Additional file 6

The adaptive community-response (ACR) method for collecting misinformation on social media

Julian Kauk[1], Helene Kreysa[1], André Scherag[2, 3], and Stefan R. Schweinberger[1, 3, 4]

[1] *Department of General Psychology and Cognitive Neuroscience, Friedrich Schiller University Jena, Jena, Germany*
[2] *Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital, Jena, Germany*
[3] *Michael Stifel Center Jena, Jena, Germany*
[4] *German Center for Mental Health (DZPG), Site Jena-Magdeburg-Halle, Germany*

August 29, 2023

Control analysis conforming that text similarity is reduced for four baselines

Welch's t-test indicated that text similarity during the story period was significantly higher for stories with four baselines than for stories with two or three baselines. However, this increase might be caused by confounding effects of decreased recall and/or a lower number of subqueries for stories having four baselines. To account for potential confounding effects, we run a linear regression model (using the STATSMODELS OLS function; see [1]) treating text similarity as a dependent measure (y), the number of baselines as an independent measure (x), and recall and the number of subqueries as covariates (z_1 and z_2). Please note that we pooled stories with two and three baselines into one category because only a very few stories had two baselines. The model was specified by the formula $y \sim x + z1 + z2$. The results of the model are shown in the output 1.

As shown in output 1, text similarity significantly increased for four baselines after controlling for recall and number of subqueries (see coefficient X[T.4 BL]). The linear regression model therefore confirmed that text similarity was increased for four baselines, as indicated by the group coefficient being significantly different from zero ($\beta = .035$; $t(344) = 2.36, p = 9.4 \cdot 10^{-3}, \text{one} - \text{tailed}$).

Output 1: Results of the linear regression model.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            np.log10(y)   R-squared:                       0.096
Model:                            OLS   Adj. R-squared:                  0.088
Method:                 Least Squares   F-statistic:                     12.17
Date:                Thu, 15 Jun 2023   Prob (F-statistic):           1.38e-07
Time:                        12:06:08   Log-Likelihood:                 205.93
No. Observations:                 348   AIC:                            -403.9
Df Residuals:                     344   BIC:                            -388.4
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -0.3058      0.017    -17.787      0.000      -0.340      -0.272
x[T.4 bl]      0.0354      0.015      2.361      0.019       0.006       0.065
z1             0.0809      0.053      1.520      0.129      -0.024       0.186
z2            -0.0369      0.007     -5.034      0.000      -0.051      -0.022
==============================================================================
Omnibus:                       96.383   Durbin-Watson:                   1.885
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              240.300
Skew:                          -1.335   Prob(JB):                     6.60e-53
Kurtosis:                       6.073   Cond. No.                         19.2
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

References

[1] Seabold S, Perktold J (2010) Statsmodels: Econometric and Statistical Modeling with Python. In: Proceedings of the 9th Python in Science Conference, DOI 10.25080/majora-92bf1922-011