

Errata

On Blair, Coppock and Moor (2020), *American Political Science Review* 114(4)

This note lists eleven corrections to “When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments,” published in the *American Political Science Review* 114(4), doi [10.1017/S0003055420000374](https://doi.org/10.1017/S0003055420000374): ten misstated numbers and one sentence that describes the estimates incorrectly as a result. **None of them changes a conclusion of the paper.** Every one is in the results text, restating a quantity the article also prints in Table 4, and in each case Table 4 is right and the sentence is not. The sentences that carry the paper’s findings, and Table 4 itself, are unaffected.

Every number in a corrected sentence below is computed from the deposited replication archive ([10.7910/DVN/YUXHZT](https://doi.org/10.7910/DVN/YUXHZT)) at the time this document is rendered, including the numbers that were not wrong. The code and output supporting each correction are at github.com/active-maintenance-aec/blair_coppock_moor_2020.

Every quantity on the percentage-point scale is stated to one decimal place, except where the article states it as an approximation, and bold marks the corrections. A corrected sentence therefore differs from the published one in more places than it is wrong: the article states most of these quantities in whole points, and setting a whole number to one decimal is not a correction. Only what is bold has changed. A figure the article hedges as approximate keeps the article’s own precision, because the approximation is the claim and such a figure is correct at the precision it states. Where the fuller value matters it is given in the note beneath the entry.

The first six entries are substantive: the number a reader would quote is wrong by more than a point, and the first two turn on a sign. The last five restate Table 4 at whole-point precision and differ from it by less than a point, which is exactly why they need a decimal to be visible at all.

1. The sexual orientation underreporting estimate carries the wrong sign

For questions thought to be subject to overreporting, the average estimate of sensitivity bias is -2 percentage points; for underreporting, the estimate is **$+2$** percentage points.

Corrected:

For questions thought to be subject to overreporting, the average estimate of sensitivity bias is -1.8 percentage points; for underreporting, the estimate is -1.0 percentage points.

Table 4 prints -0.01 for the same quantity. The paper's conclusion for this literature, that sensitivity bias in questions about attitudes towards gays and lesbians cannot be distinguished from zero, is unaffected.

2. Only one of the two sexual orientation estimates has an unexpected sign

These estimates **both have an unexpected sign** relative to the predictions and are not distinguishable from zero.

Corrected:

Only the overreporting estimate has an unexpected sign relative to the predictions, and neither is distinguishable from zero.

The sentence follows directly from the misstated figure corrected above. Underreporting predicts a negative difference between the direct question and the list experiment, and the underreporting average is -1.0 points, which is the sign predicted. Overreporting predicts a positive difference, and the overreporting average is -1.8 points, which is not. Both credible intervals contain zero, so the second half of the sentence stands as published.

3. The racial prejudice 95 per cent predictive interval

The 50% predictive interval reaches from $+1$ to $+7$ points, but the 95% predictive interval admits large negative biases (-16 points) to large positive biases up to **16** points.

Corrected:

The 50% predictive interval reaches from $+0.7$ to $+6.9$ points, but the 95% predictive interval admits large negative biases (-5.2 points) to large positive biases up to **12.8** points.

Table 4 prints $[-0.05, 0.12]$ for this interval. The lower bound in the text is roughly three times the one the table carries.

4. The width of the turnout 95 per cent predictive interval

This uncertainty is also reflected in the very wide predictive intervals: the 50% interval is 13 points wide and the 95% interval is **45** points wide.

Corrected:

This uncertainty is also reflected in the very wide predictive intervals: the 50% interval is 12.9 points wide and the 95% interval is **37.4** points wide.

Table 4's own interval for turnout is $[-0.12, 0.26]$, which is 38 points wide, so the sentence disagrees with the table as well as with the data.

5. The vote buying predictive intervals

In the bottom row of Figure 4a, we show the 50% predictive interval from -14 to -2 points and the 95% predictive interval from -27 to $+10$ points.

Corrected:

In the bottom row of Figure 4a, we show the 50% predictive interval from -14.0 to -2.7 points and the 95% predictive interval from -24.9 to $+8.1$ points.

Table 4 prints $[-0.14, -0.03]$ and $[-0.25, 0.08]$ for the same two intervals.

6. The width of the religious prejudice credible interval

This estimate is in the expected direction, but the credible interval is **13** points wide and includes zero.

Corrected:

This estimate is in the expected direction, but the credible interval is **11.4** points wide and includes zero.

Table 4's own interval for religious prejudice is $[-0.07, 0.04]$, which is 11 points wide.

7. The upper bound of the vote buying credible interval

In the top row is the estimated average sensitivity bias across studies for vote-buying questions (black diamond), -8 points with a 95% credible interval stretching from -13 to -3 points.

Corrected:

In the top row is the estimated average sensitivity bias across studies for vote-buying questions (black diamond), -8.4 points with a 95% credible interval stretching from -13.3 to -3.7 points.

Table 4 prints -0.04 for the same bound.

8. The average bias across all underreporting studies

Second, as shown in Table 4, the average sensitivity bias in the case of underreporting is -3 points (SE: 1 point).

Corrected:

Second, as shown in Table 4, the average sensitivity bias in the case of underreporting is -3.5 points (SE: 1.0 point).

Table 4 prints -0.04 , which rounds to -4 rather than -3 at whole-point precision.

9. The lower bound of the racial prejudice credible interval

For questions about racial prejudice, our summary guess is that if anything, subjects overreport racist attitudes by approximately 4 points (95% CI -3 to $+9$).

Corrected:

For questions about racial prejudice, our summary guess is that if anything, subjects overreport racist attitudes by approximately 4 points (95% CI -2.4 to $+9.5$).

Table 4 prints -0.02 for the same bound, which is -2 rather than -3 at whole-point precision. The estimate the sentence hedges as “approximately 4 points” is 3.8 and is correct at the precision it states, so it is left as published.

10. The lower bound of the authoritarian overreporting predictive interval

Based on 21 studies, our meta-analysis estimate of the average level for studies in which overreporting is predicted is +14 points and the 50% predictive interval suggests a likely level of sensitivity bias between **+4** to +24 points.

Corrected:

Based on 21 studies, our meta-analysis estimate of the average level for studies in which overreporting is predicted is +14.2 points and the 50% predictive interval suggests a likely level of sensitivity bias between **+4.5** to +23.9 points.

Table 4 prints 0.05 for the same bound.

11. The sample size at which a list experiment becomes preferable on RMSE grounds

For a study with 1,000 subjects, the bias must be greater than **5.5** points to prefer a list experiment; at 2,000, the bias must be greater than 4 points.

Corrected:

For a study with 1,000 subjects, the bias must be greater than **5.6** points to prefer a list experiment; at 2,000, the bias must be greater than 3.9 points.

The RMSE indifference curve the article's own Figure 2a plots crosses at 5.57 points at this sample size, so a bias of 5.5 points would not in fact prefer the list experiment. The companion bound at 2,000 subjects crosses at 3.94 and holds as stated.