REPLY TO SCHMIDT ET AL.:

A robust surge of cognitive distortions in historical language

Johan Bollen[a,1], Marijn ten Thij[a,b], Fritz Breithaupt[c], Alexander T. J. Barron[a], Lauren A. Rutter[d], Lorenzo Lorenzo-Luaces[d], and Marten Scheffer[e]

In their critique, Schmidt et al. (1) claim that our analysis of book language (2) cannot meaningfully reflect society. Their arguments bear no relevance to our paper. The statement that "words in books are not clinical interviews, and word frequencies are not psychiatric assessments" is irrelevant because we make no attempts at clinical diagnoses. Their observation that "Derrida" is a more frequent book word than "The Beatles" is also a red herring: We do not compare between words, but instead follow the dynamics of phrases over time. Lastly, our tracking of cognitive distortions (CDS) markers is not an attempt to identify "negative thoughts" but rather to detect markers of language involved in the expression of distorted thinking.

Furthermore, Schmidt et al. (1) claim that our results are explained by a composition shift of the Google Books data toward more fiction since 2000.

We have to disagree. We made our observations relative to a null model that specifically controls for such changes in corpus composition and other recency effects, and reported a robust signal well above that baseline (2).

Schmidt et al. (ref. 1, figure 1) perform a linear regression analysis that shows a correlation between a word's relative frequency in fiction and its rise in prevalence. Because our CDS n-grams (2, 3) are about 43% more prevalent in fiction than English overall, a shift toward more fiction does increase CDS n-gram prevalence. However, our analyses indicate that the observed rise of fiction in the data would only cause CDS prevalence to increase 16% from 1980 to 2019, much less than the magnitude of the observed shift and accounted for by our null model.

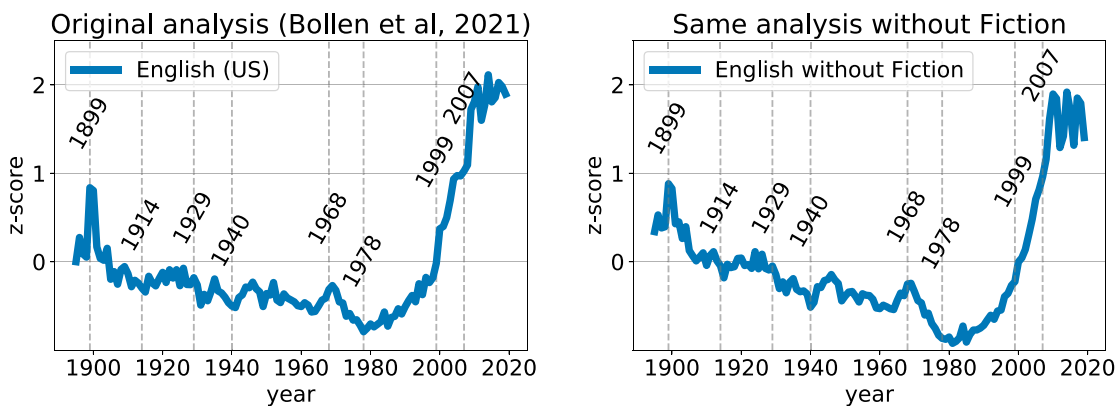Schmidt et al. (ref. 1, figure 2) make another inferential error when they draw conclusions from a

Fig. 1. (*Left*) Original results published in Bollen et al. (2). (*Right*) The same analysis with n-gram counts in the Fiction corpus subtracted from the English corpus. The comparison reveals that the original results are robust against the removal of Fiction, and can thus not be explained by the growth of Fiction in the Google Books sample.

correspondence between the prevalence of CDS n-grams and the sum of the log prevalence of their constituent words. These observations are not only compatible with our results but predicted: Changes in n-gram prevalence should match those of their constituent words. One cannot write "completely bad" without "completely" and "bad." Furthermore, both terms individually mark similar cognitive distortion types, and will thus follow a similar trajectory.

Instead of such indirect inferences or speculations, there is a more direct way to test whether our results are caused by a rise of fiction in the database. We remove the entire Fiction corpus from English by subtracting Fiction n-gram word counts from those in the English corpus. This analysis (Fig. 1) shows that the dynamics of CDS markers hardly differ from our original results. Along with the null model, this confirms that our results are unlikely to be driven by the growth of fiction in the Google Books sample.

Overly harsh critiques on the emerging field of culturomics carry the risk of throwing the baby out with the bathwater. The millions of books produced over the past centuries are not unbiased reflections of natural language. Yet, they are not uncoupled from social, cultural, and psycholinguistic changes (4–8). This implies a treasure trove of information when interpreted with care.

1 B. Schmidt, S. T. Piantadosi, K. Mahowald, Uncontrolled corpus composition drives an apparent surge in cognitive distortions. *Proc. Natl. Acad. Sci. U.S.A.*, 10.1073/pnas.2115010118 (2021).

2 J. Bollen *et al.*, Historical language records reveal a surge of cognitive distortions in recent decades. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102061118 (2021).

3 K. C. Bathina, M. t. Thij, L. Lorenzo-Luaces, L. A. Rutter, J. Bollen, Individuals with depression express more distorted thinking on social media. *Nat. Hum. Behav.* **5**, 458–466 (2021).

4 J. B. Michel *et al.*; Google Books Team, Quantitative analysis of culture using millions of digitized books. *Science* **331**, 176–182 (2011).

5 R. Iliev, J. Hoover, M. Dehghani, R. Axelrod, Linguistic positivity in historical texts reflects dynamic environmental and psychological factors. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7871–E7879 (2016).

6 P. M. Greenfield, The changing psychology of culture from 1800 through 2000. *Psychol. Sci.* **24**, 1722–1731 (2013).

7 P. S. Dodds *et al.*, Human language reveals a universal positivity bias. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2389–2394 (2015).

8 T. T. Hills, E. Proto, D. Sgroi, C. I. Seresinhe, Historical analysis of national subjective wellbeing using millions of digitized books. *Nat. Hum. Behav.* **3**, 1271–1275 (2019).

2 of 2 | PNAS
https://doi.org/10.1073/pnas.2115842118

Bollen et al.
A robust surge of cognitive distortions in historical language