Errata

On Coppock and McClellan (2019), *Research and Politics* 6(1)

This note lists six corrections to "Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents," published in *Research and Politics* 6(1), doi 10.1177/2053168018822174. **None of them changes a conclusion of the paper.** Two are counts of statistically significant differences that follow from the hypothesis test the appendix says it used rather than the one its code performs, one reverses the direction of a descriptive comparison, one is a count off by one, and two are sample sizes in the appendix study manifest. The paper's finding, that Lucid respondents track US national benchmarks well and can serve as a drop-in replacement for Mechanical Turk, is unaffected, as is its one negative result on the healthcare rumor corrections.

Every number in a corrected sentence below is computed from the deposited replication archive (10.7910/DVN/DDWWJW) at the time this document is rendered, including the numbers that were not wrong. Bold marks the corrections. The code and output supporting each correction are at github.com/active-maintenance-aec/coppock_mcclellan_2019.

The first two entries share a cause. Section 1 of the online appendix states that "we obtain two-tailed p-values under a normal approximation," and the deposited `12_DistanceTests.R` instead computes `pnorm(2 * (1 - abs(z)))`, which is not a p-value for any hypothesis. The two expressions differ in where they put the threshold: the deposit's calls a difference significant at the 0.05 level once `|z|` exceeds 1.82, where the two-tailed test needs 1.96. Of the 21 characteristics in appendix Table 1, extraversion is the only one whose verdict falls between the two, and the corrections below are what follows for the two sentences that count significant differences. The standard errors themselves come from an unseeded bootstrap and are not exactly reproducible, so both corrections are recomputed from the distance and the standard error appendix Table 1 prints, which do not move; the maintained rewrite's own run agrees with each.

1. Lucid is not significantly closer than MTurk on all five personality traits

> Formal hypothesis tests demonstrate that Lucid is significantly closer to the ANES 2012 than MTurk on **all five traits**.

Corrected:

> Formal hypothesis tests demonstrate that Lucid is significantly closer to the ANES 2012 than MTurk on **four of the five traits, extraversion excepted**.

Appendix Table 1 prints 0.100 and 0.052 for extraversion's distance and bootstrapped standard error. Those two numbers give `z` = 1.92 and a two-tailed p-value of 0.054, above the 0.05 threshold; the table's own p-value column prints 0.034, which is the deposit's expression rather than the two-tailed test. The maintained rewrite's seeded run reaches the same conclusion, at `z` = 1.89 and p = 0.059. The other four traits clear 0.05 by a wide margin under either test.

2. The number of significant differences in the online appendix

> Out of 21 opportunities, Lucid is closer to the ANES 18 times, **14** of which are significant.

Corrected:

> Out of 21 opportunities, Lucid is closer to the ANES 18 times, **13** of which are significant.

Extraversion is the difference between the two counts, for the reason given above. Recomputing every p-value from the distance and standard error appendix Table 1 prints gives the same 13. The sentence that follows it is unaffected: MTurk is closer in 3 instances and significant in one, voter turnout, under both tests.

3. The MTurk sample under-represents southerners

> The regional balance on Lucid comes very close to the 2012 ANES, whereas the MTurk sample appears to **overrepresent** southerners.

Corrected:

> The regional balance on Lucid comes very close to the 2012 ANES, whereas the MTurk sample appears to **underrepresent** southerners.

The southern share is 30.1 per cent on MTurk against the ANES 2012 benchmark of 37.2, so MTurk sits 7.1 points below it; Lucid is at 37.6. Figure 1, which is where the sentence sends the reader, plots the same thing in standard deviations: MTurk at –0.146 and Lucid at 0.009, with the benchmark at zero. MTurk over-represents the Northeast and the Midwest instead. Appendix Table 1 is unaffected: it reports the distance between the two samples rather than the direction of either, and its South row is right.

4. Six demographic differences are significant, not five

> Out of the 11 demographic variables that are measured for both Lucid and MTurk, the Lucid mean is closer to the ANES mean in nine instances, **five** of which are statistically significant.

Corrected:

> Out of the 11 demographic variables that are measured for both Lucid and MTurk, the Lucid mean is closer to the ANES mean in nine instances, **six** of which are statistically significant.

Appendix Table 1 prints a p-value below 0.05 for Female, Education, Age, Mean income, Black and South, and Lucid is the closer sample on all six. The count is the same under the two-tailed test corrected in the first two entries, so it does not turn on which p-value is used.

5. The sample size of the Lucid welfare replication

> Sample: **1,811** Lucid respondents. [Online appendix, section 2.1.3]

Corrected:

> Sample: **3,294** Lucid respondents.

1,811 is the Lucid sample of the Hiscox free-trade replication, reported in section 2.4.2 of the same appendix. The welfare data hold 3,504 Lucid rows, of which 3,294 have a usable outcome, and the manifest reports the analysed sample in every other section.

6. The sample size of the original Kam and Simas study

> Sample: **761** respondents gathered from a Knowledge Networks national probability sample. [Online appendix, section 2.3.1]

Corrected:

> Sample: **752** respondents gathered from a Knowledge Networks national probability sample.

Appendix Table 2 reports N = 752 for all three models fitted to this sample, and 752 is what the deposited data give. The manifest reports the analysed sample rather than the source study's throughout: its entry for the 1984 General Social Survey gives 943, where the deposited file holds 1,473 respondents for that survey.