

# Pre- and post-election EDA

## Exploratory data analysis

### Event expectation

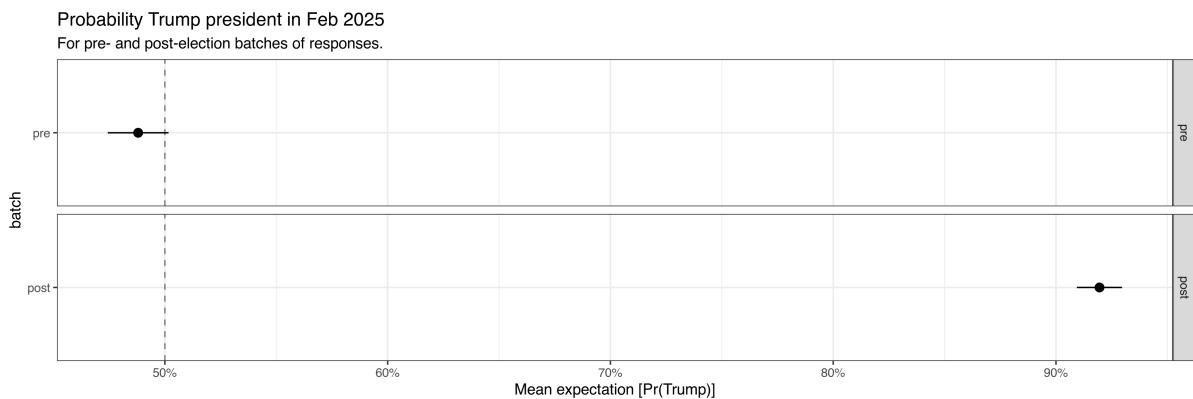


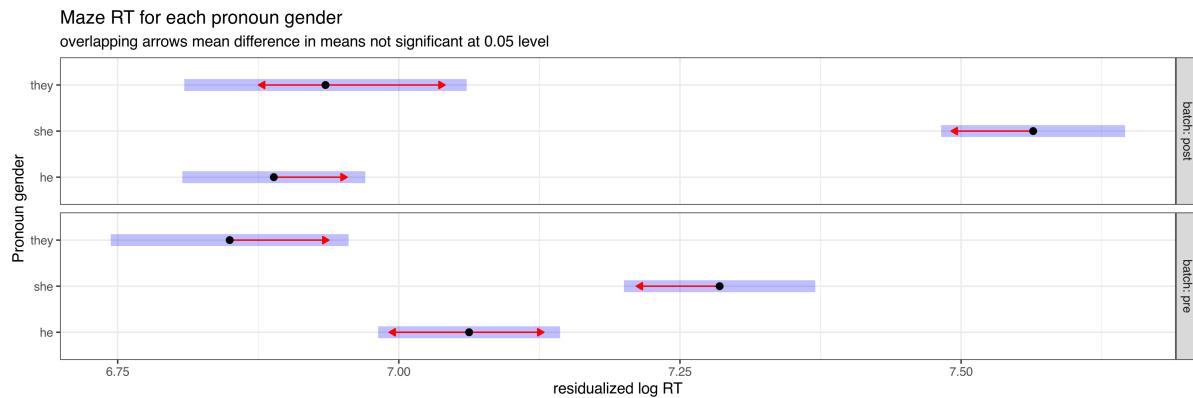
Table 1: One-sample t-tests comparing Trump probability to 50%

batch	n	mean_prob	t_stat	p_value	ci_lower	ci_upper
pre	1255	0.488	-1.724	0.085	0.474	0.502
post	1243	0.920	81.142	0.000	0.909	0.930

^ In the pre-election batch, mean expectations were 48.801%, and this was not significantly different from 50%. In the post-election batch, mean expectations were 91.954%, and this was significantly different from 50%.

## Maze

Pairwise comparisons between genders (first pronoun only): Within each batch {pre-election, post-election}, is there a significant difference in reading times for each pairwise comparison among {male,female,neutral} pronouns?



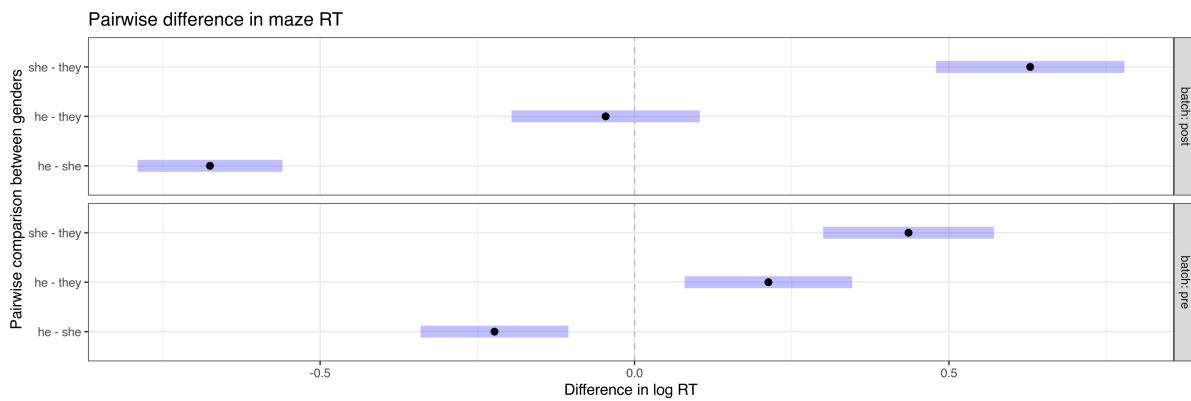
batch = post:

	pronoun_gender_pairwise	estimate	SE	df	t.ratio	p.value
he - she		-0.675	0.0587	452	-11.506	<.0001
he - they		-0.046	0.0761	452	-0.604	0.5464
she - they		0.629	0.0763	452	8.252	<.0001

batch = pre:

	pronoun_gender_pairwise	estimate	SE	df	t.ratio	p.value
he - she		-0.223	0.0597	452	-3.729	0.0002
he - they		0.213	0.0677	452	3.143	0.0018
she - they		0.436	0.0691	452	6.308	<.0001

Results are given on the log (not the response) scale.



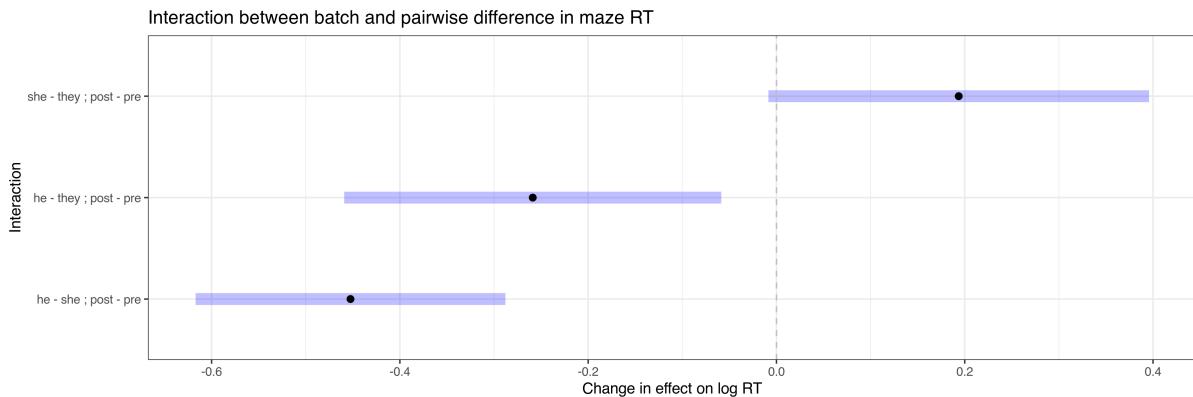
^ Pairwise comparisons between genders show, for Maze RTs:

- in pre-election data, RT of female pronouns is larger than that of male pronouns, which in turn are slower than gender-neutral pronouns. All differences are significant at 0.05 level.
- in post-election data, RT of female pronouns is significantly larger than that of either male or neutral pronouns respectively, and difference between male and neutral pronouns is not significant.

Now, looking at interaction between each pairwise contrast, pre-vs-post election:

pronoun_gender_pairwise	batch_pairwise	estimate	SE	df	t.ratio	p.value
he - she	post - pre	-0.452	0.0837	452	-5.404	<.0001
he - they	post - pre	-0.259	0.1019	452	-2.540	0.0114
she - they	post - pre	0.194	0.1029	452	1.882	0.0605

Results are given on the log (not the response) scale.

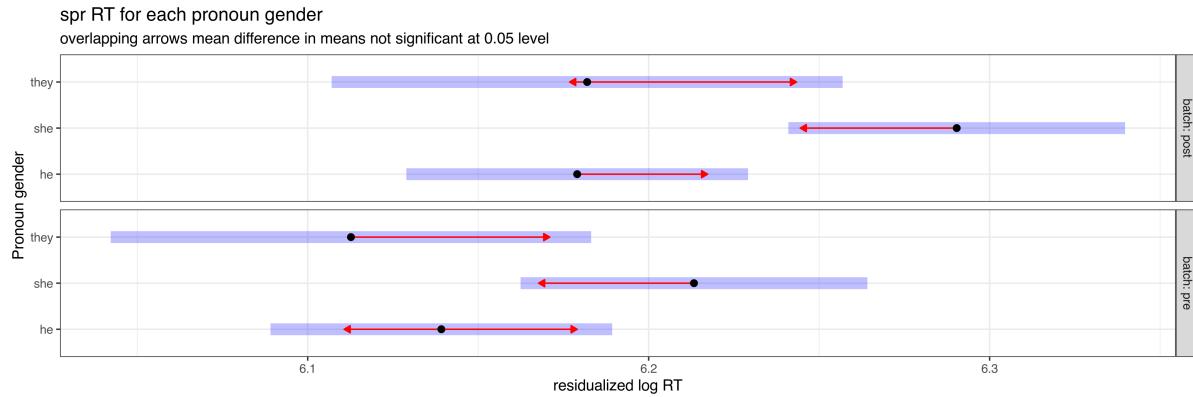


^ interaction between each pairwise contrast among pronouns and experiment batch (pre- vs post-election).

- Comparing post-election minus pre-election
  - log RT difference he-she is negative — that is, the bias for reading male faster than female pronouns increased.
  - log RT difference he-they is negative but smaller, and likewise log RT difference she-they is positive but pretty small.

## SPR

Pairwise comparisons between genders (first pronoun only): Within each batch {pre-election, post-election}, is there a significant difference in reading times for each pairwise comparison among {male,female,neutral} pronouns?



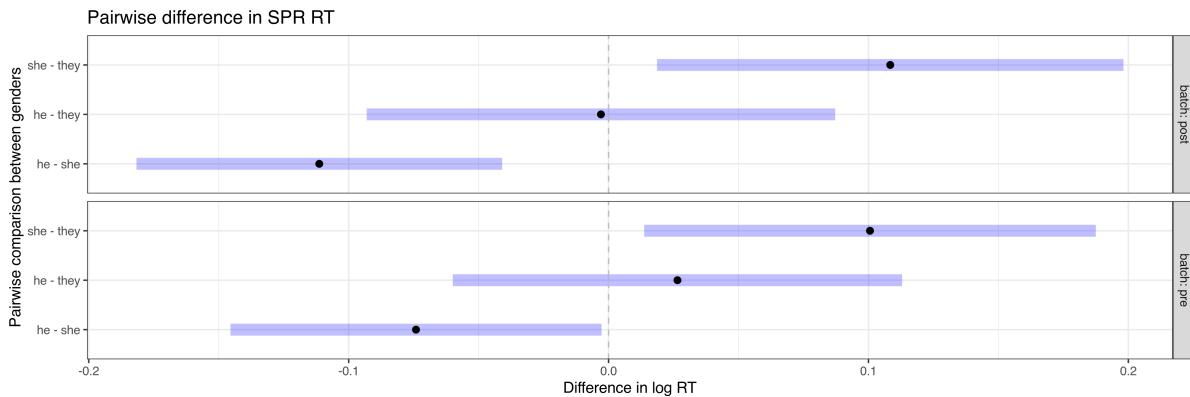
batch = post:

	pronoun_gender_pairwise	estimate	SE	df	t.ratio	p.value
he - she		-0.11131	0.0358	836	-3.106	0.0020
he - they		-0.00292	0.0459	836	-0.064	0.9492
she - they		0.10839	0.0457	836	2.371	0.0180

batch = pre:

	pronoun_gender_pairwise	estimate	SE	df	t.ratio	p.value
he - she		-0.07409	0.0364	836	-2.037	0.0420
he - they		0.02651	0.0440	836	0.602	0.5474
she - they		0.10060	0.0443	836	2.273	0.0233

Results are given on the log (not the response) scale.



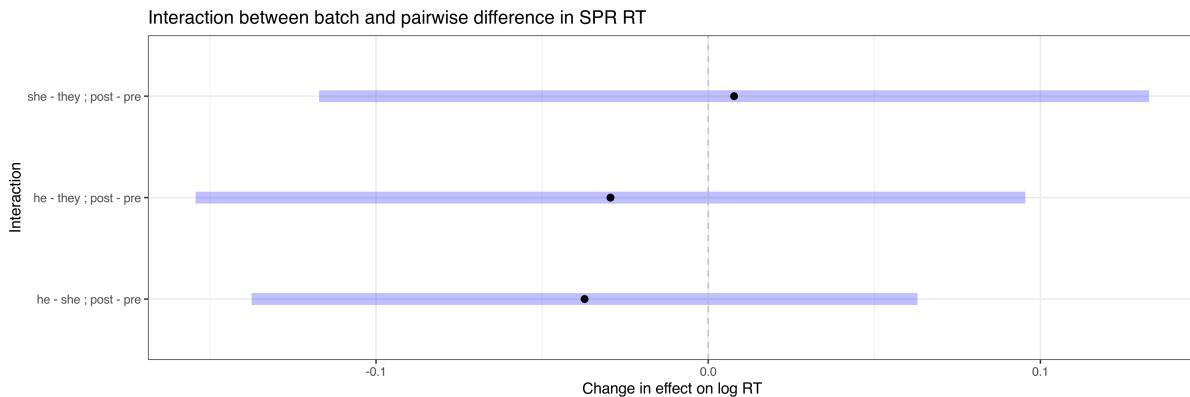
^ Pairwise comparisons between genders show, for SPR RTs:

- in pre-election data, RT of female pronouns is slightly larger than that of neutral pronouns, at 0.05 level (estimate=0.0312, p=0.0238). No other pairwise comparisons are significant.
- in post-election data, RT of female pronouns is larger than that of either neutral or male pronouns, respectively, at 0.05 level. Comparison between male and neutral pronouns is not significant.

Next, looking at interaction between each pairwise contrast, pre-vs-post election:

pronoun_gender_pairwise	batch_pairwise	estimate	SE	df	t.ratio	p.value
he - she	post - pre	-0.03722	0.0511	836	-0.729	0.4663
he - they	post - pre	-0.02944	0.0636	836	-0.463	0.6438
she - they	post - pre	0.00779	0.0636	836	0.122	0.9027

Results are given on the log (not the response) scale.



^ interaction between each pairwise contrast among pronouns and experiment batch (pre- vs post-election).

- Comparing post-election minus pre-election
  - log RT difference he-she is negative (estimate = -0.0792, p < .0001) — that is, the bias for reading male faster than female pronouns increased.
  - other contrasts are not significant.

## Other exploratory plotting

Number of participants per task pre- and post- election.

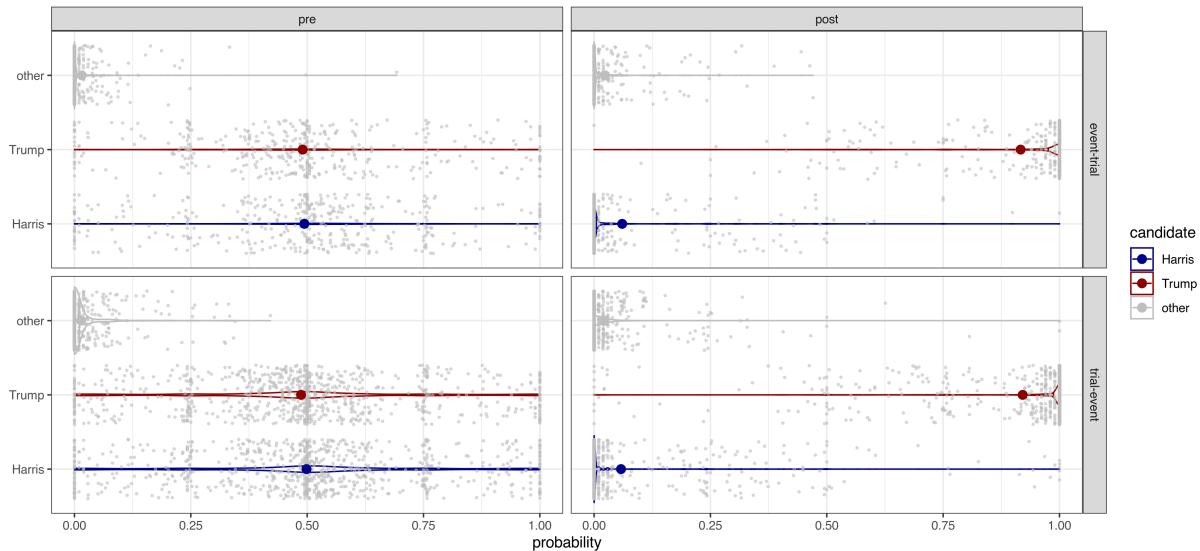
batch	cloze	maze	spr	total
pre	485	244	445	1174
post	497	235	432	1164

## Expectations

Check for order effect on belief estimation results, in pre- and post- election data.

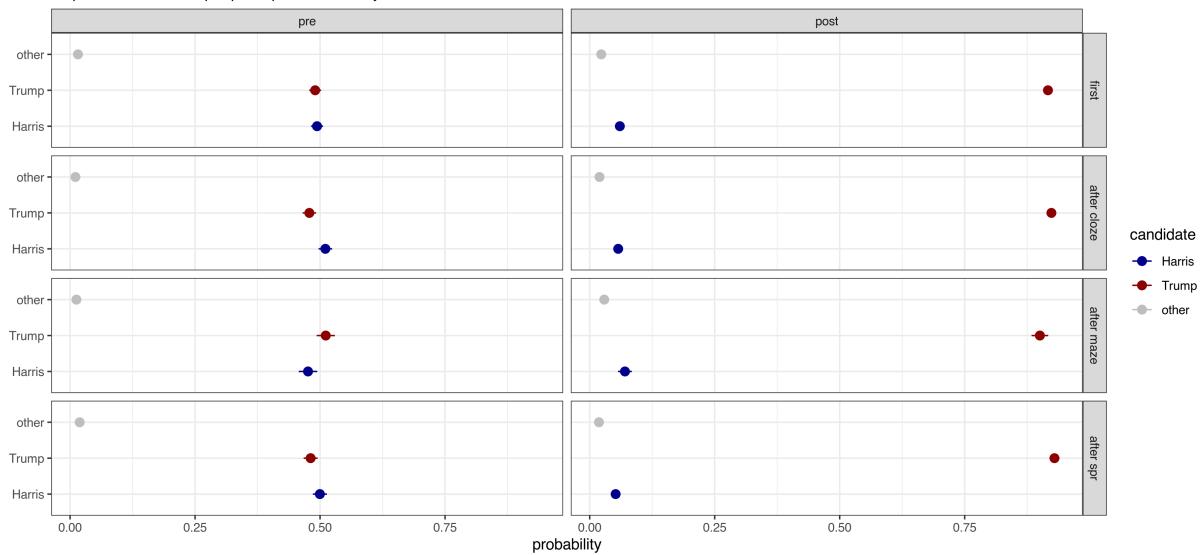
"Who do you think will be the US president in February 2025?"

responses normalized per participant

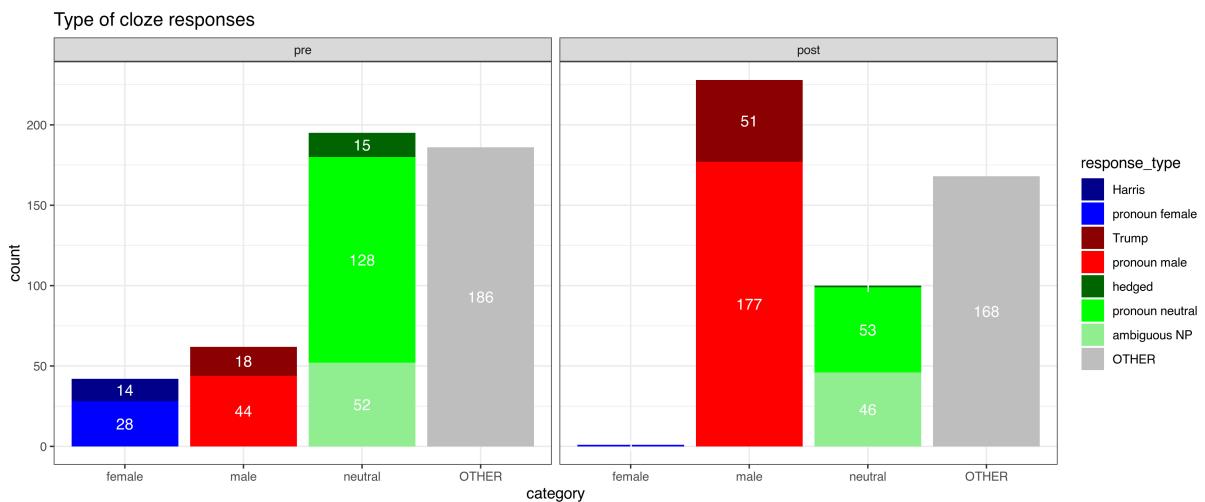


Looking for order effects, (combining all orders that start with “event” to one, for this plot, since all are identical wrt this data)

"Who do you think will be the US president in February 2025?"  
responses normalized per participant, faceted by order

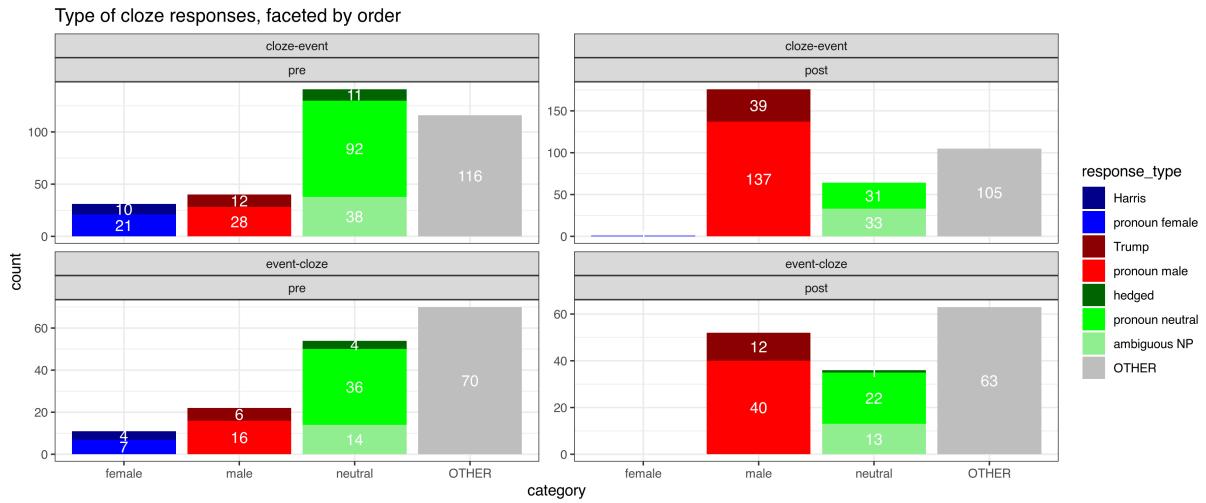


## Cloze



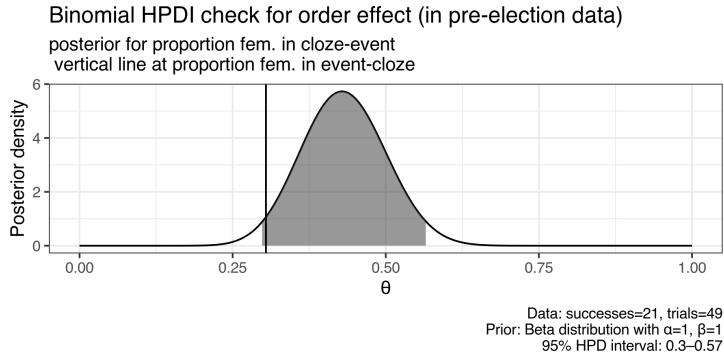
```
prop.test(F, M+F) p-value  
0.07709987
```

Faceting by order:



In the event-cloze order there seems to be a somewhat larger bias toward producing male pronouns or male referents/Trump.

*But it seems the difference between orders here is actually not significant, per Titus' binomial 95%-HPDI test:*



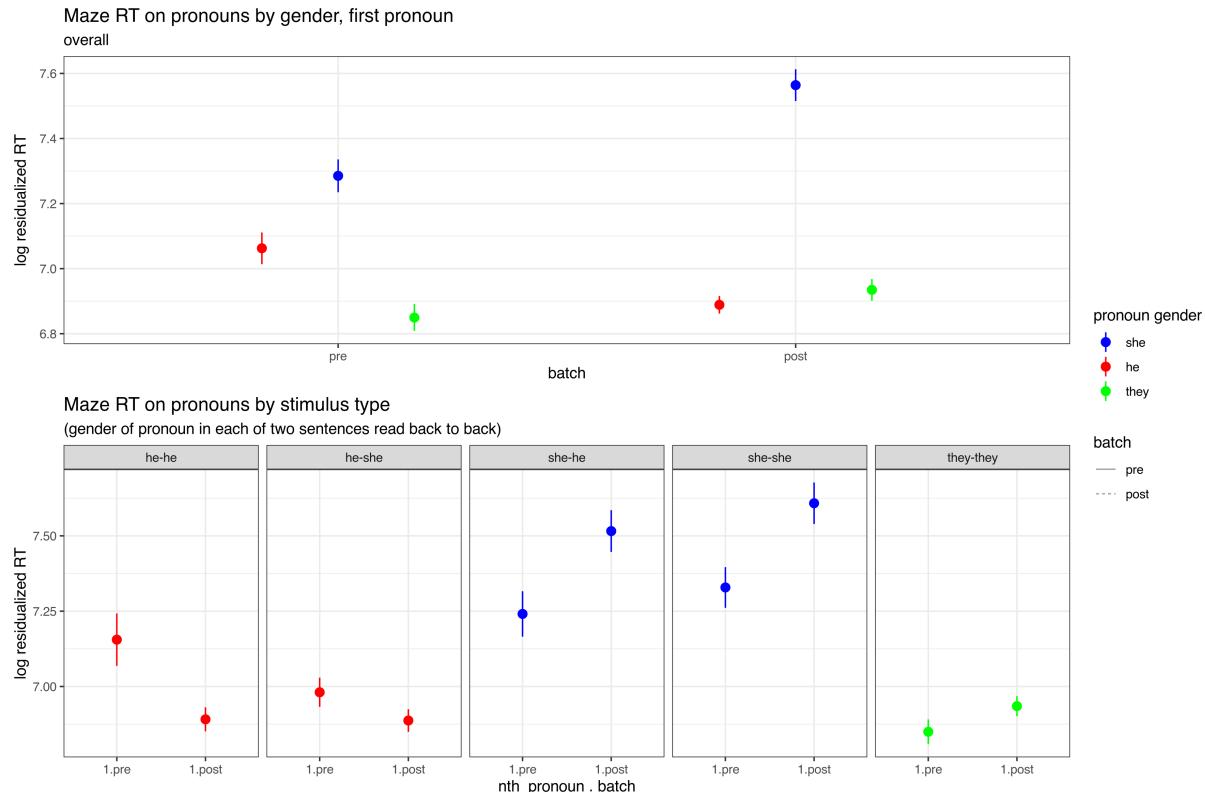
## By item cloze breakdown

Type of cloze responses, by item  
(OTHER responses not shown)

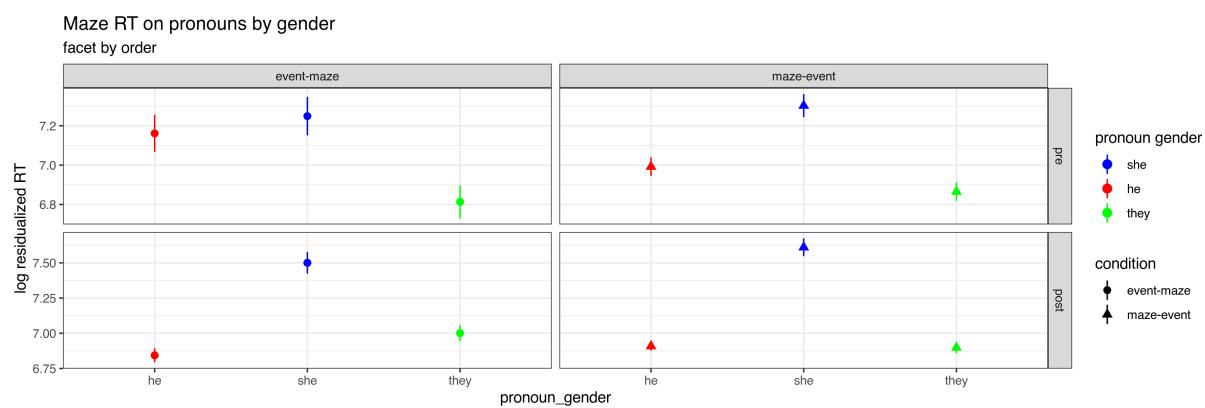


## Maze

Look at average Maze RT by gender of pronoun, and for each condition of two sentences presented. Shows an effect of gender (male pronoun is read faster than female, neutral is perhaps even faster than male). These Maze RTs are residualized as with the SPR (control for participant mean reading speed and eg word length, punctuation, and by item effects).

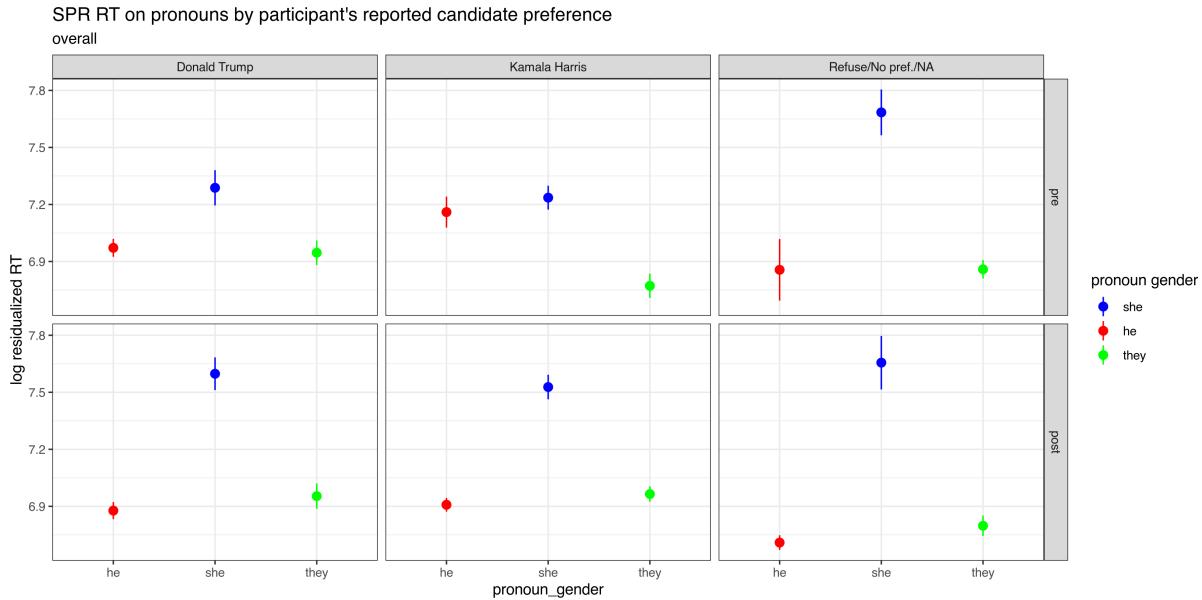


## Faceting by order



^ In the pre-election data (top facets above), there may be a slightly less clear male pronoun bias for participants who did the maze task following the event estimation (event-maze) versus before (maze-event). In the post-election data, the bias was strong in either order.

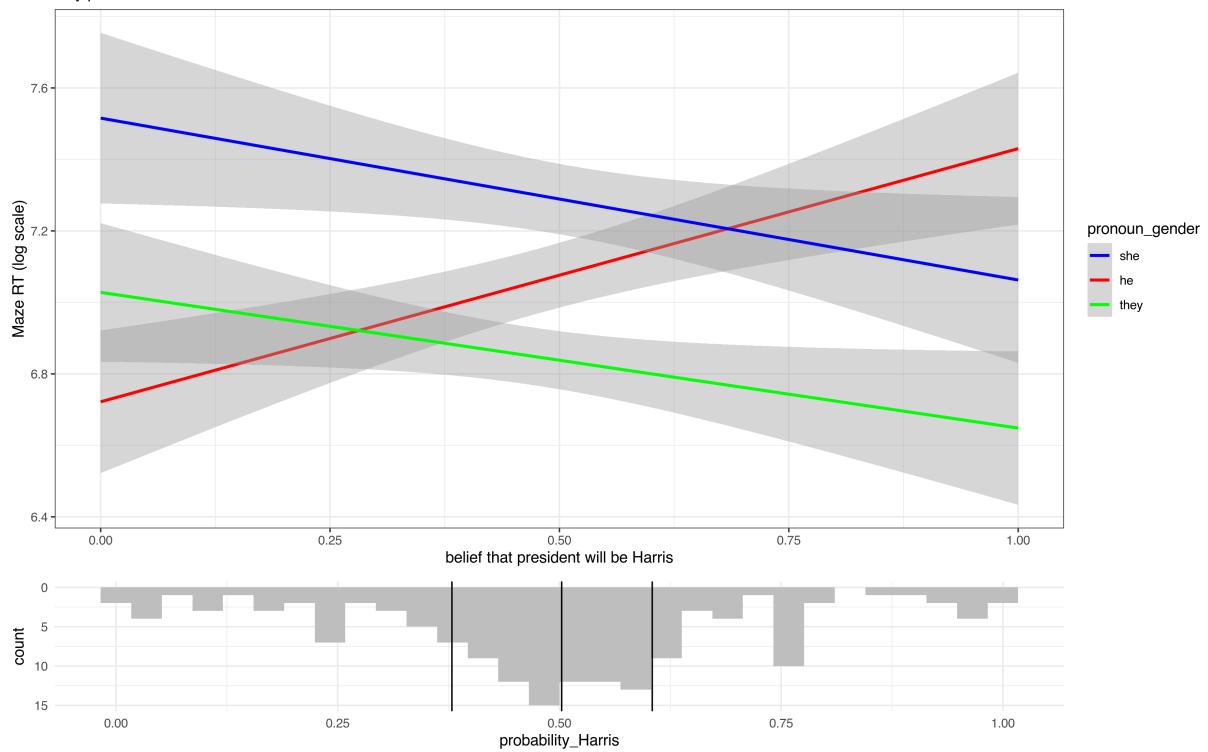
And broken down by participant reported preference



^ The pre-election bias is least strong among participants whose preferred candidate was Harris.

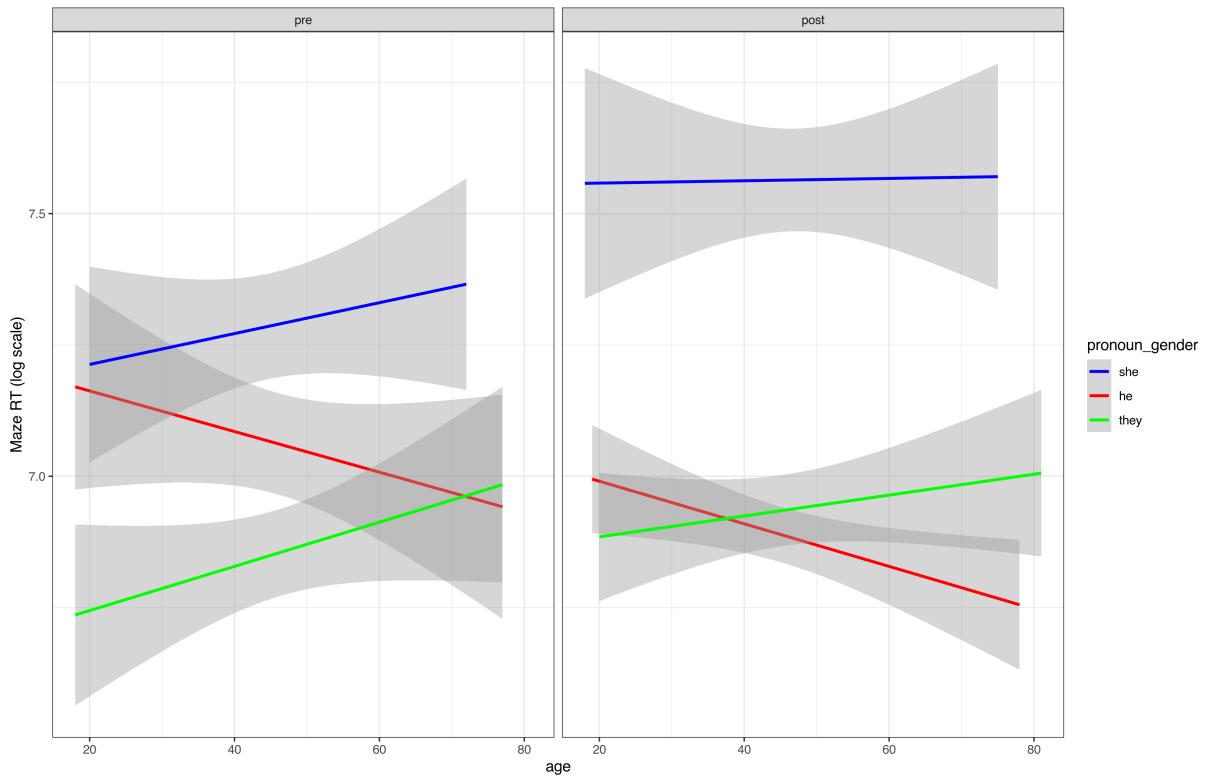
We could look at a related effect as a continuous interaction plotting event-probability (who the participant thinks will be president) on the x-axis versus RT on the y-axis:

Maze RT on pronouns, plotted against probability of female president  
only pre-election data



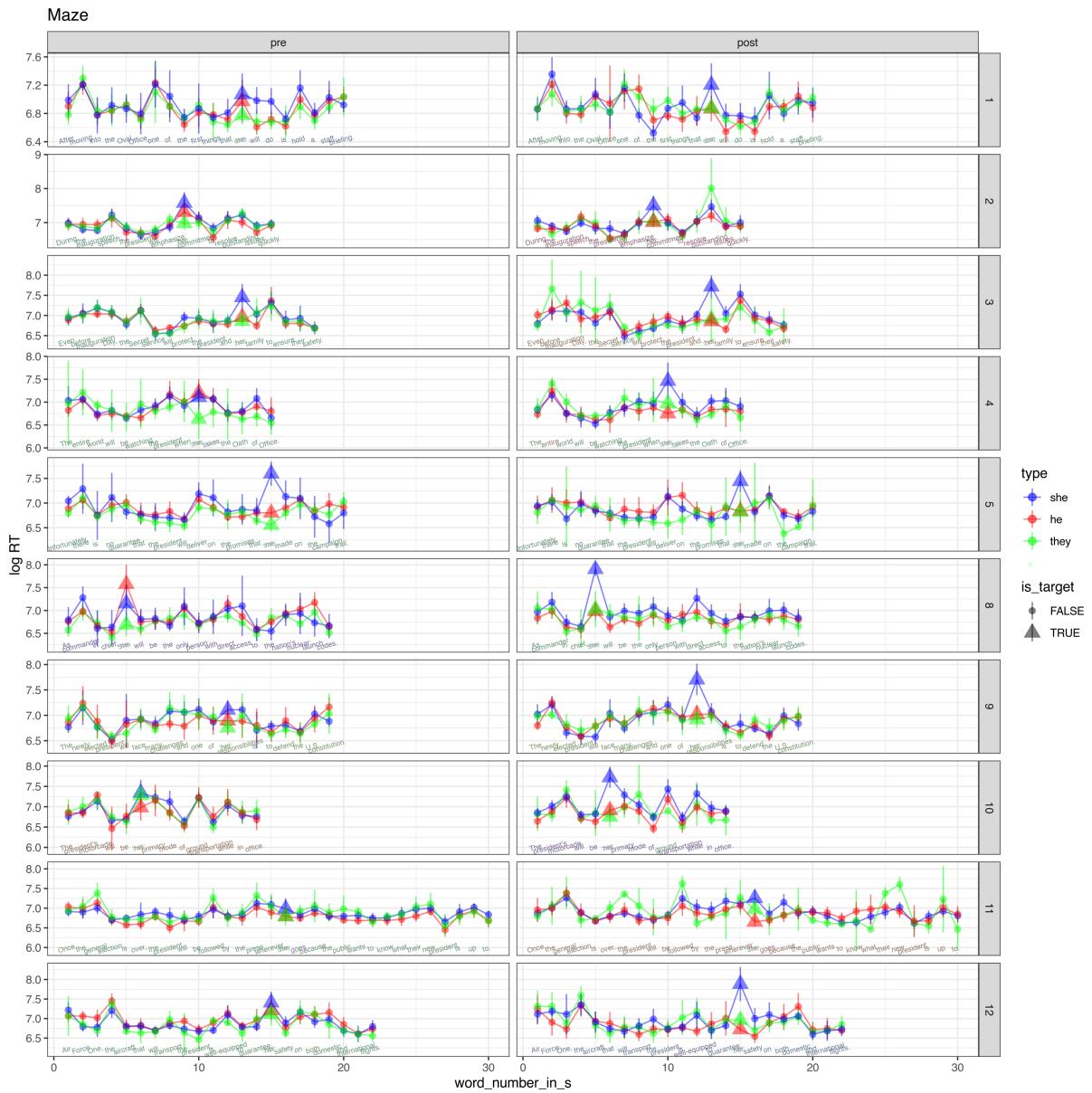
^ It looks like the male-pronoun processing bias (/ female-pronoun processing penalty) gets smaller as belief that the president would be female is higher. But note the majority of participants are in the middle of the distribution.

Maze RT on pronouns, plotted against participant age



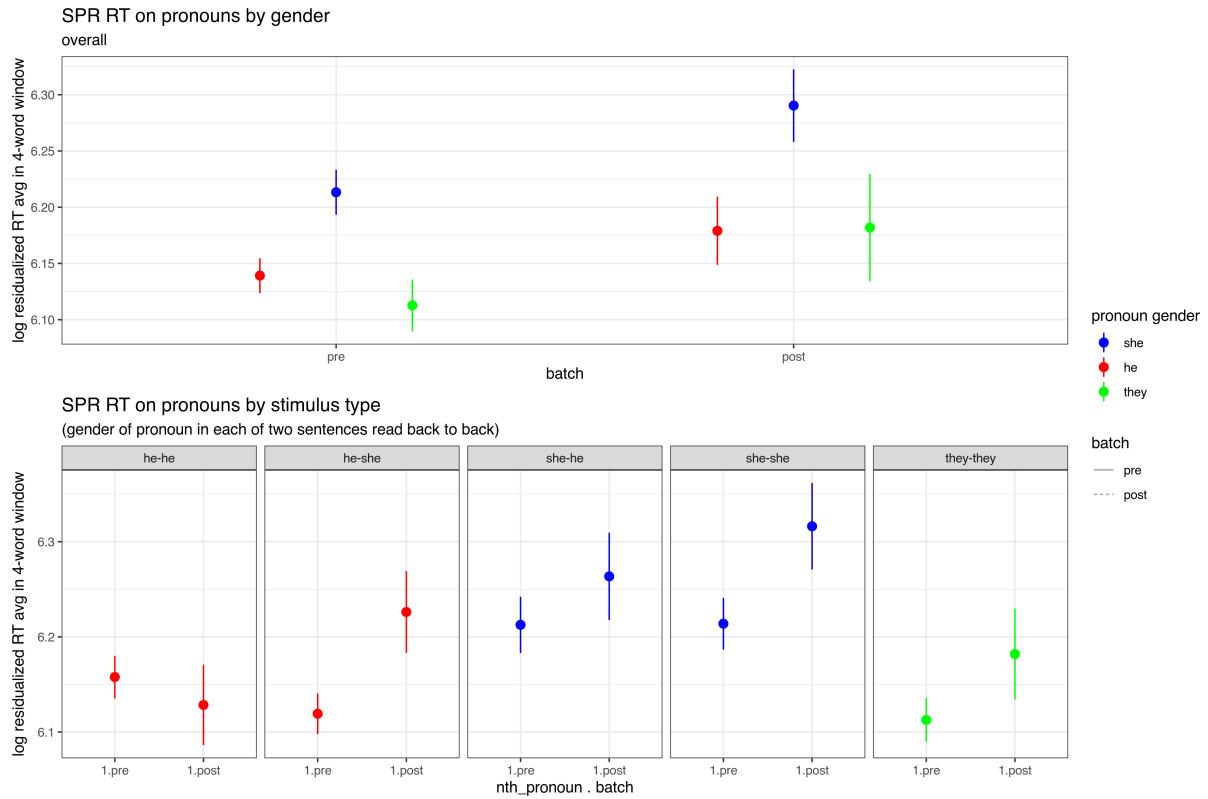
^Looking at participant age, there also could be something interesting going on, at least with the decreasing processing cost of the neutral pronoun for younger readers.

## Maze reading time tracks on each of the target sentences:

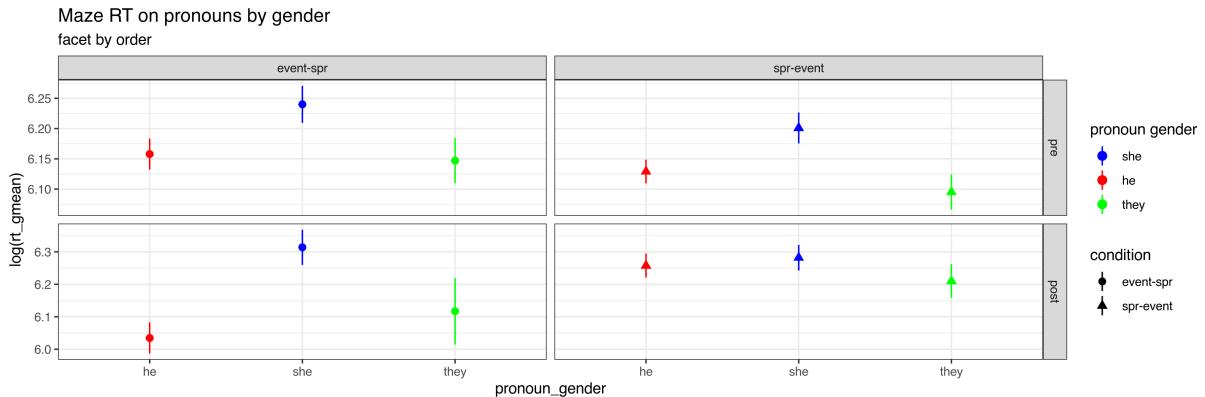


## SPR

Likewise look at average SPR RT by gender of pronoun, and for each condition of two sentences presented. These RTs are residualized.



### Faceting by order

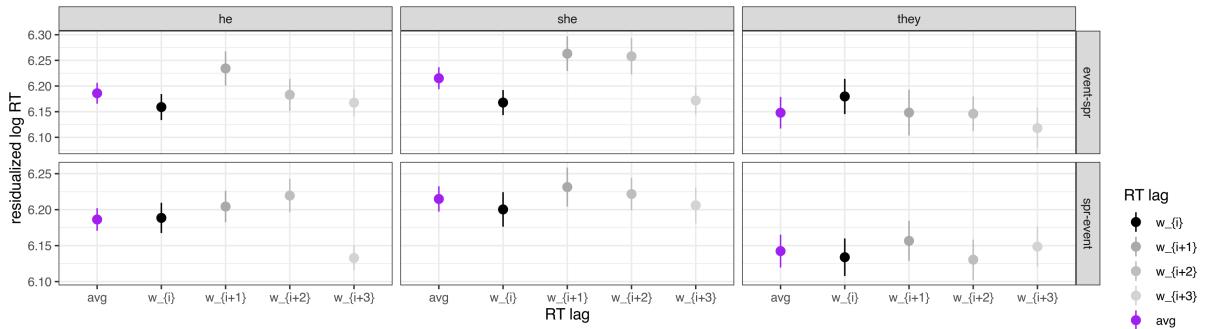


^ Probably not enough power to detect, but it may be true that the bias is stronger in the spr-event order, as with Maze data.

### Pre-election data

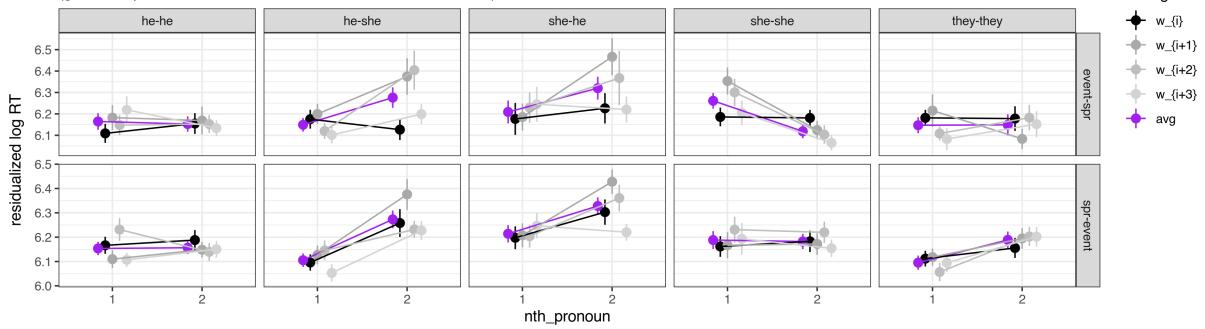
#### SPR by gender

Looking at RT with different lags  $w_{-i+n}$ , where  $w_{-i}$  = target pronoun.  
Also showing geometric\_mean( $w_{-i+n}$ ) for  $n = \{0, \dots, 3\}$



#### Maze RT on pronouns by stimulus type

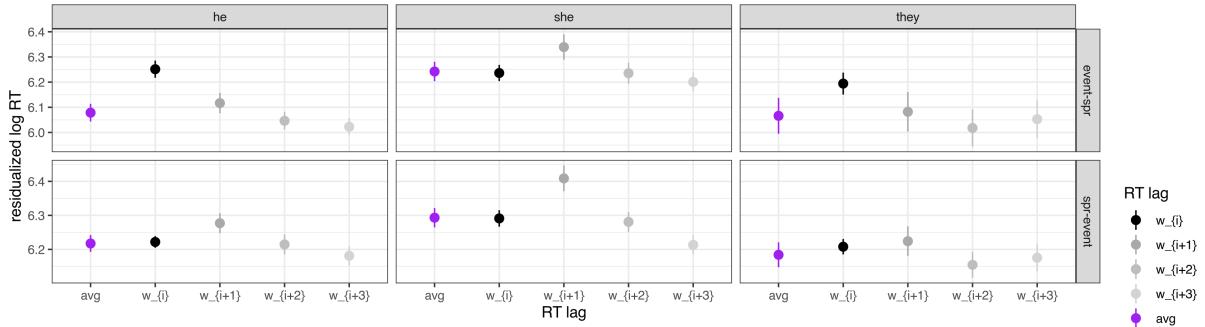
(gender of pronoun in each of two sentences read back to back)



### Post-election data

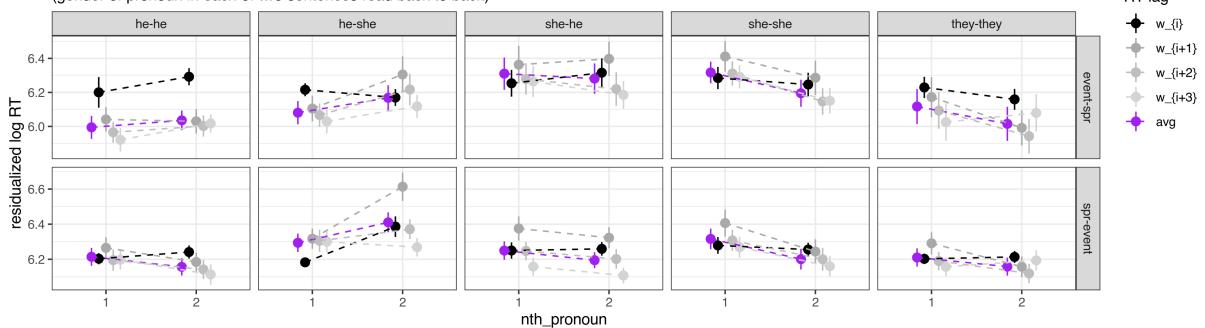
#### SPR by gender

Looking at RT with different lags  $w_{\{i+n\}}$ , where  $w_i$  = target pronoun.  
Also showing geometric\_mean( $w_{\{i+n\}}$ ) for  $n = \{0, \dots, 3\}$

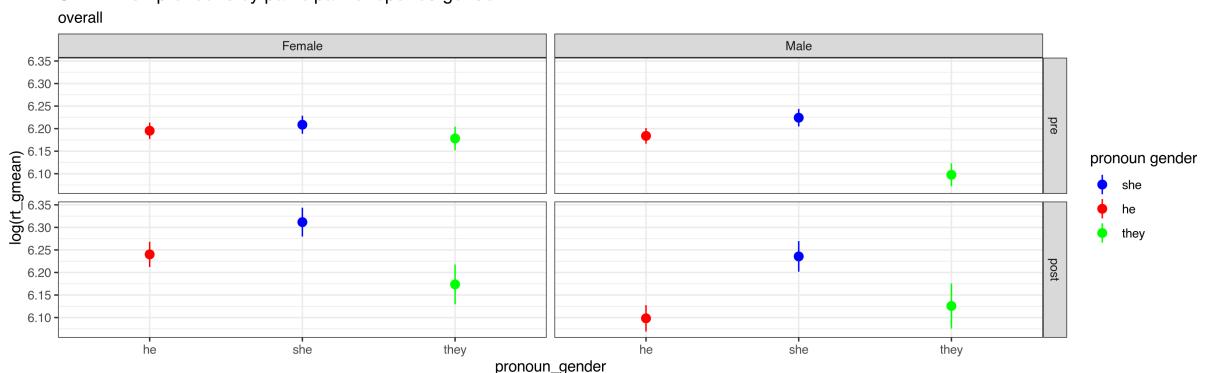


#### Maze RT on pronouns by stimulus type

(gender of pronoun in each of two sentences read back to back)

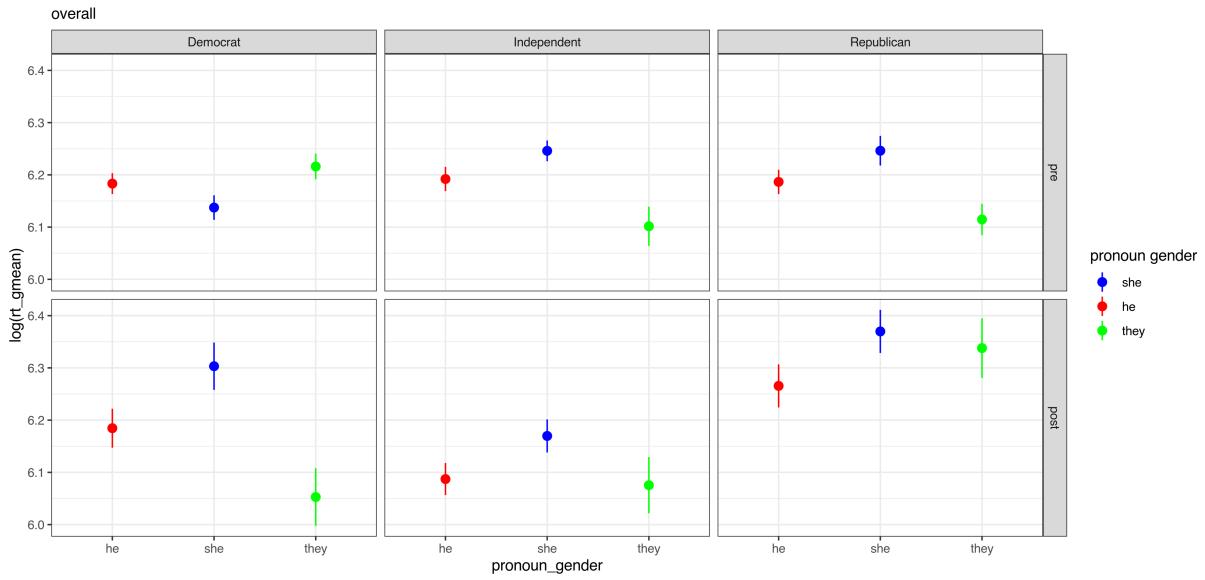


#### SPR RT on pronouns by participant's reported gender overall



^ SPR averages broken down by participant gender may be useful if we're interested in the difference that shows up in Maze as well: bias seems larger in males.

SPR RT on pronouns by participant's reported political affiliation.



^ SPR averages broken down by participant political affiliation, as with Maze above.

## SPR tracks on each of the target sentences:

