

Errata for Peyton, Huber and Coppock (2022)

The Generalizability of Online Experiments Conducted During the COVID-19 Pandemic

Eleven corrections to the article and its online appendix, none of which changes a conclusion. Every corrected value below is computed at render time from the maintained reanalysis of the deposited replication archive, 10.7910/DVN/38UTBF. The article is 10.1017/XPS.2021.17; a corrigendum published in 2023, 10.1017/XPS.2022.23, corrects the significance counts in the article's Figures 2 and 3 and the first paragraph of page 6, and nothing below repeats a correction that corrigendum already made. The reanalysis is at active-maintenance-aec/peyton_huber_coppock_2022.

Values on the percentage-point scale are stated to one decimal place throughout, including numbers that were not wrong. Bold marks the tokens that changed.

1. Online appendix, Section A.8, page 18: the four estimates attributed to the original study are the pre-COVID replication's

Published:

In the original study, the 90/70 condition caused an increase in the proportion of subject that preferred the nuclear option by about **37** percentage points relative to the 90/90 condition. The estimated effect of the 90/45 condition, relative to the 90/90 condition, was **51** percentage points. Similarly, the 90/70 condition caused an increase in the proportion of subjects that approved of the nuclear option by about **17** percentage points, and the 90/45 condition caused an increase of about **27** percentage points.

Corrected:

In the original study, the 90/70 condition caused an increase in the proportion of subject that preferred the nuclear option by about **29.7** percentage points relative to the 90/90 condition. The estimated effect of the 90/45 condition, relative to the 90/90 condition, was **47.6** percentage points. Similarly, the 90/70 condition caused an increase in the proportion of subjects that approved of the nuclear option by about **5.1** percentage points, and the 90/45 condition caused an increase of about **27.5** percentage points.

All four published figures are the estimates from the Aronow, Baron and Pinson (2019) pre-COVID replication rather than from Press, Sagan and Valentino (2013). The appendix's own Table 2 prints both columns and gives the original study 0.30, 0.48, 0.05 and 0.28 against the replication's 0.37, 0.51, 0.17 and 0.27.

2. Online appendix, Section A.10, footnote 5, page 25: the count of significant differences

Published:

Although all 6 replication estimates are smaller in magnitude than those in the original study, only **two** of these differences (**John Podesta and Trump** stories) are statistically distinguishable from zero.

Corrected:

Although all 6 replication estimates are smaller in magnitude than those in the original study, only **four** of these differences (**John Podesta, Trump, Scaramucci and Vermont** stories) are statistically distinguishable from zero.

The two additional differences are significant at conventional levels before any adjustment for multiple comparisons, which is what the footnote describes; three of the four survive a Benjamini-Hochberg adjustment across the six stories. The 2023 corrigendum corrected the same class of error in the article's Figures 2 and 3 and did not reach this footnote.

3. Online appendix, Section A.1, page 1: the replication estimate and its distance from the benchmark

Published:

Pooling across the pre-COVID studies, the summary effect size estimate is a 30 percentage point increase in support for Russian journalists. Our effect estimate of **25.5** points is 83% of this magnitude, and this difference of **5.5** percentage points is not statistically significant ($P = 0.16$).

Corrected:

Pooling across the pre-COVID studies, the summary effect size estimate is a 30.5 percentage point increase in support for Russian journalists. Our effect estimate of **25.4** points is 83% of this magnitude, and this difference of **5.1** percentage points is not statistically significant ($P = 0.16$).

The sentence's own three figures do not agree with each other: 30 less 25.5 is 4.5, not 5.5. The ratio and the p-value it also prints are both right, and both follow from the corrected estimate and difference.

4. Page 6: the non-conjoint correspondence

Published:

The 24 of 28 correctly signed estimates from the other replications were, on average, **49%** as large as the pre-COVID summary effect sizes.

Corrected:

The 24 of 28 correctly signed estimates from the other replications were, on average, **48%** as large as the pre-COVID summary effect sizes.

The mean ratio over those 24 pairs is 0.483. The overall 73% and the conjoint 87% stated in the same paragraph both reproduce exactly, and the three are consistent with one another only at 48 per cent.

5. Page 10: the average duration among browser respondents

Published:

Those from web applications spent approximately 7 min less time completing surveys than those from web browsers, who spent an average of **21.5** min.

Corrected:

Those from web applications spent approximately 7.2 min less time completing surveys than those from web browsers, who spent an average of **21.0** min.

21.5 minutes is the average among nonmobile respondents, which the next sentence uses.

6. Page 11: the inattentive share

Published:

If 73% of respondents were attentive and **26%** inattentive, our replication estimate would, on average, match our pre-COVID benchmarks.

Corrected:

If 73% of respondents were attentive and **27%** inattentive, our replication estimate would, on average, match our pre-COVID benchmarks.

The correspondence is 73.0 per cent, so the two shares are the attentive figure and its complement.

7. Online appendix, Section A.9, page 22: the number of estimated AMCEs

Published:

In total, there are **81** estimated AMCEs in Figure A.10 – 41 for the original study and 41 for the replication.

Corrected:

In total, there are **82** estimated AMCEs in Figure A.10 – 41 for the original study and 41 for the replication.

The two component counts are right and Figure A.10 plots 82 points.

8. Online appendix, Section A.8, page 17: the 90/70 condition's success rate

Published:

2) a “90/70” condition in which the conventional strike had a **75%** chance of success

Corrected:

2) a “90/70” condition in which the conventional strike had a **70%** chance of success

The condition's two siblings are described at the rates their labels name, a 90% chance for the 90/90 condition and 45% for the 90/45 condition, and the label is written 90/70 in the same paragraph, in the sentence that follows it and on Figure A.8.

9. Online appendix, Section A.5, page 11: a reference to a table that does not exist

Published:

Table A.5 provides a summary of differences between the original study and the replications by treatment arm.

Corrected:

Table 1 provides a summary of differences between the original study and the replications by treatment arm.

The appendix carries no Table A.5. The table this sentence points at is captioned Table 1 and appears on the following page; the deposited file that builds it is named `appendix_table_a5.tex`, so it is the printed caption rather than the reference that departed from the intended numbering.

10. Page 6: whether every replication estimate was smaller

Published:

All replication summary estimates were smaller in magnitude than their pre-COVID estimates, with 24 of 28 signed in the same direction.

Corrected:

All but one of the replication summary estimates were smaller in magnitude than their pre-COVID estimates, with 24 of 28 signed in the same direction.

The exception is the support-for-redistribution pair, whose pre-COVID summary effect is 0.00008 and whose COVID-era estimate is -0.044. It is one of the four pairs signed against their benchmark, so the count in the second half of the sentence is unaffected.

11. Online appendix, Section A.7, page 15: the standard error of the difference between the two intentionality arms

Published:

The estimated treatment effect for the COVID-specific replication is about 0.01 points smaller than the direct replication (SE = **0.04**, $P = 0.46$).

Corrected:

The estimated treatment effect for the COVID-specific replication is about 0.01 points smaller than the direct replication (SE = **0.06**, $P = 0.46$).

0.04 is the standard error of the difference against the pre-COVID benchmark, given in the sentence that follows. The published $P = 0.46$ is the one-tailed p-value at the corrected standard error, so the sentence is internally consistent once the standard error is corrected.