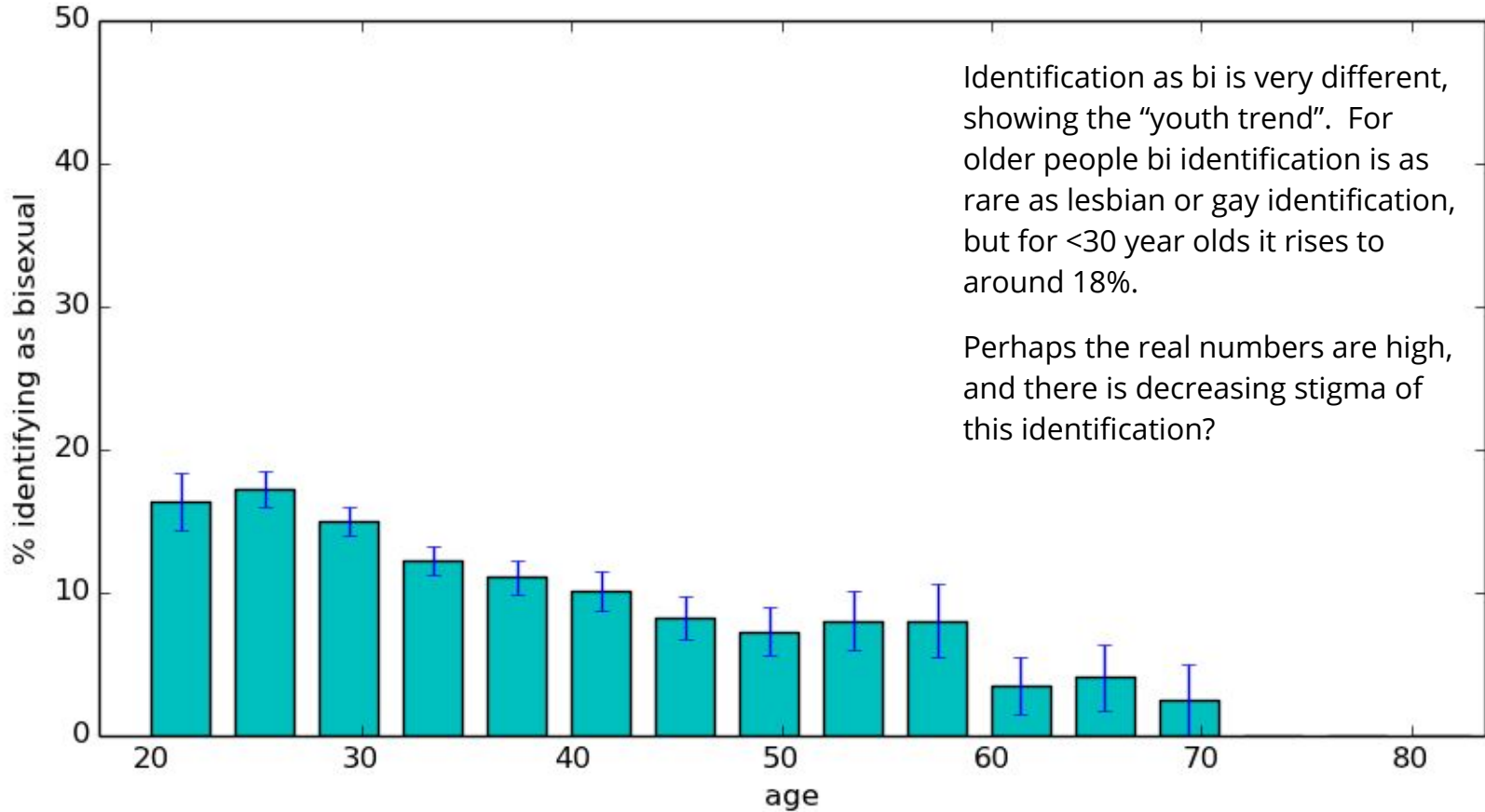


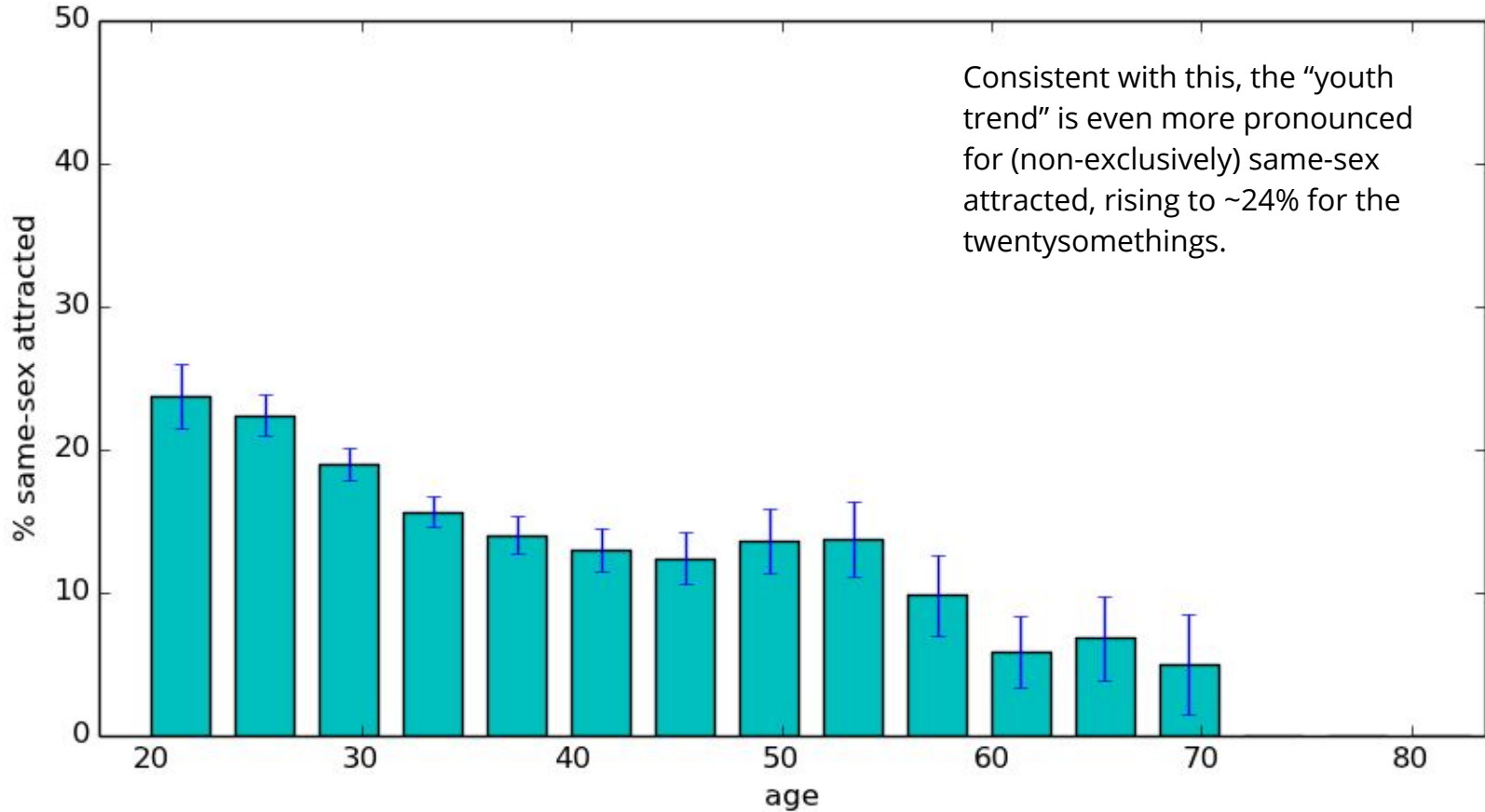


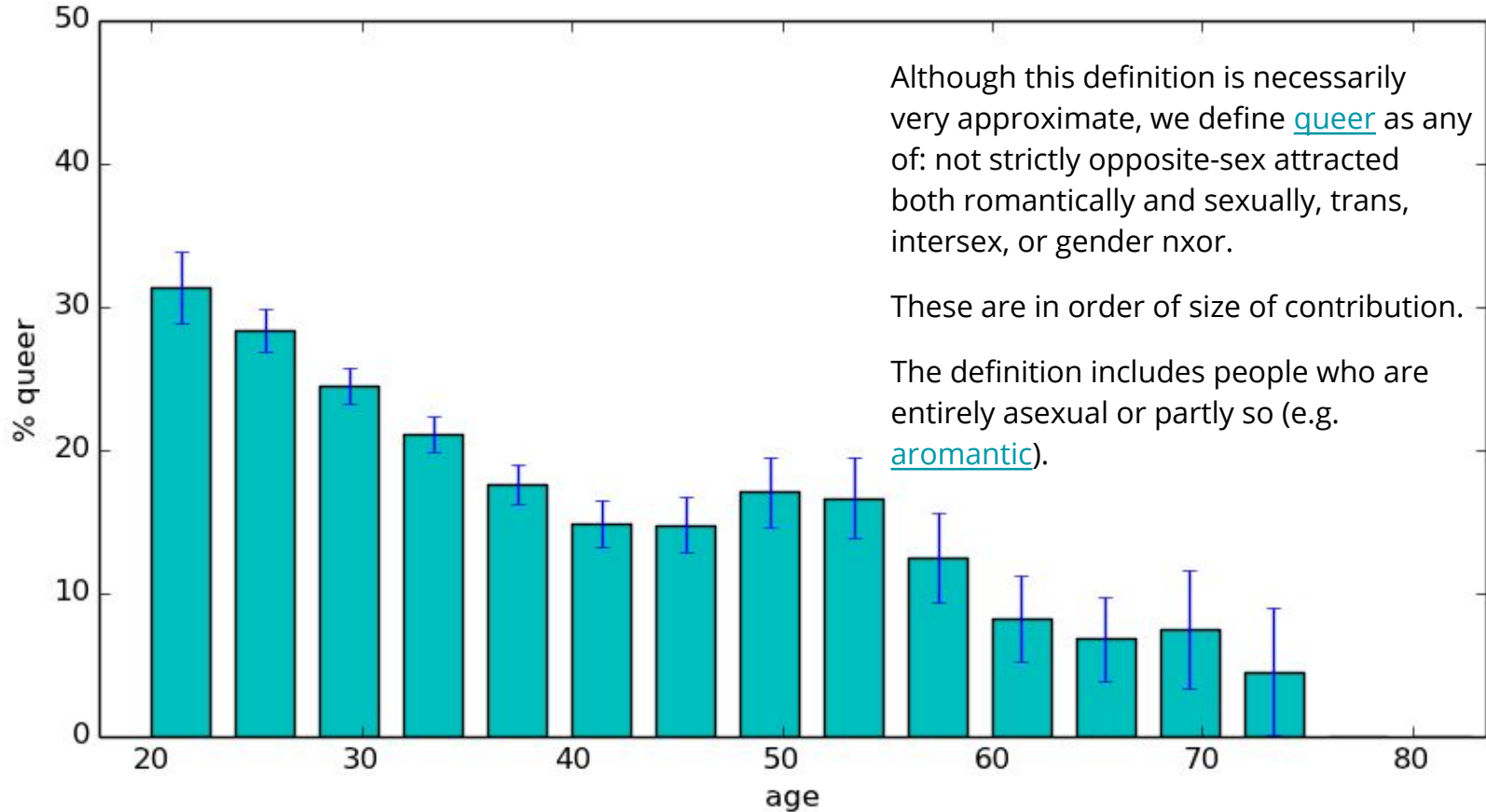








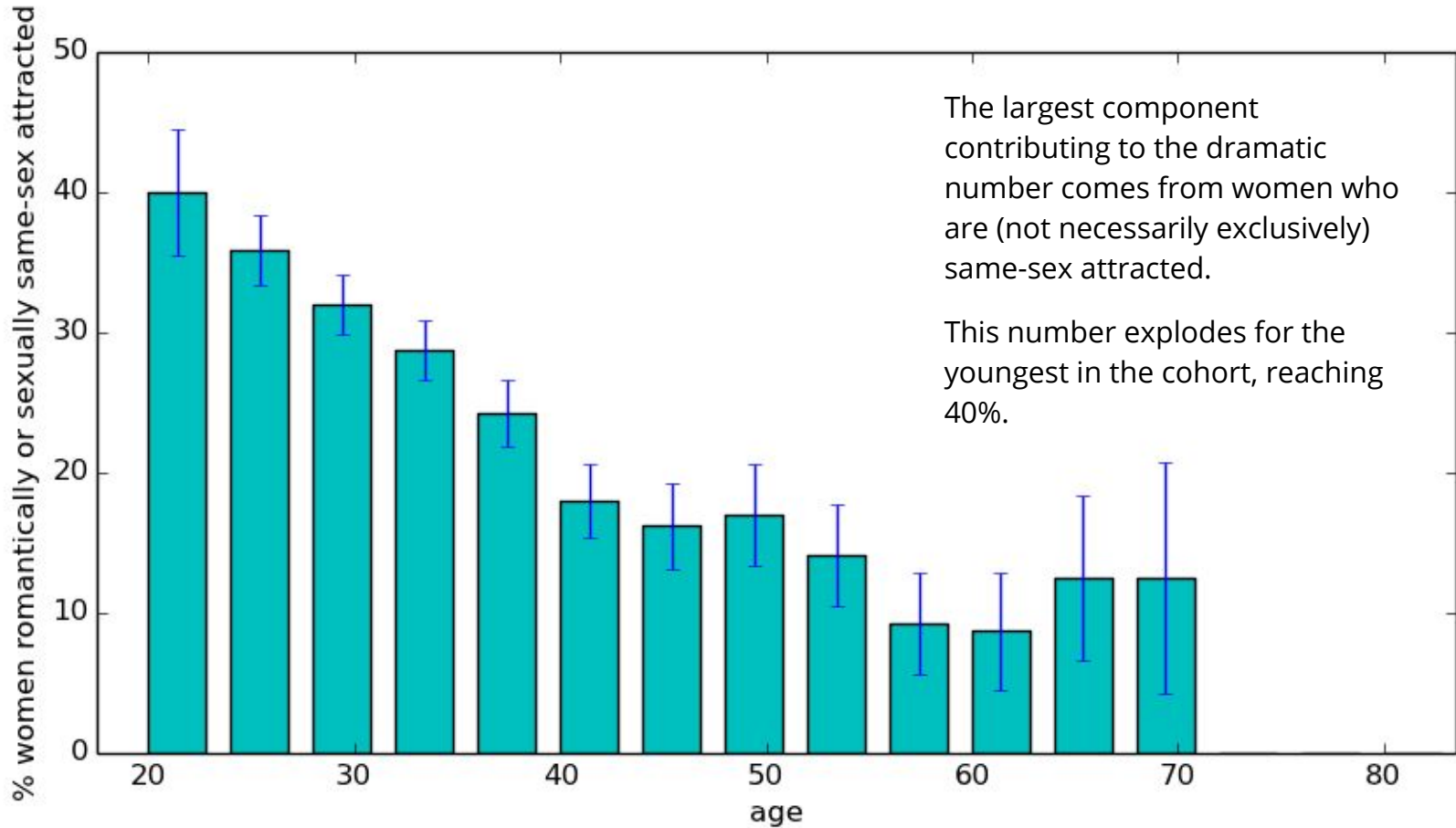


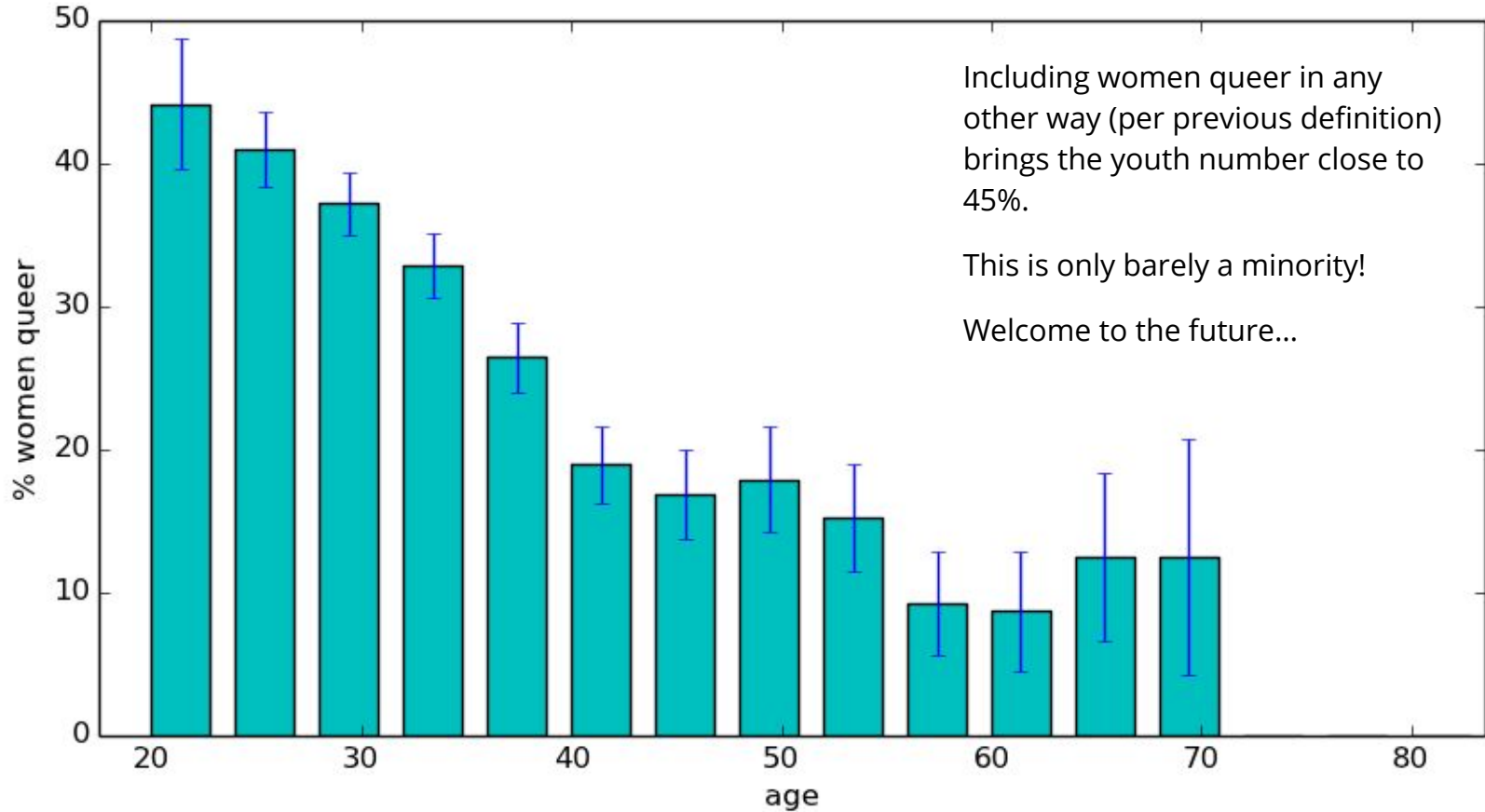


Although this definition is necessarily very approximate, we define [queer](#) as any of: not strictly opposite-sex attracted both romantically and sexually, trans, intersex, or gender nxor.

These are in order of size of contribution.

The definition includes people who are entirely asexual or partly so (e.g. [aromantic](#)).





Including women queer in any other way (per previous definition) brings the youth number close to 45%.

This is only barely a minority!

Welcome to the future...

So again:

There are no rules for identity.

We see the gender binary.

We see the gender “spectrum”.

We see that sex-gender-sexuality-orientation is in fact higher-dimensional.

Presentation predicts identity more strongly than body data does.

Body and behavior are strong predictors of gender identity, weaker predictors of other identities.

More people are intersex, trans, or gender nonconforming than you probably think.

Same-sex sexual attraction is more common than same-sex romantic attraction.

Same-sex attraction among women is extremely common.

Young people are increasingly nonconformant to either old identities or gender/sexuality binaries.

Attempts to validate identity with a body correlate will fail and do harm (cf. [sex testing in the Olympics](#)).