Commentary

Psychological Science
1–3

DOI: 10.1177/0956797616680396
www.psychologicalscience.org/PS

No Evidence That an Ebola Outbreak Influenced Voting Preferences in the 2014 Elections After Controlling for Time-Series Autocorrelation: A Commentary on Beall, Hofer, and Schaller (2016)

Leonid Tiokhin[1,2] and Daniel Hruschka[1,2]
[1]School of Human Evolution and Social Change, Arizona State University, and [2]Center for Evolution & Medicine, Arizona State University

Received 6/19/16; Revision accepted 11/1/16

In a recent article, Beall, Hofer, and Schaller (2016) used observational time-series data to test the hypothesis that the 2014 Ebola outbreak influenced the 2014 U.S. federal elections. This represents one example of a recurring psychological interest in using observational data (a) to assess long-term temporal predictions of psychological theories in naturalistic settings (Jebb, Tay, Wang, & Huang, 2015) and (b) to examine how psychological theories can predict cross-population variation in attitudes and behavior (Eppig, Fincher, & Thornhill, 2010; Fincher & Thornhill, 2012; Gelfand et al., 2011; Murray, Schaller, & Suedfeld, 2013; Schaller & Murray, 2008). While such nonexperimental designs hold considerable promise, they also introduce analytic challenges that can lead to spurious inferences if left unaddressed (Hackman & Hruschka, 2013; Hruschka & Hackman, 2014; Hruschka & Henrich, 2013; Jebb et al., 2015; Pollet, Tybur, Frankenhuis, & Rickard, 2014). Here, we use Beall et al.'s analyses to illustrate how using observational data without attention to one long-recognized threat to inference in time-series data—temporal autocorrelation—can lead to spurious inferences (Yule, 1926).

Beall et al. used the coincidence of the 2014 Ebola epidemic and the 2014 U.S. federal elections (as well as ancillary analyses of Canadian elections) to assess two hypotheses derived from theories of the behavioral immune system (Schaller & Murray, 2008). First, they hypothesized that perceived threat of disease should increase political conservatism. Second, they hypothesized that disease threats may increase conformism and lead to a *bandwagon effect*, "the phenomenon in which voters show an increased inclination to support whichever political candidate is leading in recent polls" (p. 596). Beall et al. assessed these hypotheses by correlating 2-month time series of (a) online searches for the term "Ebola" and (b) daily polling data for U.S. congressional elections, a month before and a month after the Centers for Disease Control and Prevention's announcement of the first Ebola case in the United States (September 30, 2014). Beall et al. found strong correlations between daily Ebola search volumes during the months of September and October and support for conservative candidates at national and state levels over that same time period. They interpreted this correlation between time series as support for their first hypothesis. Beall et al. also found that correlations between Ebola searches and Republican support were stronger in states that started off with greater support for Republican candidates and with long-standing Republican voting norms, and they interpreted this result as support for the bandwagon effect.

These analyses relied on correlations between two time-series variables—Ebola search volume and daily polling—taken over 2 months. When two variables evolve over time, they can frequently look highly correlated, even without any underlying causal relationship between them (Yule, 1926; see Koplenig & Müller-Spitzer, 2016, for an illustrative example). This results from temporal autocorrelation—greater similarity in data points that are closer to each other in time—and the common existence of long-run trends in time-series data that can create

Corresponding Author:
Leonid Tiokhin, School of Human Evolution and Social Change, Arizona State University, 900 Cady Mall, Tempe, AZ 85281
E-mail: ltiokhin@asu.edu

many non-independent data points (Jebb et al., 2015). One simple method for dealing with such threats is to *detrend* (i.e., remove the long-term trend from) the time series by analyzing the changes between time points rather than their absolute values. This removes first-order autocorrelation and is often the first step in time-series analysis (Jebb et al., 2015). Calculating changes between absolute values leads to the "loss" of the first observation in the time series. However, in time series in which observations are highly autocorrelated, this does not necessarily represent the real loss of an independent data point, because data points are highly nonindependent.

Here, we applied this simple detrending procedure to the Beall et al. time series and reanalyzed the data (see the Supplemental Material available online for further details). First, we found exceedingly high levels of temporal autocorrelation in the time-series variables (rs > .90). In other words, each observation was nearly perfectly correlated with the observation that came directly before it in the time series. This indicated that detrending was a necessary first step in analyzing the time series (see Table S1 in the Supplemental Material). By detrending the data, we were then able to compare changes between adjacent observations rather than simply compare the absolute values of those observations.

After detrending the data, we found no empirical support for either of the original two hypotheses (Table 1). At both national and state levels, there were no longer strong or significant associations between Ebola search volume and preference for conservative candidates in the U.S. federal elections. The strong correlation in the Canadian elections (based on only nine data points) was still strong but no longer significant and had exceedingly wide confidence intervals. Moreover, there was no support for a moderating bandwagon effect: States leaning Republican

Table 1. Comparison of Correlations Between "Ebola" Search Volume and Measures of Voter Intentions

Measure of voter intentions	Beall, Hofer, and Schaller's (2016) analysis	Detrended analysis (first-order autocorrelation removed)
U.S. national elections	.51*	.30
Canadian national elections	.69*	.65
State elections	.31*	.04
Republican-led polls	.51***	.09
Democrat-led polls	−.08	−.03
Positive (Republican) Partisan Voter Index score	.55***	.13
Negative (Democratic) Partisan Voter Index score	−.12	−.05

Note: Beall et al. examined U.S. national elections in Study 1 and Canadian national elections in Study 3. All other correlations refer to the state-level analyses of Study 2.
*$p < .05$. ***$p < .001$.

Table 2. Comparison of Differences (Cohen's d) Between Correlations of "Ebola" Search Volume and Measures of Voter Intentions

Measure of voter intentions	Beall, Hofer, and Schaller's (2016) analysis	Detrended analysis (first-order autocorrelation removed)
Republican vs. Democrat leading polls	0.92*	0.24
Positive vs. negative Partisan Voter Index states	1.11**	0.37

*$p < .05$. **$p < .01$.

in either current or past elections did not show correlations greater than zero or correlations greater than those observed in Democratic states (Table 2). These results were robust to the composition of the sample (including or excluding outliers and excluding or including six states with insufficient data on daily changes; see Table S1 in the Supplemental Material).

Given that Beall et al.'s findings were not robust to basic time-series controls and were based on particularly small samples, this strongly suggests that either (a) these initial findings were spurious or (b) the study design used by Beall et al. was insufficiently powered to detect any potential associations or to test the proposed hypotheses. The latter is a clear possibility. For example, the statistical power to detect a statistically significant correlation between fully detrended time series would have been less than 0.5 in both the U.S. study (n = 23, observed r = .3, α = .05) and Canadian study (n = 8, observed r = .6, α = .05), whereas data from both studies still exhibit substantial second-order correlation (see Table S2 in the Supplemental Material). Many sources of randomness, such as measurement error in either the dependent or independent variables, would further increase the likelihood of null findings. These are all potential limitations of the data used in the original Beall et al. study and reanalyzed here.

We have described one of the simpler tools—detrending to remove first-order correlation—to deal with inferential threats that arise in observational data analysis. Autocorrelation of observed time series is by no means the only threat to inference when working with observational data. For example, using smoothed data, as in Beall et al.'s article (an issue we describe in more detail in the Supplemental Material), can also lead to spurious correlations. Many other useful analytic techniques exist for observational data analysis and are necessary for avoiding common pitfalls. For time-series data, one can also model and remove higher-order trends and seasonality, as well as other factors that introduce temporal autocorrelation (Jebb et al., 2015; in the Supplemental Material, we describe additional simulation approaches for checking inferences that can be used if researchers choose not to detrend their data). For cross-population comparisons that may be subject to pseudoreplication of units (e.g., Mississippi and Alabama may not really be independent observations in analyses across the 50 U.S. states), one can introduce controls for macroregional variation (Hruschka & Henrich, 2013), conduct spatially autocorrelated regressions (Anselin & Bera, 1998), or remove cultural autocorrelation by looking at changes over cultural phylogenies (Mace & Holden, 2005). To deal with potentially unmeasured confounding variables that are particularly pernicious in observational data, there are fixed-effects models for panel data (Allison, 2009) and instrumental-variable analyses (Angrist, Imbens, & Rubin, 1996). There is a rich literature addressing each of these that includes checks on the assumptions and appropriate implementation of these techniques to best avoid inferential threats introduced by these myriad issues.

Action Editor

D. Stephen Lindsay served as action editor for this article.

Author Contributions

L. Tiokhin and D. Hruschka conceived and designed the study. L. Tiokhin analyzed the data. L. Tiokhin and D. Hruschka drafted, revised, and approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797616680396

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/d9jfz/. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797616680396. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

References

Allison, P. D. (2009). *Fixed effects regression models* (Vol. 160). Thousand Oaks, CA: Sage.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455. doi:10.1080/01621459.1996.10476902

Anselin, L., & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah & D. E. A. Giles (Eds.), *Statistics: Textbooks and monographs: Vol. 155. Handbook of applied economic statistics* (pp. 237–289). New York, NY: Marcel Dekker.

Beall, A. T., Hofer, M. K., & Schaller, M. (2016). Infections and elections: Did an Ebola outbreak influence the 2014 U.S. federal elections (and if so, how)? *Psychological Science*, *27*, 595–605. doi:10.1177/0956797616628861

Eppig, C., Fincher, C. L., & Thornhill, R. (2010). Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3801–3808. doi:10.1098/rspb.2010.0973

Fincher, C. L., & Thornhill, R. (2012). Parasite-stress promotes in-group assortative sociality: The cases of strong family ties and heightened religiosity. *Behavioral & Brain Sciences*, *35*, 61–79. doi:10.1017/S0140525X11000021

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., . . . Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*, 1100–1104. doi:10.1126/science.1197754

Hackman, J., & Hruschka, D. (2013). Fast life histories, not pathogens, account for state-level variation in homicide, child maltreatment, and family ties in the U.S. *Evolution & Human Behavior*, *34*, 118–124. doi:10.1016/j.evolhumbehav.2012.11.002

Hruschka, D. J., & Hackman, J. (2014). When are cross-group differences a product of a human behavioral immune system? *Evolutionary Behavioral Sciences*, *8*, 265–273. doi:10.1016/j.evolhumbehav.2012.11.002

Hruschka, D. J., & Henrich, J. (2013). Institutions, parasites and the persistence of in-group preferences. *PLoS ONE*, *8*(5), Article e63642. doi:10.1371/journal.pone.0063642

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology, 6*, Article 727. doi:10.3389/fpsyg.2015.00727

Koplenig, A., & Müller-Spitzer, C. (2016). Population size predicts lexical diversity, but so does the mean sea level – why it is important to correctly account for the structure of temporal data. *PLoS ONE*, *11*(3), Article e0150771. https://doi.org/10.1371/journal.pone.0150771

Mace, R., & Holden, C. J. (2005). A phylogenetic approach to cultural evolution. *Trends in Ecology & Evolution*, *20*, 116–121. doi:10.1016/j.tree.2004.12.002

Murray, D. R., Schaller, M., & Suedfeld, P. (2013). Pathogens and politics: Further evidence that parasite prevalence predicts authoritarianism. *PLoS ONE*, *8*(5), Article e62275. doi:10.1371/journal.pone.0062275

Pollet, T. V., Tybur, J. M., Frankenhuis, W. E., & Rickard, I. J. (2014). What can cross-cultural correlations teach us about human nature? *Human Nature*, *25*, 410–429. doi:10.1007/s12110-014-9206-3

Schaller, M., & Murray, D. R. (2008). Pathogens, personality, and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology*, *95*, 212–221. doi:10.1037/0022-3514.95.1.212

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series?–a study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, *89*, 1–63.

SUPPLEMENTAL MATERIALS FOR:

No evidence that Ebola outbreak influenced voting preferences in 2014 elections, after controlling for autocorrelation in time series.

Leonid Tiokhin[1,2] , Daniel Hruschka[1,2]

[1] Arizona State University School of Human Evolution and Social Change

[2] Arizona State University Center for Evolution & Medicine

In Press, Psychological Science

Open access materials available at https://osf.io/d9jfz/

MATERIALS AND METHODS

Sample. We used the same data and variables analyzed by Beall et al. in their analyses. Study 1 considered national-level polling data and Ebola search volume data between September 1 and November 1 aggregated across the U.S. Beall et al. analyzed the relationship between U.S. country-level voter intentions and U.S. country-level Google searches for "Ebola. In the time frame of September 1, 2014 to November 1, 2014, aggregate nationwide polling results were available for 24 days – 9 days preceding the initial Ebola outbreak and 15 days following the initial outbreak. Beall et al. conducted analyses using this entire sample, as well as an 8-value subset of this sample (the last week of September and first week of October).

Study 2 considered the same variables and time frame as study 1, but aggregated at the state Senate-election level for 34 out of 36 elections in which data was available (Kansas and Alabama had insufficient data and were excluded). The variables were either measured at the level of country (Google searches for "Ebola") or individual Senate elections (Within-state voter intentions for each senate election). Of these 34 states, two (Hawaii and Rhode Island) had outlier state-Senate Voter Intention Index scores, and Beall et al. conducted analyses both with and without them. Beall et al. excluded Virginia from moderation analyses of Partisan Voter Index, at it had score of 0 (See “Variables”).

Study 3 considered the same variables and time frame as study 1, but used Canadian nationwide polling results and Google Searches for “Ebola”. In the time frame of September 1, 2014 to November 1, 2014, aggregate nationwide polling results were available for 9 days. Beall et al. solely conducted analyses using this sample.

While we attempted to use the same dataset as Beall et al., there were some differences between our dataset and theirs. While the original Beall et al. analyses for study 1 contained data points for each of 24 days, our study 1 analyses contained change scores for 23 pairs of days. Controlling for first-order autocorrelation required calculating changes from one data value to the next. This resulted in the loss of the first data-point in the time series, for which changes could not be calculated. In study 2, calculating changes to control for first-order autocorrelation resulted in the exclusion of six state-Senate elections (Idaho, Mississippi, Nebraska, Tennessee, West Virginia, and Wyoming). These had insufficient data points to calculate correlations (i.e. less than three) and were excluded from our analyses. Because our resulting sample of state-senate elections differed from that of Beall et al., we also re-conducted the original Beall et al. analyses, excluding these states. In study 3, the original Beall et al. analyses contained data points for 9 days, while our analyses contained change scores for 8 pairs of days. As in study 1, controlling for autocorrelation resulted in the loss of the first data-point in the time series, for which changes could not be calculated. Table S1 summarizes exclusions, inclusions, and results across all analyses.

Table S1 – Comparison of results across three approaches: (1) original analysis, (2) original analysis on the subsample used for detrending, (3) detrended analysis (i.e. first-order autocorrelation removed) on subsample.

	Original analysis (before detrending)	Subsample analysis (before detrending)	Detrended analysis on subsample
Correlations of Ebola Search Volume & Voter Intentions	**r**	**r**	**r**
U.S. National[a]	0.51*	NA	0.30
Canada National [c]	0.69*	NA	0.65
All States [b]	0.31*	0.19	0.04
Republican-Leading Elections [b]	0.51***	0.39*	0.09
Democratic-Leading Elections [b]	–0.08	–0.08	–0.03
Positive (Republican) PVI [b]	0.55***	0.43*	0.13
Negative (Democratic) PVI [b]	–0.12	–0.12	–0.05
Differences between correlations	**d**	**d**	**d**
Republican vs. Democratic-Leading Elections [b]	0.92*	0.70	0.24
Positive vs. Negative PVI states [b]	1.11**	0.85*	0.37

Table S1 lists correlations between the Ebola Search Volume Index (ESVI) and the Voter-Intention Index (VII) across all analyses. Column 1 lists correlations in the original Beall et al. analysis. Detrending (Column 3) was only possible on a subsample of the original Beall et al. data. To demonstrate that using this subsample did not fundamentally alter the relationships in the original Beall et al. data, we replicated the original Beall et al. analyses using this same subsample (Column 2). * Significant at $p < 0.05$ ** Significant at $p < 0.01$ *** Significant at $p < 0.001$. [a] refers to Study 1. [b] refers to Study 2. [c] refers to Study 3.

Variables. *Voter Intention Index (VII).* Voter intentions were estimated from nationwide polling data from Pollster, a poll-aggregation website. For each day on which data was available, the Pollster website specified the percentage of potential voters within the United States who indicated an intention to vote for each candidate from either the Republican or Democratic Party. To generate VII values, Beall et al. subtracted the percentage of voters who intended to vote Democrat from the percentage of voters who intended to vote Republican. As such, positive VII values indicated nationwide preference for Republican candidates and negative VII values indicated nationwide preference for Democrats (See Fig. S1). For study 2, Beall et al. used the same procedure as in study 1 to generate the VII for each state-Senate election on the days such polling data was available. For study 3, Beall et al. used the same procedure as in study 1 to generate the VII based on Canadian nationwide polling data regarding preferences for the Conservative Party versus New Democratic Party. Positive VII values indicated Canadian nationwide preference for the Conservative Party and negative VII values indicated Canadian nationwide preference for the New Democratic Party (See Fig. S2).

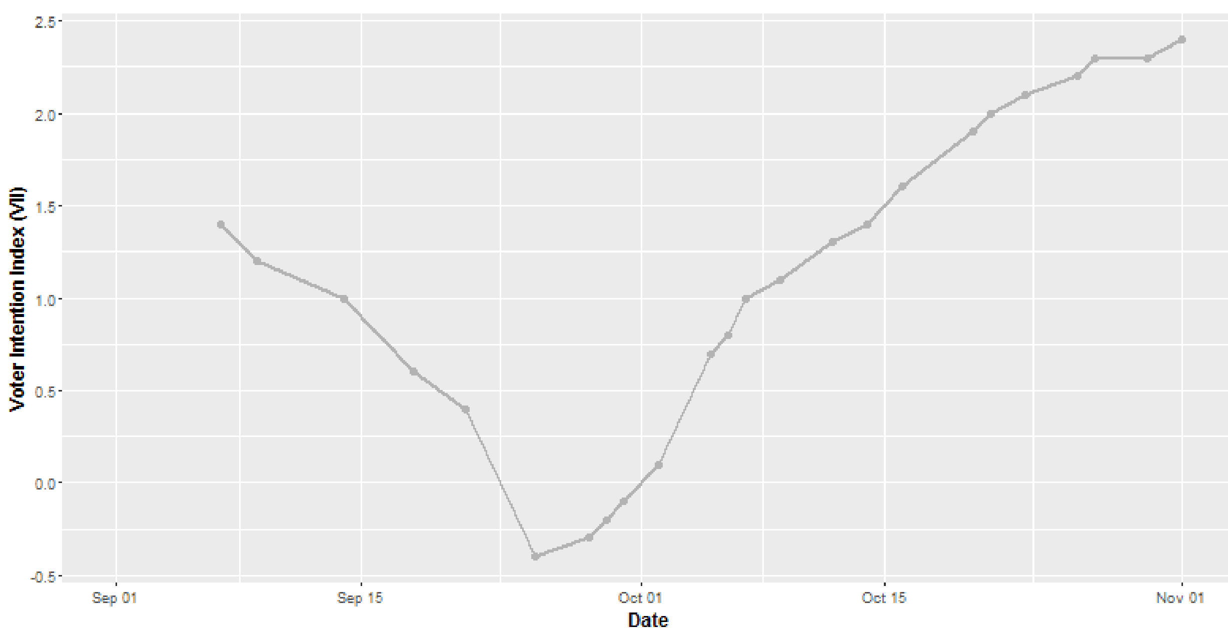

Fig. S1. Nationwide Voting Intentions across Time, Operationalized via the Voter Intention Index (VII). The Voter Intention Index (VII) steadily decreases throughout most of September, and steadily increases from late September to November 1.

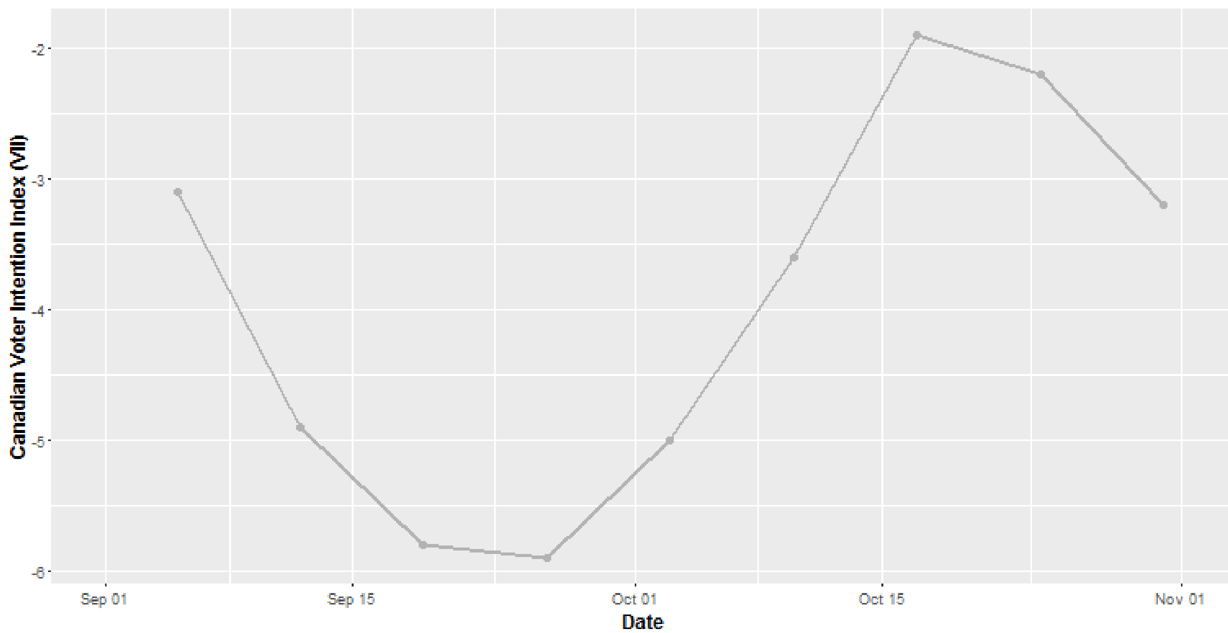

Fig. S2. Canadian Nationwide Voting Intentions across Time, Operationalized via the Canadian Voter Intention Index (VII). The Canadian VII decreases throughout most of September, and steadily increases from late September to late October, before again decreasing in late October.

Ebola Search Volume Index (ESVI). For the time period of August 26, 2014 to November 1, Beall et al. obtained internet search volume data for the term "Ebola" from Google Trends. For each day from September 1, 2014 to November 1, 2014, Beall et al. took the arithmetic mean "Ebola" daily search volume from the previous 7 days, ending on and including the specified day. This resulted in an Ebola search-volume index (ESVI). The ESVI was used in both study 1 and study 2 as the measure of Ebola's psychological salience (See Fig. S3.). For study 3, Beall et al. used Google Trends data from Canada to assess nationwide internet search volumes for "Ebola" (See Fig. S4.).

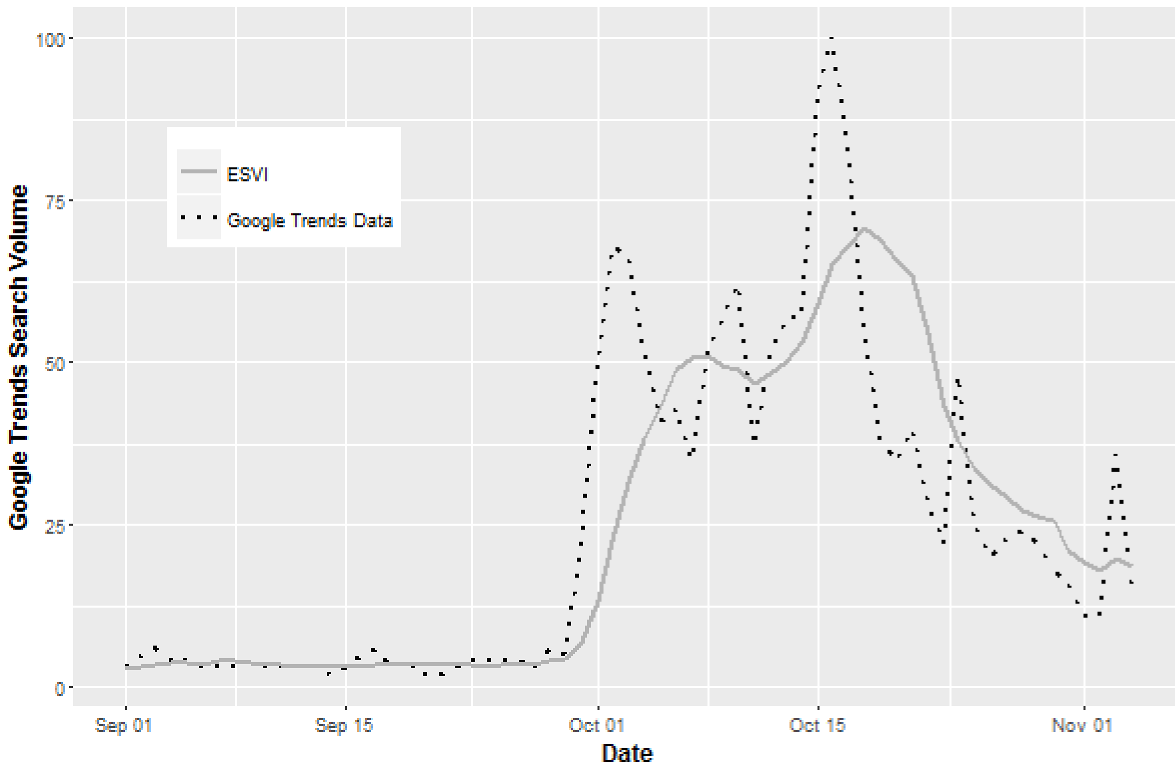

Fig. S3. Ebola's Psychological Salience across Time, Operationalized via Google Search Volumes for the term "Ebola". The Ebola Search Volume Index (ESVI) is insensitive to daily fluctuations in raw Google search volumes for "Ebola"

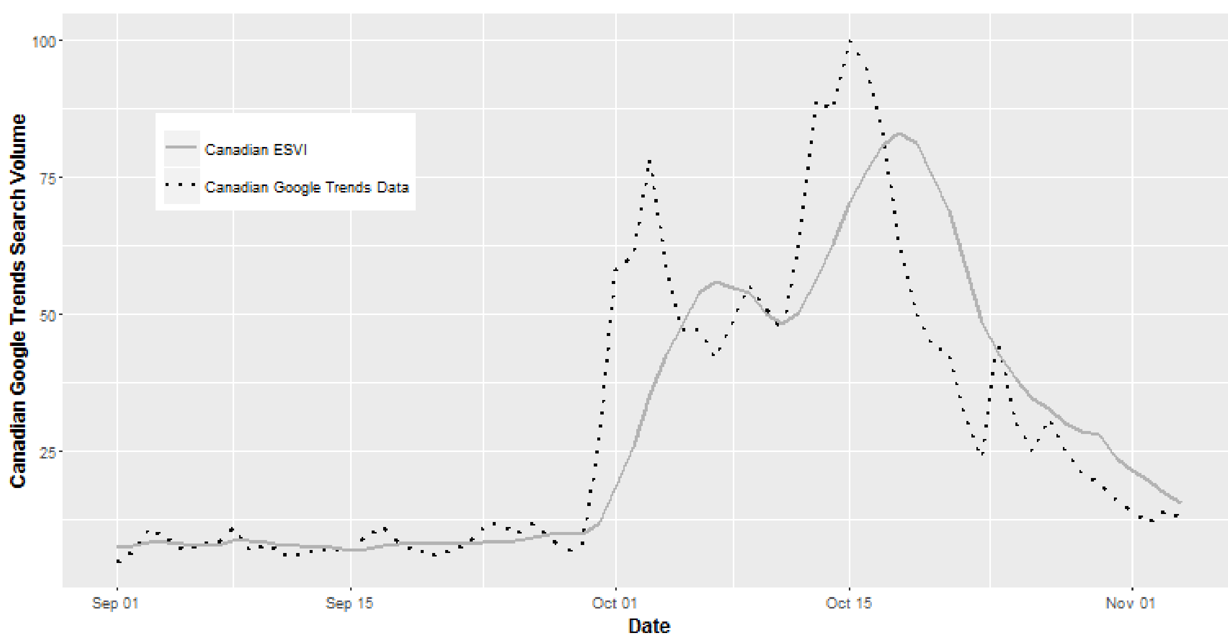

Fig. S4. Ebola's Psychological Salience in Canada across Time, Operationalized via Canadian Google Search Volumes for the term "Ebola". The Canadian Ebola Search Volume Index (ESVI) is insensitive to daily fluctuations in raw Canadian Google search volumes for "Ebola".

Candidate Leading Polls at Time of Ebola Outbreak: In study 2, Beall et al. used VII values to categorize each state-Senate election as being led by a Democratic (n=11) or Republican (n=22) candidate at the time of the Ebola outbreak. This categorization was based on the most recent poll preceding the outbreak.

Partisan Voter Index (PVI): Beall et al. use data from the 2014 Cook Political Report's Partisan Voter Index (httw://www.cookpolitical.com/story/5604) as a measure of a state's "enduring political norms". While this link only lists district-specific PVI scores for 2014, Beall et al. use state-level PVI scores in their analyses. How state-level PVI scores were generated was not reported in Beall et al.'s paper, but it appears that a state's PVI score is the arithmetic mean of the PVI scores of all districts in that state. States with positive PVI scores were categorized as generally Republican (n=19) and states with negative PVI scores were categorized as generally Democratic (n=12). Virginia had a PVI score of 0 and was excluded from PVI analyses.

ANALYTICAL STRATEGY

In study 1, Beall et al. examined correlations between national-level VII time series and 1) ESVI time series for the 24 days in September and October for which this data was available, as well as 2) raw Ebola search volumes for the 8 days during the two-week period that included the last week of September and the first week of October. In study 2, Beall et al. assessed correlations between state-specific VII values and ESVI values, and examined whether these correlations were greater in states that had higher initial support for Republicans. In study 3, Beall et al. examined correlations between Canadian national-level VII time series for the 9 days in September and October for which this data was available.

These initial analyses did not investigate first-order autocorrelation or fully detrend the data to address it. To assess levels of first-order autocorrelation, we re-analyzed the same variables by lagging an observation from each variable (x) to the temporally adjacent observation (x + 1), and then correlated x with x+1. The resulting correlation is a measure of how predictive the value of a variable at timepoint x is of the value of that same variable at timepoint x + 1.

To remove first-order autocorrelation from the variables, we subtracted the value of that variable at time x from its value at time x+1. We then repeated the Beall et al. analyses with these detrended variables. For study 1, we created a lagged VII variable (lagged from the temporally prior observation), and subtracted the original VII data from this lagged VII variable. This resulted in a new variable, “VII Changes”, which contained changes for 23 pairs of adjacent days compared to 24 days for which the original VII had data (1 day lost due to calculating the changes). “VII Changes” measures the marginal increase or decrease in the VII during any given time period. We used this same strategy to calculate the changes in ESVI, for each day that data on “VII Changes” was also available. This resulted in an “ESVI Changes” variable, which contained data for the same 23 days for which “VII Changes” data was available. Once we created these variables, we ran bivariate correlations, for both the entire time period for which data was available and for just the two-week period that included the last week of September and the first week of October, replicating Beall et al.’s analysis.

For study 2, we used the same strategy as study 1 to calculate the changes in VII and ESVI for each state-Senate election. Because daily changes were calculated by subtracting a variable’s value at time x from its value at time x+1, this resulted in “State-Specific VII Changes” and “State-Specific ESVI Changes” variables of size n-1, where n was the total number of data points for that election in the original data. We then re-conduct Beall et al.’s correlational time-series analyses, controlling for first-order autocorrelation, by using "State-Specific VII Changes" instead of "VII" and "State-Specific ESVI Changes" instead of "ESVI".

For study 3, we used the same strategy as study 1 to calculate the changes in Canadian VII and Canadian ESVI, for the 9 days for which data was available. This resulted in "Canadian VII Changes" and "Canadian ESVI Changes" variables that contained data regarding changes for 8 pairs of adjacent days (1 day lost due to calculating the changes). As in study 1, we ran bivariate correlations between these variables for the entire time period for which data was available.

We also directly replicated Beall et al.'s results and report 95% Confidence Intervals for their correlations.

RESULTS

Is there temporal autocorrelation in the time series?

Table S2 summarizes first-order autocorrelations of the key variables in studies 1, 2, and 3. All variables in the original Beall et al. data exhibited extremely high levels of first-order autocorrelation. This was especially true for the 7-day aggregate ESVI used in the original study, because each aggregate ESVI value was made up of 6 of the 7 same values as the previous aggregate ESVI value. Calculating the changes between data values resulted in strongly attenuated levels of autocorrelation, although some autocorrelation still remained for VII Changes and ESVI changes. Three state-Senate elections (Montana, South Carolina 1, and South Carolina 2) had insufficient data points to calculate autocorrelation in VII Changes and were excluded from this analysis. There was no evidence that these exclusions affected the level of autocorrelation in the raw state-Senate election VII data ($r = 0.75$, $N = 23$, $p < 0.001$, 95% CI [0.58, 0.92]).

Table S2: Autocorrelation of key variables.

	N - Beall et al. Data	Raw Beall et al. Data	Detrended Beall et al. data
Raw Ebola Search Volumes (U.S. national) [a, b]	23	0.64*** [0.32, 0.84]).	–0.37 [–0.69, 0.06]
ESVI (U.S. national) [a, b]	23	0.94*** [0.85, 0.97]	0.60** [0.23, 0.81]
ESVI (Canada national) [c]	8	0.63 [–0.14, 0.92]	0.11 [–0.70, 0.80])
VII (U.S. national) [a]	23	0.95*** [0.89, 0.98]),	0.43* [0.10, 0.72]
VII (Canada national) [c]	8	0.69 [–0.04, 0.94]	0.58 [–0.31, 0.93]
VII (U.S. State-level average) [b]	26	0.76*** [0.61, 0.92]	0.03 [–0.16, 0.22]

Table S2 lists first-order autocorrelations of the key variables in studies 1, 2, and 3 (Column 1). Column 2 lists autocorrelations after detrending (i.e. removing 1st-order autocorrelation). 95% Confidence Intervals are listed in brackets. * Significant at $p < 0.05$ ** Significant at $p < 0.01$ *** Significant at $p < 0.001$. [a] refers to Study 1. [b] refers to Study 2. [c] refers to Study 3. N of detrended data is always one less than the raw data.

Do internet searches for "Ebola" predict people's intentions to vote for conservative candidates at the national level?

In study 1, Beall et al. found that the ESVI was positively and significantly correlated with the VII (r = .51, p = 0.012, N = 24 days, 95% CI [0.13, 0.76]). They also found a large, positive but non-significant correlation between raw Google Trends "Ebola" search volumes and the VII for the two-week period that included the last week of September and the first week of October (r = 0.61, p=0.111, N = 8 days, 95% CI [–0.17, 0.92]).

When removing first-order autocorrelation, the correlation between VII Changes and ESVI Changes was substantially lower, and no longer significant (r = 0.30, p = 0.159, N = 23, 95% CI [–0.12, 0.64]) for the months of September and October. When we replicated the additional Beall et al. correlation of raw Google Trends "Ebola" search volume with VII over the two-week period that included the last week of September and the first week of October, the correlation of detrended variables was still non-significant, but also no longer had the suggestively high, positive magnitude ($r = –0.19$, $p > .250$, $N = 8$ days, 95% CI [–0.79, 0.60].

In study 3, Beall et al. found that the Canadian ESVI was positively and significantly correlated with the Canadian VII ($r = .69$, $p = 0.042$, $N = 9$ days, 95% CI [0.04, 0.93]), as well as the Canadian voter-intention-change-index ($r=0.76$, $p = 0.017$, $N = 9$ days, 95% CI [0.20, 0.95]).

When removing first-order autocorrelation, the correlation between Canadian ESVI Changes and Canadian VII Changes was slightly reduced, and no longer significant ($r = 0.65$, $p = 0.08$, $N = 8$ days, 95% CI [–0.11, 0.93]). We did not analyze the relationship between Canadian ESVI and the Canadian *voter-intention-change-index* (See Detailed Discussion for justification)

Do internet searches for "Ebola" predict people's intentions to vote for Republican candidates at the state-level?

In study 2, Beall et al. found that, across 32 Senate elections, the arithmetic mean of the correlations between ESVI and the state-specific VII was significantly greater than 0 (mean $r = 0.31$, $p = 0.016$, $N = 32$, CI [0.06, 0.55]), though it reduced (mean $r = 0.24$, $p = 0.057$, $N = 34$, 95% CI [–0.01, 0.49]) if outliers Hawaii and Rhode Island were included. Most of these correlations took on either extreme positive or extreme negative values (see Fig. S5.). When considering only those 28 states with sufficient data for later comparison with a detrended analysis (at least three time points), the arithmetic mean of the correlations between ESVI and the state-specific VII was lower with the two outliers excluded (mean $r = 0.19$, $p = 0.18$, $N = 26$, 95% CI [–0.09, 0.47]) and included (mean $r = 0.12$, $p = 0.18$, $N = 28$, 95% CI [–0.16, 0.39]).

Once detrending the two key variables, the arithmetic mean of correlations between "ESVI Changes" and "VII Changes" across 28 Senate elections was reduced to nearly zero and was no longer significant (mean r = 0.04, p > .250, N = 26, CI [–0.16, 0.23]) and was (mean r = 0.03, p > .250, N = 28, CI [–0.15, 0.21]) when outliers Hawaii and Rhode Island were included. Neither of these were significantly different from 0. The detrended correlations no longer took on extreme negative or extreme positive values, and their distribution more closely resembled a normal distribution (see Fig. S5.).

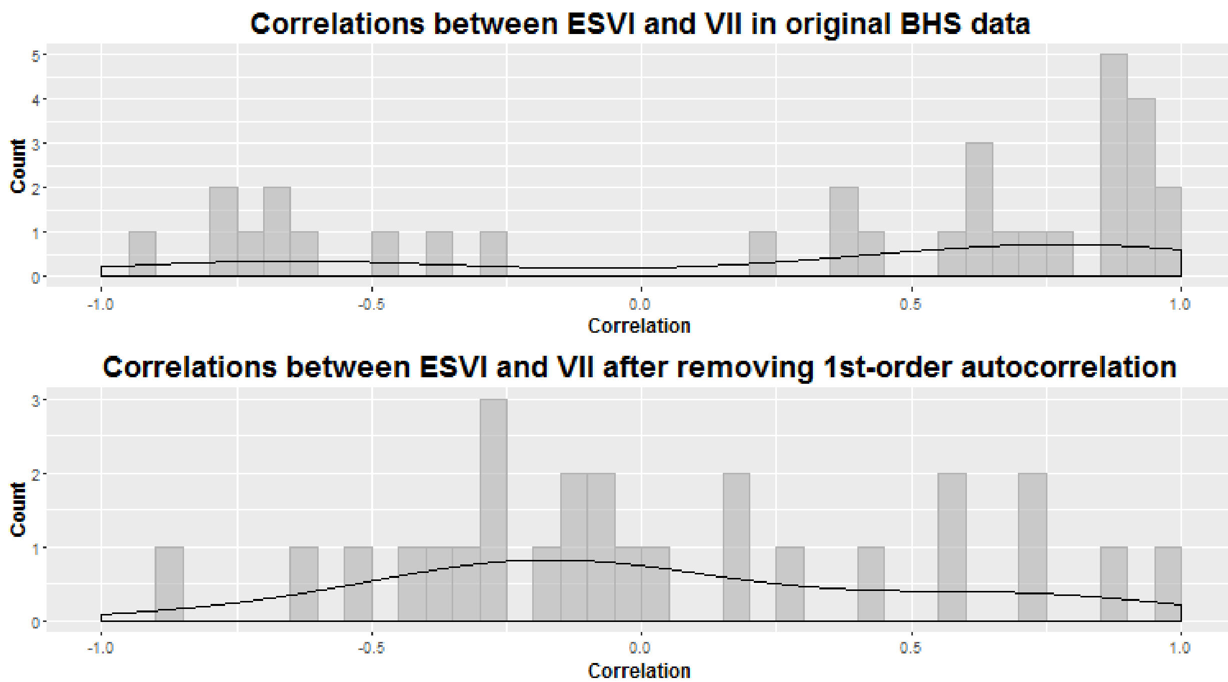

Fig. S5. Distribution of Correlations between Ebola Search Volume Index (ESVI) and Voter Intention Index (VII) for 1) The original 32-election Beall et al. data and 2) A 26-election subset of the Beall et al. data after removing 1^{st}-order autocorrelation. Correlations in the original 32-election Beall et al. data take on extreme values and were bimodally distributed. Removing 1^{st}-order autocorrelation resulted in a distribution of correlations more closely resembling a normal distribution.

Does a state-level bandwagon effect moderate the relationship between internet searches for “Ebola” and people’s intentions to vote for Republican candidates?

In the second part of study 2, Beall et al. found that, across 32 Senate elections, the arithmetic mean correlation between ESVI and the state-specific VII was positive for elections in which a Republican led the polls at the time of the initial Ebola Outbreak (mean r = 0.51, N = 21,

$p < 0.001$, CI [0.26, 0.75]) but not for elections in which a Democrat led the polls at that time (mean $r = –0.08$, $N = 11$, $p > .250$, CI [–0.60, 0.44). This difference was statistically significant ($d = 0.92$, $t(30) = 2.52$, $p = 0.017$). With outliers Hawaii and Rhode Island included, they found a similar pattern for Republican leading states (mean $r = 0.51$, $N = 21$, $p < 0.001$, 95% CI [0.26, 0.75]) and for Democratic leading states (mean $r = –0.19$, $N = 13$, $p > .250$, 95% CI [–0.65, 0.27]). When considering only those 28 states that had at least three time points (for comparison with a later detrended analysis), the arithmetic mean correlation between ESVI and the state-specific VII was still positive and statistically significant (mean $r= 0.39$, $N = 15$, $p = .021$, 95% CI [0.07, 0.70]) for Republican leading states but not for Democratic leading states (mean $r = –0.08$, $N = 11$, $p > .250$, CI [–0.60, 0.44]). The difference between these was not statistically significant (Welch Two Sample t-test ($t = 1.69$, $df = 17.64$, $p = 0.109$)). With outliers Hawaii and Rhode Island included, the pattern was similar for Republican (mean $r = 0.39$, $N = 15$, $p = .021$, 95% CI [0.07, 0.70]) and Democratic (mean $r = –0.19$, $N = 13$, $p > .250$, 95% CI [–0.65, 0.27]) leading states.

In contrast, when removing 1st order autocorrelation, we found no evidence of a moderating effect of party leading the polls at the time of the outbreak on the correlation between Ebola searches and people's voting intentions. The arithmetic mean correlation between "ESVI Changes" and "VII Changes" across 26 Senate elections was (mean $r = 0.09$, $N = 15$, $p >.250$, 95% CI [-0.18, 0.35]) for elections in which a Republican led the polls at the time of the outbreak and (mean $r = –0.03$, $N = 11$, $p > .250$, 95% CI [–0.38, 0.31]) for elections in which a Democrat led the polls at the time of the outbreak. This pattern held when outliers Hawaii and Rhode Island were included for Republican leading elections (mean $r = 0.09$, $N = 15$, $p >.250$, 95% CI [–0.18, 0.35]) and Democrat leading elections (mean $r = –0.03$, $N = 13$, $p >.250$, 95% CI

[–0.32, 0.25]). Neither of these were significantly different from 0. Further, they were not significantly different from each other (Welch Two Sample t-test, outliers excluded: $t = 0.59$, $df = 20.60$, $p > .250$; Welch Two Sample t-test, outliers included: ($t = 0.66$, $df = 25.49$, $p > .250$)).

In a related analysis, Beall et al. found that, across 31 Senate elections, the arithmetic mean correlation between ESVI and the state-specific VII was positive for elections with Republican-leaning PVI scores (mean $r = 0.55$, $N = 19$, $p < 0.001$, 95% CI [0.30, 0.81]) but not for those with Democrat-leaning PVI scores (mean $r = –0.12$, $N = 12$, $p > .250$, 95% CI [–0.58, 0.34]). This difference was statistically significant ($d = 1.11$, $t(29) = 3.00$, $p = 0.005$). With outliers Hawaii and Rhode Island included, they found a similar pattern (mean $r = 0.55$, $N = 19$, $p < 0.001$, 95% CI [0.30, 0.81]) for positive PVI states and (mean $r = –0.22$, $N = 14$, $p > .250$, 95% CI [–0.63, 0.19]) for negative PVI states. When considering only those 27 states that had at least three time points for comparison with a later detrended analysis, the arithmetic mean correlation between ESVI and the state-specific VII was positive, and statistically significant for Republican-leaning states (mean $r = 0.43$, $N = 13$, $p = .020$, 95% CI [0.08, 0.78]) but not for Democrat-leaning states (mean $r = –0.12$, $N = 12$, $p > .250$, 95% CI [–0.58, 0.34]). This difference was statistically significant (Welch Two Sample t-test: ($t = 2.11$, $df = 21.17$, $p = 0.047$). With outliers Hawaii and Rhode Island included, the pattern was similar for Republican-leaning (mean $r = 0.43$, $N = 13$, $p = 0.020$, 95% CI [0.08, 0.78])) and Democrat-leaning (mean $r = –0.22$, $N = 14$, $p > .250$, 95% CI [–0.63, 0.19) states.

In contrast, when controlling for first-order autocorrelation, we found that the moderating effect of state PVI on the correlation between Ebola searches and people’s voting intentions disappeared. The arithmetic mean correlation between “ESVI Changes” and “VII Changes” across 25 Senate elections was small and non-significant for both Republican-leaning states (mean $r = 0.13$, $N = 13$, $p > .250$, 95% CI [–0.19, 0.45]) and Democrat-leaning states (mean $r = –0.05$, $N = 12$, $p > .250$, 95% CI [–0.34, 0.24]). This pattern held when outliers Hawaii and Rhode Island were included (Republican-leaning states: mean $r = 0.13$, $N = 13$, $p > .250$, 95% CI [–0.19, 0.45]; Democrat-leaning states: mean $r = –0.05$, $N = 14$, $p > .250$, 95% CI [–0.29, 0.19]). Neither of these were significantly different from zero. Further, they were not significantly different from each other (Welch Two Sample t-test, outliers excluded: $t = 0.92$, $df = 22.94$, $p > .250$; Welch Two Sample t-test, outliers included: ($t = 0.97$, $df = 23.07$, $p > .250$).

DETAILED DISCUSSION

In their paper, Beall et al. used time-series data to test the hypothesis that the 2014 Ebola outbreak influenced the 2014 U.S. Federal elections. They found that the volume of Google searches for "Ebola" strongly covaried with support for conservative candidates, at both the national and state level, and that the correlation between support for Republican candidates and Ebola search volume was strongest in U.S. states with greater support of Republican candidates and with longstanding Republican voting norms. However, Beall et al. assumed that time series observations were independent of each other, whereas we showed that all analyzed variables exhibited extremely high degrees of temporal autocorrelation (See Table S2). Here we reanalyzed the Beall et al. data, controlling for first-order autocorrelation, and found that essentially all relationships between the Ebola outbreak and people's voting intentions became attenuated and non-significant (See Table S1). Because the Beall et al. findings were not robust to such basic controls, this strongly suggests that either: (1) many of the initial findings were either spurious or (2) the study design used by Beall et al. was insufficiently powered to detect any associations that may actually exist.

In study 1, we found that the positive correlation between ESVI and VII in September and October dropped substantially and became non-significant after removing first-order autocorrelation. If we focused on only the two-week period that included the last week of September and first week of October, and instead analyzed the relationship between raw Google Trends "Ebola" search volumes and VII, we found that the positive correlation between these two variables disappeared entirely after removing first-order autocorrelation.

In study 2, we found that the positive correlations between national ESVI and state-Senate election VII across all available states became indistinguishable from zero after removing those states that didn't have sufficient data for detrending, and dropped to values very close to zero when removing first-order autocorrelation. Further, we also found that the moderating effects of candidate leading the polls at time of the outbreak and state-level PVI disappeared after removing first-order autocorrelation. These results were robust to the composition of the sample (including/excluding outliers; excluding/including six states with insufficient data on daily changes). The state-Senate election data was the most fine-grained of all the Beall et al. data and contained the largest number of data points. As such, it afforded the best test of the hypothesis that the psychological salience of Ebola affected people's voting intentions in during the 2014 U.S. elections. Our analyses revealed no evidence for this purported relationship: there is no evidence from this data that increases in national-level Google searches for "Ebola" were associated with people's tendency to favor Republican or Democratic candidates in state-Senate elections. Further, there is no evidence for an increased inclination to conform to popular opinion with increases in national-level Google searches for "Ebola".

In study 3, we found that the positive correlation between Canadian ESVI and Canadian VII in September and October dropped slightly and became non-significant after removing first-order autocorrelation. Study 3's small sample size (9 data points, 8 after controlling for first-order autocorrelation) meant that it provided the least informative test of any hypothesis (while the point estimate for the correlation between Canadian VII changes and Canadian ESVI changes was $r = 0.65$, the confidence interval was CI[-0.11 to 0.93]). This sample size is much too small for an appropriate time series analysis (Jebb, Tay, Wang, & Huang, 2015), and thus the original study design was likely insufficiently powered to detect potential associations or test the proposed hypotheses. Given that underpowered research designs increase the probability of generating false positives and result in exaggerated effect-size estimates (Button et al., 2013), we are hesitant to make much of study 3.

The dangers of analyzing smoothed time-series data. Beall et al. obtained their voter-intention data from "Pollster", a poll-aggregation website (http://elections.huffingtonpost.com/pollster). Pollster has several settings for data smoothing (i.e. Less Smoothing, Moderate, More Smoothing) and Beall et al. used the "more smoothed" version of the Pollster data in their analyses. This is problematic: smoothing can lead to greater temporal autocorrelation in data points, because their value is increasingly artificially determined by the value of neighboring data points. As a consequence, smoothing can lead to a number of problems including misestimation of parameters and dramatic underestimates of standard errors, resulting in spurious inferences for correlations. Beall et al.'s study 1 provides a salient example. Analyzing the "more smoothed" version of the Pollster data, Beall et al. show a plot (Figure 1) illustrating that the upward trend in intentions to vote for conservative candidates (i.e. VII) one week after the Ebola outbreak in the U.S. was stronger than the trend during the week prior to the outbreak. However, when analyzing the "less smoothed" version of the same data, we found the *opposite* result: the Ebola outbreak was associated with a reverse in the temporal trajectory of

VII scores (i.e. increased voter intentions to vote for *democratic* candidates). This illustrates one of the perils of using smoothed data, and suggests that Beall et al.'s findings may crucially depend on the smoothing procedure that generated the data[1]. (See https://osf.io/d9jfz/ for data and code. "Study1_Data_VariousLevelsOfSmoothing.csv" contains Study 1 data in its various forms. The "VoterIntentionIndex_LessSmoothed_Pollster" column is the "less smoothed" pollster data. The "VoterIntentionIndex_FromStudy1" column is the "more smoothed" data used by Beall et al. data. The "VoterIntentionIndex_Smoothed_Pollster" column is the "more smoothed" data when downloaded from Pollster (https://goo.gl/aB8VR1) on May 1, 2017. "Study1_AdditionalAnalyses.R" contains the relevant R code.).

An alternative to detrending data. In those cases when detrended time series are not appropriate for answering a specific research question, there are other potential approaches to dealing with temporal autocorrelation. One approach is to simulate the expected distribution of cross time-series correlations at different levels of lag-1 autocorrelation. Doing so allows one to plot two *causally unrelated* time series with varying levels of autocorrelation and generate the expected distribution of correlations between them. It takes as input the observed first-order temporal autocorrelation in the two empirical time series of interest (as well as the sample size). For example, below is a simulation for two time series with first-order autocorrelation of 0.90 (ar = c(0.9)) and sample size n = 23. Below, we provide the R code (also available at https://osf.io/d9jfz/, in the file titled "Study1_AdditionalAnalyses.R") for plotting this distribution. This simple simulation demonstrates that highly-autocorrelated time series (such as those analyzed by Beall et al.) will frequently be highly correlated with one another, even though

[1] This example arose out of productive conversations with the original study authors about different ways of analysing the data from Study 1.

they are causally unrelated. This approach does not necessarily remove threats from higher-order autocorrelations or other factors, but it is another useful first-line defense.

```
simnum<-10000
simul <- matrix(nrow=simnum, ncol=1, 0)
for (i in 1:simnum){
ar.sim<-arima.sim(model=list(ar=c(.9)),n=23)
ar.sim2<-arima.sim(model=list(ar=c(.9)),n=23)
simul[i] <- cor(ar.sim,ar.sim2)
}
hist(simul)
quantile(simul,c(0.025,0.975))

ar.sim<-arima.sim(model=list(ar=c(.94)),n=23)
acf(ar.sim)
```

Caveats. Here, we focused our analyses on the more fine-grained time-series data analyzed by Beall et al., as it provided the strongest test of their hypotheses. The fact that support for Republican candidates was higher after announcement of the Ebola epidemic than before the announcement was obviously suggestive, but was not particularly informative for the proposed hypotheses given that it represented 2 observations (before and after) from a single data point—national-level means. Beall et al. also constructed a *voter-intention-change-index*, which assessed the difference between that day's VII and the VII 7 days before, resulting in 16 data points (study 1) and 9 data points (study 3). This was a useful start toward a fully detrended analysis, but it only detrended one of the variables. Further, we were unable to determine why study 3's *voter-intention-change-index* contained 9 instead of 8 data points (one data point should have been lost when calculating changes). Given these issues, we created our own variables for assessing changes between data points. Finally, the Beall et al. analysis of how state-level pre-post changes differed across Republican- or Democrat-leaning states was suggestive. However, it also raises interesting questions about spatial and cultural autocorrelation that are not the direct topic of this paper, but deserve additional exploration (Hruschka & Henrich, 2013).

Our re-analysis of the Beall et al. data had a number of limitations. Because of the small sample size in the original data, the analyses were insufficiently powered to detect potential associations. Of course, this is a concern with any analyses of the existing Beall et al. data, suggesting the current data may not be sufficiently powered to test these hypotheses. Indeed, some guidelines suggest a minimum of 50 data points for time series analyses (Jebb et al., 2015). After controlling for first-order autocorrelation, none of the purported associations between Google searches for "Ebola" and people's voter intentions remained significant, though it is possible that such associations can only be detected in a longer time-series. Thus, while the current analyses have a number of limitations, these are the same caveats that apply to the original Beall et al. analyses.

References

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. http://doi.org/http://dx.doi.org/10.1038/nrn3475

Hruschka, D. J., & Henrich, J. (2013). Institutions, parasites and the persistence of in-group preferences. *PloS One*, *8*(5), e63642. http://doi.org/http://dx.doi.org/10.1371/journal.pone.0063642

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: examining and forecasting change. *Frontiers in Psychology*, *6*. http://doi.org/http://dx.doi.org/10.3389/fpsyg.2015.00727