

## **‘We Shall Overcome’: First-Person Plural Pronouns from Search Volume Data Predict Collective Action across the United States**

Collective action is a key driver of social and political change within societies. So far, the main factor mobilizing individuals into collective action remains the extent to which they feel identified with a protesting group (i.e., social identification). Although the link between social identification and collective action is well-established, current evidence relies mostly on self-report data. To tackle this issue, we combined real-life protest counts in the US (2017-2020) with online search data (Google Trends) for pronouns indicating a ‘group’ mindset (first-person plural pronouns; e.g. ‘we’, ‘us’). Time series analyses indicated that weekly fluctuations in searches ( $N = 164$ ) predict both protest and protester counts over time. Confirmatory mixed-models then showed that a 1% increase in pronoun searches was linked with +13.67% protests 95%CI[4.02,23.32] and +47.45% protesters 95%CI[26.54, 68.36] the following week. These original results have important implications for the ecological study and quantification of collective action dynamics in psychology.

**Keywords:** social identification, first person plural pronouns, collective action, protests, search volume data

## INTRODUCTION

Collective action in the form of protests and demonstrations is a powerful driver of social and political change. Psychological science suggests that the main motivational factor mobilizing individuals into collective action is the extent to which they feel identified with a protesting group (i.e., social identification). However, current evidence relies on self-report data and mostly cross-sectional designs, so this link was yet to be demonstrated using real-world outcomes. In this paper, we describe a simple – yet robust – procedure that circumvents these issues to study the social psychological dynamics of protests *in vivo*. For the first time, we demonstrate that population-level social identification predicts real protest occurrence ( $N = 23,481$ ) over time by leveraging search volume data on First Person Pronoun use (e.g., ‘I’, ‘me’, ‘we’, ‘us’) in the US.

### **The psychology of collective action**

What motivates individuals to engage in collective action? Social psychological research has historically highlighted three main factors to explain mobilization, namely perceived injustice (Kawakami & Dion, 1995), group-efficacy towards collective action (Bandura, 2000) and a sense of belonging to a common ingroup (i.e., social identity; Klandermans, 2001). Building on these factors, the modern social psychology of collective action now relies on a unified theoretical model: the Social-identity Model of Collective Action, (SIMCA; Van Zomeren, Spears, Fischer, & Leach, 2004; see Van Zomeren, Postmes, & Spears, 2008). In a nutshell, the SIMCA predicts that emotions linked with perceptions of injustice (e.g., anger), perceived efficacy of collective action in bringing about social change and a sense of identification with the protesting group additively increase the likelihood that one will engage into collective action.

So far, the SIMCA’s predictive power has received extensive empirical corroboration upon a wide range of issues (Mahfud & Adam-Troian, 2020; Tausch et al., 2011; 2015; Van den Bos, 2018; van Zomeren, 2013). The model has also been extended to incorporate moral convictions and group-based emotions such as fear, contempt and anger to explain recourse to both normative (e.g., legal demonstrations) and non-normative (e.g., riots) collective actions (Ayanian & Tausch, 2016; Shuman, Cohen-Chen, Hirsch-Hoefler, Halperin, 2016; van

Zomeran, Postmes, & Spears, 2013). Still, experimental studies point at a unique causal role of social identification in shaping mobilization intentions (Bos et al., 2020; Shi, Hao, Saeri, & Cui, 2015; Subašić, Schmitt, & Reynolds, 2011; Thomas, Mavor, & McGarty, 2012; Ufkes, Calcagno, Glasford, & Dovidio, 2016). As an illustration, recent longitudinal investigations downplayed the temporal role of efficacy and showed that identification with a deprived group predicts both subsequent perceptions of injustice and intentions to engage in collective action to defend that group's rights (Thomas, Zubielevitch, Sibley, & Osborne, 2019). In other words, identification dynamically 'feeds' other predictors from the SIMCA (see Hasan-Aslih et al., 2020).

Yet, current evidence for the causal role of social identity in driving collective actions remains incomplete for three main reasons. First, studies are bounded to self-reports of behavioural intentions (or past behaviour), which – although theoretically informative – tell us little in terms of actual behaviour in context (Baumeister, Vohs, & Funder, 2007). Second, collective action unfolds over time. Though many researchers acknowledge the dynamic nature of collective action (e.g., Becker, & Tausch, 2015), even the few longitudinal tests of the SIMCA include only a handful points in time. Finally, protests and demonstrations are *collective* phenomena by definition, which questions the exclusive reliance of collective action researchers on individual-level data. We thus propose to overcome these issues by (1) focusing on *real* collective action (2) over extensive periods of *time* (3) using ecological measures of *population*-level social identification.

### **An ecological approach to protests**

To do so, one would need accessible data on collective actions that occurred in time. Incidentally, this type of data can now be retrieved from newly available databases on protest counts over time in the US (with daily updates, see Fisher et al., 2019). Therefore, the main challenge remains to find an ecological measure of social identification that could be paired with these protest counts. One possibility consists in gathering word counts from online sources (e.g., social media, blogs posts, online searches, see Al-Mosaiwi & Johnstone, 2018; Mooijman, Hoover, Lin, Ji, & Dehghani, 2018) that reflect aggregated psychological and emotional states (Pennebaker & King, 1999).

More specifically, research suggests that one type of words reflects psychological processes of group identification: First Person Pronouns (FPP, Pennebaker, Mehl, & Niederhoffer, 2003). In fact, pronouns are a crucial means of highlighting entities at the center of a narrative (Gordon, Grosz, & Gilliom, 1993). In the case of human entities (agents), pronouns are powerful indicators of their centrality in the discourse (Brown & Levinson, 1987). Accordingly, the use of FPP can indicate the psychological salience of a group entity versus the individual (e.g., 'I' vs. 'us'). For instance, research on dyadic relationships found that use of FPP ('we' instead of 'he/she') predicts increased relationship commitment, satisfaction and lower divorce rates (Kay & Fitzsimmons, 2004; Seider, Hirschberger, Nelson, & Levenson, 2009; Simmons, Gordon, & Chambless, 2005).

In the context of groups, research indicates that plural-FPP use also increases after experiencing social stressors (e.g., terror attacks, personal small-scale tragedies; Cohn, Mehl, & Pennebaker, 2004; Gortner & Pennebaker, 2003), reflecting a heightened sense of group belonging and cohesiveness. Studies using experimental group tasks revealed that use of plural-FPP during a preliminary group discussion predicts subsequent group performance and cohesiveness (see Gonzales, Hancock, & Pennebaker, 2010). Conversely, analyses of political speeches in the UK and Australia show that successful candidates use – on average – more plural-FPP than their opponents (Steffens & Haslam, 2013). Relatedly, experimental psychologists routinely use FPP circling tasks to activate interdependent-independent self-concepts (for a summary see Oyserman, 2016) while macro-level analyses have successfully used FPP from cultural productions (i.e., books) as a proxy for individualistic-collectivistic values in American and Chinese cultures (Hamamura, & Xu, 2015; Twenge, Campbell, & Gentile, 2012). Therefore, use of FPP seems to be a reliable linguistic marker for social identification, albeit devoid of reference to specific groups.

### **The Present Study**

Here, we decided to perform an ecological test of the SIMCA by examining the dynamics of collective action at the population-level. In line with the SIMCA, we wished to test whether FPP use from online data (as a proxy for social identification) could predict

*subsequent* changes in collective action. Indeed, if social identification leads individuals to mobilize in protests, this should be observable from population-level traces. Increased salience of group identities in the population should be reflected in online traces of characteristic markers such as FPP. To the extent that prevalence of online FPP-use indicates how much individuals define themselves as part of groups – regardless of the specifics – these should be linked with the magnitude of ongoing collective actions within that population. More specifically, we hypothesized that prevalence of plural-FPP (e.g., ‘we’, ‘us’) at any point in time would predict higher protest mobilization at subsequent points in time (H1). The reverse should be true for single-FPP prevalence (H2).

## METHODS

To test our hypotheses we gathered observations aggregated at the population level and spanning as many time points as possible to enable fine-grained causal tests using statistical tools from time series analyses (for an introduction to time series analyses, see Lütkepohl, 2013). All data, scripts and analyses supporting this section can be openly accessed at [https://osf.io/q36vy/?view\\_only=8e001ef3df5f4af692523dace4e582ae](https://osf.io/q36vy/?view_only=8e001ef3df5f4af692523dace4e582ae)

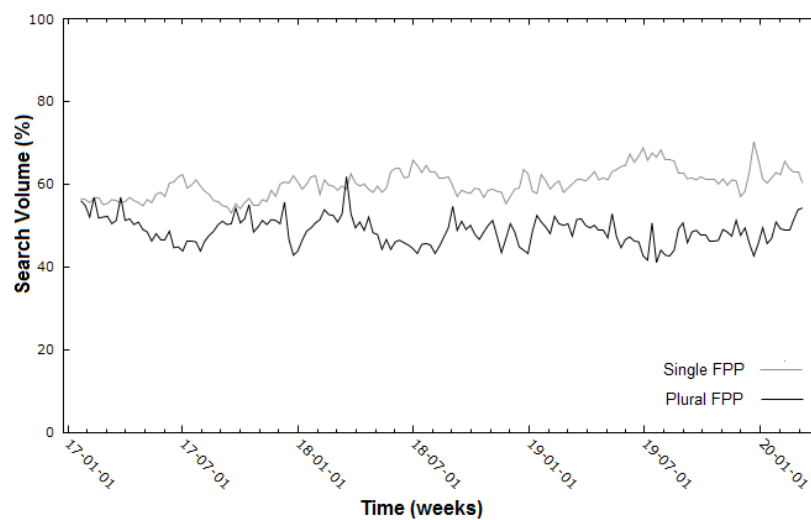
### Data

#### *First person Pronouns*

To compute our measures of FPP, we decided to use search volume data from Google Trends. In short, Google Trends provides researchers with the frequency at which a specific search term is typed in Google’s search engine (normalized by maximum frequency of term searched) within specified languages and geographical areas (Google, 2017). Following the FPP measures proposed by Twenge et al. (2012) and Hamamura and Xu (2015), we decided to extract search volume indices for the words ‘I’, ‘me’, ‘mine’, ‘myself’ to create a single averaged FPP time series and for the words ‘us’, ‘we’, ‘our’, ‘ours’, ‘ourselves’ for a plural FPP time series aggregated at the US-level from January the 15<sup>th</sup> of 2017 to March the 8<sup>th</sup> of 2020 ( $N_{weeks} = 164$ ; figure 1).

Although it is possible to use textual data from social media posts and online blog articles, search volume indices present four main advantages. First, these data are completely anonymized, publicly available and can be easily retrieved (at

<https://trends.google.com/trends/?geo=US>). Second, data from search engines can tap into a broader range of the population than social media users. Third, search volume data is constantly updated at a fixed frequency (weekly) which facilitates aggregation and interpretation. Finally, a growing literature has consistently established the usefulness of search volume measures to predict various offline behaviours ranging from stock markets fluctuations (Preis, Reith, & Stanley, 2010; Preis, Moat & Stanley, 2013) to political mobilization (DiGrazia, 2017; Oullier & Marzouki, 2015) and state-level suicide rates (Adam-Troian & Arciszewski, 2020; see Jun, Yoo, & Choi, 2018 for a review).



*Figure 1. Chronogram of weekly search volumes for Single and Plural First Person Pronouns (FPP) in the US from 01-17-17 to 03-08-20. X axis represents time in weeks, with dates in yy-mm-dd format.*

### *Protest mobilization*

We extracted protest mobilization data from the recently created Count Love database (openly accessible at <https://countlove.org/faq.html>). This database covers protests in the US starting from the 15<sup>th</sup> January of 2017, by using machine learning tools that identify reports of collective actions from online articles, press coverage and other available sources. Briefly, the database is assembled automatically and routinely updated from constantly extended URLs of organization websites, with the help of natural language processing techniques, cross-checked by human agents. The events are then classified and only comprise protests, excluding gatherings (e.g., fundraising events) and rituals (e.g., commemorative celebrations). The database also includes estimates of attendance (number of participants), which are

approximate (e.g., a report of ‘dozen’ is counted as ‘10’). More details about the inner workings of Count Love can be found in Fisher et al. (2019). We extracted the data on March 8<sup>th</sup> of 2020, yielding  $N = 23,481$  protests including 12,641,229 demonstrators. From these, we computed two protest mobilization time series, one for protest counts and the other for participant counts (normalized per 1,000 inhabitants). These were aggregated over  $N = 164$  weeks to match the FPP search volume series (figure 2).

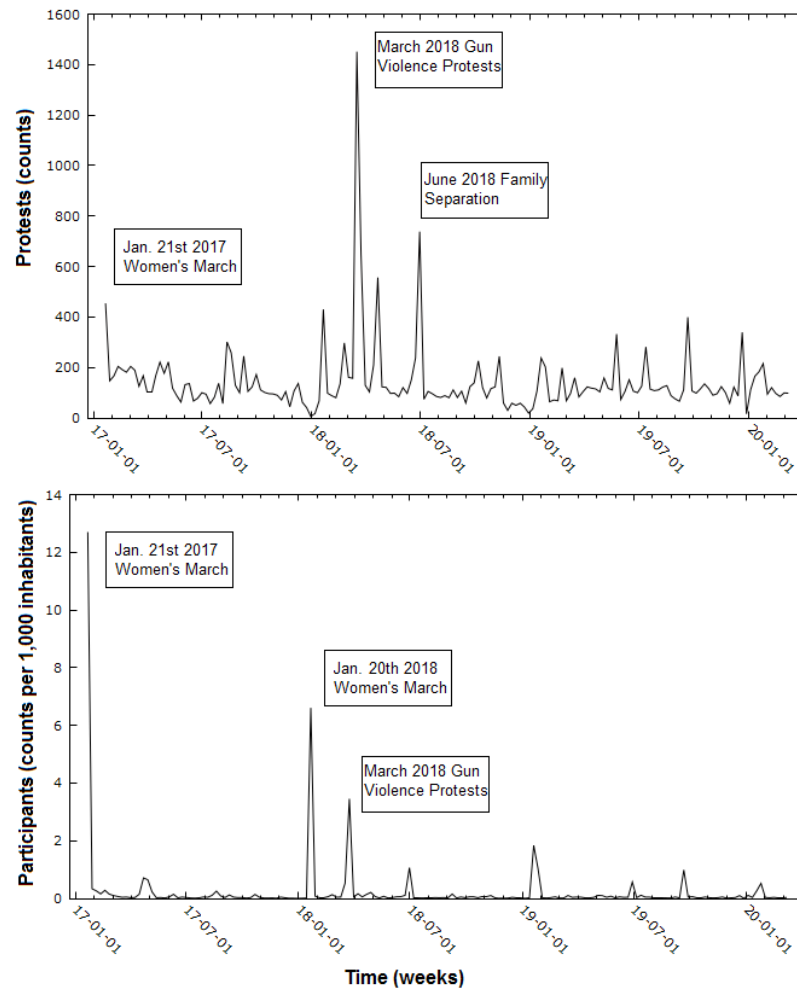


Figure 2. Chronogram of weekly protest counts (above) and participants per 1,000 inhabitants (below) in the US from 01-17-17 to 03-08-20. X axis represents time in weeks, with dates in yy-mm-dd format. Comments indicate significant events at play behind substantial protest count increases.

### Analytical strategy

To assess the relationship between FPP and protest mobilization time series, we implemented a three-step approach using complementary statistical frameworks.

Because our data are time series, simple regression models would be biased due to time-induced dependence between observations. One way to circumvent this issue is to use sets of OLS regressions including past values of  $Y$  as predictors to assess links between time series (i.e., so-called ‘vector autoregressive models’; Sims, 1980). Implementing these models then allows for testing a key statistical property: Granger causality. Briefly, an independent variable  $X$  ‘Granger Causes’ a dependent variable  $Y$  if  $Y$  can be predicted (even partially) by  $X$  over and above  $Y$ ’s own past (lagged) values. Granger causality can be tested by regressing  $Y$  at  $t + 1$  on both  $Y$  and  $X$  and interpreting the results (Granger, 1969). One issue with these simple tests of Granger causality is that they need stationary series, often obtained through seasonal adjustment and time-differencing, which tends to produce spurious associations and to render parameter estimates difficult to interpret. Also, these procedures bear risks associated with misidentifying time series’ order of integration (Mavrotas and Kelly, 2001), resulting in erroneous conclusions. To bypass these issues, we tested Granger causality between our time series using the Toda-Yamamoto procedure (Toda & Yamamoto, 1995), which can be carried out directly on non-stationary variables (for more details and a previous example of how this procedure is applied to study collective action in psychology, see Mooijman et al., 2018).

Testing for Granger causality implies a number of analytical decisions that could affect results from the critical Wald tests used in the Toda-Yamamoto procedure. A confirmatory analysis (i.e., ‘multiverse analysis’; Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016) would therefore be carried out to cross-check results over a range of plausible alternative model specifications to ensure reliable causal inference. After robust evidence for Granger causality is obtained, reliable estimates will be obtained through generalized mixed effects models. This will be done because the Yamamoto-Toda procedure can only test a specific statistical property (i.e., absence of Granger causality) but does not necessarily provide for accurate model parameters. Time series related tests were run using the software *GRETL* (v.2020a; Mixon, 2010) and generalized-mixed models were computed on *R* (v.4.0; R Core Team, 2013, ‘*lme4*’ package).



## *Power Analyses*

Because Granger causality tests rely on sets of OLS estimates, power analyses can be approximated through estimates for multiple regression. Analyses were conducted in GPower (v3.1, Faul, Erdfelder, Buchner, & Lang, 2009). Typically, the link between social identification and collective action intentions is large (ranging from  $r = .30$  to  $r > .60$ ; e.g., Mahfud & Adam-Troian, 2020). This is reflected by meta-analytical evidence from Van Zomeren et al. (2008) which pointed at an estimated  $r = .38$  or  $d = .82$ . Yet, because our methods differ drastically from previous studies, we decided to use a more conservative estimate and to anticipate an effect size of  $d = .42$  (median size of social psychological effects; see Richard, Bond, & Stokes-Zoota, 2003), which translates into a  $f^2$  value of .044. To achieve 80% power to detect a regression coefficient of  $f^2 = .044$  at  $\alpha = .05$  (two-tailed), recommended  $N = 181$ . With a  $N = 164$ , our tests would thus be slightly underpowered (76%; under conservative assumptions). However, we lacked information for many of the critical parameter estimates needed to estimate *a priori* power in the specific case of estimating power for generalized linear models. Therefore, a reliable power estimate could not be obtained for that analysis. In any case, our sample size was conditioned to data availability and thus limited to a maximum  $N = 164$  when the study was carried out.

## **RESULTS**

### **Granger Causality Tests**

#### *Normality tests*

A pre-requisite for Granger causality tests using vector autoregressive models is to prevent severe deviations from normality, especially with regards to count data like our protest time series. Shapiro-Wilks tests showed departures from normality for Single FPP,  $W = 0.98$ ,  $p = .045$ , Plural FPP,  $W = 0.99$ ,  $p = .068$ , Protest counts,  $W = 0.51$ ,  $p < .001$  and Participant counts,  $W = 0.17$ ,  $p < .001$ . They were thus transformed using a logarithmic function.

#### *Determination of Toda-Yamamoto parameters*

Following the Toda-Yamamoto procedure, we determined the maximum order of integration  $p$  across our variables comparing estimates from augmented Dickey-Fuller

(Paparoitis & Politis, 2018) and Kwiatkowski-Phillips-Schmidt-Shin (Kwiatkowski, Phillips, Schmidt, & Shin, 1992) unit root tests. Augmented Dickey-Fuller tests indicated that the time series were all stationary at  $l = 0$  (all  $p$ s < .01). However, results from KPSS tests suggested an integration order at  $l = 0$  for log-transformed Protest counts  $LM = 0.14, p = .071, l = 1$  for Plural FPP,  $LM = 0.12, p = .092$  and Participant counts,  $LM = 0.14, p = .073$  and  $l = 2$  for Single FPP,  $LM = 0.12, p < .10$ . Therefore,  $m = \text{maximum } l = 2$ .

To determine the maximum number of lags  $m$ , we first inspected ACF and PACF plots as well as Q statistics for autocorrelations (Ljung-Box, 1978). These tended to indicate a converging  $m = 1$ . A multivariate vector autoregression lag selection procedure indicated a convergent fit for the lowest values of AIC, BIC and HQC at  $l = 1$ , thus  $m = 1$ .

### *Critical tests*

In line with the Toda-Yamamoto procedure, we set up a multivariate vector autoregression model containing four equations (one per log-transformed time series) including lagged variables for all time series at  $d_{\max} = p + m = 1 + 2 = 3$  lags each. A linear coefficient for time was also included since ACF plots for Single FPP suggested an upward trend over time. The models indicated that – indeed – Single FPP increased with time,  $b = .003, 95\%CI[.001,.005], p = .023$ . As per the Toda-Yamamoto procedure's final step, we performed a series of Wald  $\chi^2$  tests to probe the null hypothesis of Granger *non-causality*.

When predicting log Protest counts, Wald tests indicated that the slopes for both Single FPP,  $b = -.22, 95\%CI[-.45,-.01]$  and Plural FPP,  $b = .27, 95\%CI[.09,.44]$  differed substantially from zero, with respectively  $\chi^2(1) = 3.32, p = .068$  and  $\chi^2(1) = 8.65, p = .003$ . For log Participant counts, Plural FPP remained a predictor,  $b = .26, 95\%CI[.08,.44], \chi^2(1) = 9.13, p = .003$ , while the weak effect of Single FPP vanished,  $b = -.16, 95\%CI[-.38,.06], \chi^2(1) = 2.06, p = .15$ . Finally, tests suggested no noticeable Granger causality from log Participant counts to both FPP variables,  $\chi^2(1) = .04, p = .83$  for Plural FPP and  $\chi^2(1) = .97, p = .32$  for Single FPP. Protest counts too did not affect subsequent Plural FPP counts,  $\chi^2(1) = .22, p = .64$  and Single FPP,  $\chi^2(1) = 2.56, p = .11$ .

This overall pattern of results corroborated H1 and provided mixed evidence regarding H2. They also suggested strong unilateral Granger causality from Plural FPP upon both measures of protest mobilization, as theoretically expected.

### **Confirmatory Tests**

#### *Multiverse analyses*

Granger causality tests rely on model specifications that may vary according to analysts' choices. Although the validated procedure from Toda & Yamamoto (1995) constrains the sequence of tests we performed, slightly different choices in the exact operationalization of the procedure might have affected our results. To gain more robust evidence to test our hypotheses, we carried out confirmatory tests relying on a 'multiverse analysis' (see Steegen et al., 2016 for more details). We did so by re-running the Toda-Yamamoto procedure under plausible alternative choices and then reporting  $p$ -values from the Wald tests in matrices with codes for each methodological alternative.

Regarding normalization, we wished to compare our results (all variables normalized, coded 'N<sub>2</sub>'), to normalization of only the severely non-normal count time series ('N<sub>1</sub>') and to results if no normalization were applied ('N<sub>0</sub>'). When determining,  $m$ , KPSS tests for Single FPP at  $l = 1$  indicated a  $p = .045$ , which could be interpreted as indicating stationarity (i.e., close to the cutoff  $p = .05$ ) so we also ran alternative tests with  $m = 1$ ,  $d_{\max} = 2$  ('C<sub>0</sub>' vs. 'C<sub>1</sub>' for the current  $m = 2$ ;  $d_{\max} = 3$ ). PACF and ACF indicated the presence of quarterly seasonality for most series and a time trend for Single FPP, so we re-ran the tests using no-adjustment ('S<sub>0</sub>'), adjustment on time only (current 'S<sub>1</sub>') and time plus seasonal variation ('S<sub>2</sub>') using seasonal dummy covariates ( $N = 3, 4$  quarter contrasts minus 1). Finally, we specified multivariate models to keep the number of tests - thus type I errors - low ('M<sub>1</sub>'), but we wanted to run analyses using univariate models, testing causality among pair of time series separately ('M<sub>0</sub>'). This left us with  $N = 3 \times 2 \times 3 \times 2 = 36$  combinations of tests for each causal pair ( $N = 8$ ; Single FPP ↔ Protest counts, Single FPP ↔ Participant counts, Plural FPP ↔ Protest counts, Plural FPP ↔ Participant counts), yielding a total of  $36 \times 8 = 288$  tests to perform.

Results from the multiverse analysis corroborated what was found in our initial tests (figure 3). There was no substantial evidence for causal effects from our collective action measures to FPP variables. Single FPP did not seem to exert noticeable influence on collective action measures in most tests ( $p < .05$  in approximately 17% of models). Plural FPP emerged as a robust predictor of both collective action measures, with 31.7% ( $N = 13$ ) of tests outcomes below  $p < .01$  for Protest counts and 41.5% ( $N = 17$ ) for Participant counts. Subsequent P-curve analyses (Simonsohn, Nelson, & Simmons, 2013; conducted online at <http://www.p-curve.com/app4/>) confirmed strong evidential value for the effect of Plural FPP on both Protests,  $Z = -5.4$ ,  $p < .001$  and Participants counts  $Z = -6.3$ ,  $p < .001$  (figure 4).

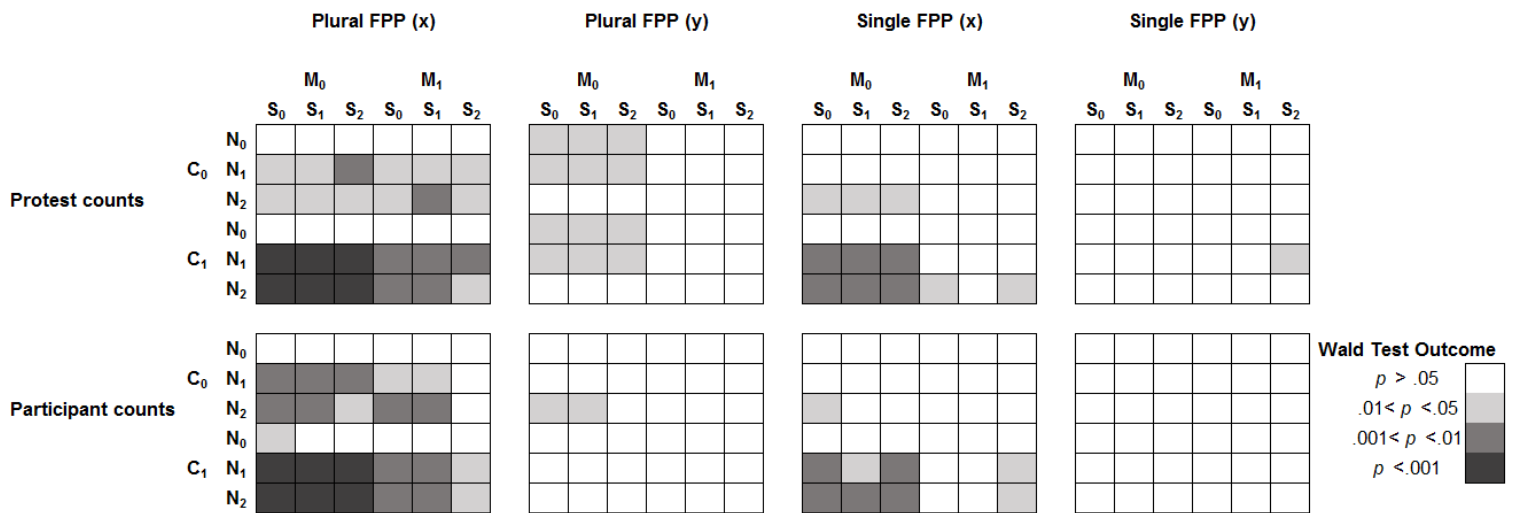


Figure 3. Matrices for the multiverse analysis of causality tests with FPP time series specified as independent (x) or dependent (y) variables ( $N = 288$ ). Each cell represents one model specification according to the coding scheme described in the ‘Confirmatory tests’ section. Shades of grey indicate the significance level of each tests’ outcomes.

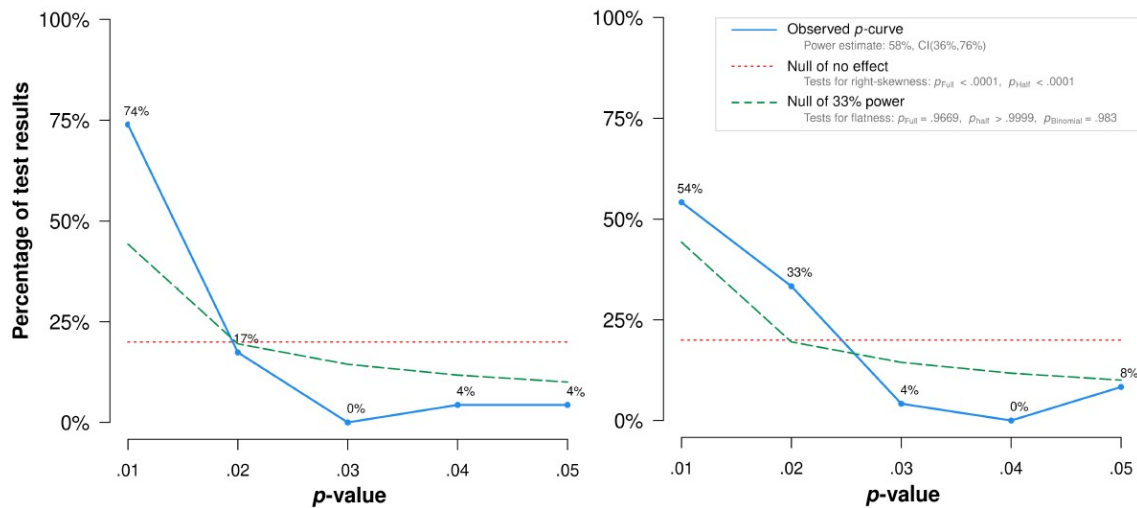


Figure 4. *P*-curve plots for Wald tests of causality tests from Plural FPP to Participant (left,  $N = 36$ ) and Protest counts (right;  $N = 36$ ).

### Mixed-effects analyses

After providing solid evidence for a causal link between Plural FPP and collective action measures at  $t+1$ , we still needed accurate parameter estimates, which could not be obtained using normalized time series (O’Hara & Kotze, 2010). Also, the multiverse analysis confirmed that Toda-Yamamoto tests – because they are sensitive to departures from normality – were not able to detect causality using non-log transformed time series (see ‘ $N_0$ ’ lines in figure 4). Therefore, we needed tests that would provide for accurate parameter estimates and a final cross check on our previous analyses without log-transformation this time. To do so, we decided to model the effects of lagged ( $t-1$ ) FPP using non-linear functions to model both Protest and Participant counts with random factors to account for the multilevel structure of our time series data (weeks within months nested in quarters within years; Schielzeth, & Nakagawa, 2013).

We first Z-transformed our predictors to avoid scaling issues when estimating the models. Then, a series of generalized mixed models using a Poisson function were computed including a parameter for time and values of the dependent variables at  $t-1$  to control for potential confounds. Results from these models yielded significant parameters for all variables ( $ps < .001$ ) but were due to biases in models’ estimates of standard error. Indeed, analyses of residual deviance showed that both models suffered from severe overdispersion, with

deviance to *df* ratios of 42.5 Protest Counts and of 40619.13 for Participant counts (when these should be as close to 1 as possible; Hinde & Demétrio, 1998).

To solve these overdispersion issues, we re-ran the same models using a Negative binomial function. Overdispersion was removed, with ratios of now respectively .96 and 1.01 for Protest and Participant counts. The better fit of these models was evident from comparison test results for Protest  $\chi^2(1) = 6050.5, p < .001$  and Participant counts,  $\chi^2(1) = 6294941, p < .001$ . Confirming our prior results, these highlighted a significant contribution of *only* Plural FPP<sub>*t-1*</sub> in predicting Protest,  $b = .15, 95\%CI[.01, .29], p = .039$  and Participant counts  $b = .63, 95\%CI[.34, .91], p < .001$ . Because the other variables in the models were not substantial predictors of collective action measures, we dropped them to compute final models for proper parameter estimates according to the following:

$$(1) M_{\text{Protest}} = \text{Protest counts} \sim \text{Plural FPP}_{t-1} + (1\text{month}/\text{quarter}/\text{year})$$

$$(2) M_{\text{Participant}} = \text{Participant counts} \sim \text{Plural FPP}_{t-1} + (1\text{month}/\text{quarter}/\text{year})$$

Comparison tests indicated that these reduced models performed as well as previous ones,  $\chi^2(3) = .29, p = .96$  for Protest counts and  $\chi^2(3) = 3.84, p = .28$  for Participant counts. As can be seen in table 1, Plural FPP is indeed a positive predictor of subsequent Protest,  $b = .17, 95\%CI[.05, .29], p = .004$  and Participant counts  $b = .59, 95\%CI[.33, .85], p < .001$ . As per conversion of these coefficients into incidence ratios – I.R. =  $\exp(\beta \times 100)$  – an increase in one standard deviation (3.38%) of online searches containing Plural FPP predicted 46.21%, more Protests 95%CI[13.59, 78.83] and 160.38%, more demonstrators 95%CI[89.70, 231.05] the subsequent week. Assuming that the standard deviation from our sample accurately reflects the ‘population’ standard deviation of Plural FPP search volume in the US, a 1% increase in Plural FPP searches would therefore predict 13.67% more Protests 95%CI[4.02, 23.32], and 47.45% more demonstrators 95%CI[26.54, 68.36] the following week. This provided strong evidence for H1 and allowed us to confidently reject H2.

Table 1. Summary of generalized negative binomial mixed *models of Plural FPP's effects at t-1* ( $M_1 = \text{Protest counts}$ ;  $M_2 = \text{Participant counts}$ ;  $N = 163$ ).

	LL	$R^2_m$	$R^2_c$	Estimate (SE)	95% CI	Z	P
$M_1$	-927.8	.06	.25				
<i>Intercept</i>				4.86(.07)	[4.73,4.99]	73.71	< .001
<i>Plural FPP<sub>t-1</sub></i>				.17(.06)	[.05,.29]	2.88	.004
$M_2$	-1834.0	.12	.46				
<i>Intercept</i>				10.01(.25)	[9.52,10.50]	40.02	< .001
<i>Plural FPP<sub>t-1</sub></i>				.59(.13)	[.33,.85]	4.46	< .001

Note. LL = Log Likelihood,  $R^2_m$  = Marginal *R*-squared,  $R^2_c$  = Conditional *R*-squared, Estimate = standardized coefficients, SE = Standard Error.

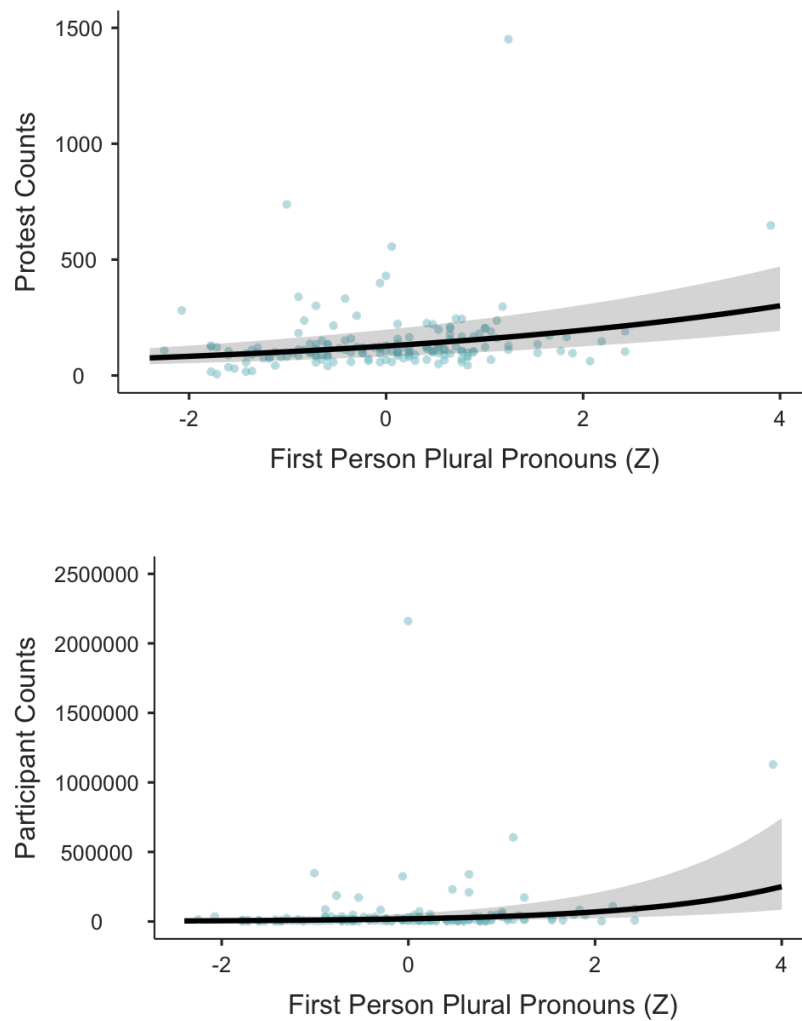


Figure 5. Scatterplots of the negative binomial relationship between FPP Plural at *t-1* (*x*-axis), Protest counts (*y*-axis, top) and Participant counts (*y*-axis, bottom;  $N = 163$ ).

## DISCUSSION

Can we predict offline social mobilization from online linguistic traces of ‘groupthink’? In this study, we combined three different analytical approaches to test whether social identification – as measured by FPP use from search volume – could predict *real life* social mobilization in the US. In a first sequence relying on time series analyses, we obtained evidence that Plural FPP (e.g., ‘we’, ‘us’) positively predicted Protest and Participant counts the following week and that the reverse was not true. This unilateral causal pathway from Plural FPP to collective action outcomes was corroborated by a second round of tests combining a multiverse analysis and follow-up *P*-curve analyses. Confirmatory analyses using generalized mixed models allowed us to conclude that 1% increases in Plural FPP at any given week in time predicted 13.67% more protests and 47.45% more participants the subsequent week. Finally, our analyses downplayed the role of Single FPP (e.g., ‘I’, ‘me’) in this process. What does this mean for the psychological science of collective action?

Theoretically, the present study constitutes a first test of the SIMCA (van Zomeren et al., 2008) that relied exclusively on ecological behavioural indicators (i.e., search volume data; number of protests and demonstrators). This test confirmed the role of social identification as a key predictor of collective action and extended it to actual behaviour. Interestingly, the observed increase in Single FPP use over time provided an incidental replication of Santos, Varnum and Grossmann (2017) which highlighted global increases in individualism over the last five decades. Given the current replicability crisis in psychology (Pashler, & Wagenmakers, 2012), and especially social psychology (Stroebe, 2016), demonstrating the external validity of research findings from laboratory and self-report studies is crucial. In fact, our analyses were able to quantify meaningful effect sizes (behavioural frequencies) and to identify causal pathways across *time*. The fine-grained results allowed by time series data could be tapped into for making more specific predictions, in a field that suffers from concerning theoretical under-specification (Muthukrishna & Henrich, 2019). More importantly maybe, our study highlights how psychological researchers can study the *in vivo* dynamics of collective action in a way that circumvents traditional methodological limitations (Fisher et al., 2019).



In line with this, the present study opens new possibilities for methodological and applied purposes. For instance, future studies could leverage our approach to set up social mobilization forecasts, using time series analyses and the opportunity of *real-time* data update from search volume (weekly) for monitoring purposes. Like epidemiological models of pandemic dynamics, tools like these could be useful for policy makers and activists, since it is now possible to predict the likelihood that social movements succeed in bringing about political change using measures of ‘dissent momentum’ (frequency of protests x number of attendees; see Chenoweth, & Belgioioso, 2019). Moreover, dynamic models of collective behaviour sometimes yield inaccurate predictions for lack of theoretical inputs from psychological theories (i.e., social identity theory; see Drury, 2020; Templeton, Drury, & Philippides, 2015). Therefore, combining search volume indices of psychological states (social identification, emotions) to ‘feed’ models of protest-riot contagion (e.g., Bonnasse-Gahot, Berestycki, Depuiset, Gordon, Roché, Rodriguez, & Nadal, 2018) could prove useful in predicting collective action dynamics.

Before concluding, some caveats that limit the scope of our results need to be acknowledged. As our analyses highlighted, Single FPP did not substantially predict collective action outcomes – despite displaying some theoretically expected negative links. This may indicate a measurement asymmetry when using Single FPP, where Plural FPP could tap into a form of ‘groupthink’ and Single FPP would not reflect ‘non-groupthink’ (see Steffens & Haslam, 2013). This null-result might also be due to statistical issues (e.g., hard-to-remove stochastic components that introduce noise in the measure) or specific cultural psychological dynamics. For instance, one could argue the prevalence of individualistic thinking is not discriminant enough in an individualistic cultural setting like the US, but that Single FPP could be more (negatively) predictive of collective action in collectivistic cultural contexts. In any case, these hypotheses warrant further investigation. Second, these results come from data collected in the US, and should not be generalized beyond the US population at this point (see Henrich, Heine, & Norenzayan, 2010). Likewise, our study highlights effects pertaining to non-violent, legal collective action and might not necessarily hold in the case of non-legal or violent action (e.g., riots). Finally, these were obtained over the past four years

only, under a specific administration (Republican, President Trump) which warrants caution when generalizing the findings temporally. In line with this, we also emphasize that our results were obtained using a weekly frequency for data-points. In sum, future confirmatory studies should be carried out using data generated from other cultural contexts, for various types of collective action with daily or hourly frequencies if possible (e.g., Mooijman et al., 2019) and spanning longer time periods. Within the boundaries of the above mentioned limitations, we are confident that the present study opens exciting new avenues for the psychological study of collective action and social mobilization.

### REFERENCES

- Adam-Troian, J., & Arciszewski, T. (2020). Absolutist Words From Search Volume Data Predict State-Level Suicide Rates in the United States. *Clinical Psychological Science*, 2167702620916925.
- Al-Mosaiwi, M., & Johnstone, T. (2018). In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clinical Psychological Science*, 6(4), 529-542.
- Ayanian, A. H., & Tausch, N. (2016). How risk perception shapes collective action intentions in repressive contexts: A study of Egyptian activists during the 2013 post-coup uprising. *British Journal of Social Psychology*, 55(4), 700-721.
- Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current directions in psychological science*, 9(3), 75-78.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?. *Perspectives on Psychological Science*, 2(4), 396-403.
- Becker, J. C., & Tausch, N. (2015). A dynamic model of engagement in normative and non-normative collective action: Psychological antecedents, consequences, and barriers. *European Review of Social Psychology*, 26(1), 43-92.

- Bonnasse-Gahot, L., Berestycki, H., Depuiset, M. A., Gordon, M. B., Roché, S., Rodriguez, N., & Nadal, J. P. (2018). Epidemiological modelling of the 2005 French riots: a spreading wave and the role of contagion. *Scientific reports*, 8(1), 1-20.
- Brown, P., Levinson, S. C., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge university press.
- Chenoweth, E., & Belgioioso, M. (2019). The physics of dissent and the effects of movement momentum. *Nature human behaviour*, 3(10), 1088-1095.
- Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological science*, 15(10), 687-693.
- DiGrazia, J. (2017). Using internet search data to produce state-level measures: The case of tea party mobilization. *Sociological Methods & Research*, 46(4), 898-925.
- Drury, J., Stott, C., Ball, R., Reicher, S., Neville, F., Bell, L., ... & Ryan, C. (2020). A social identity model of riot diffusion: From injustice to empowerment in the 2011 London riots. *European journal of social psychology*, 50(3), 646-661.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, 41(4), 1149-1160.
- Fisher, D. R., Andrews, K. T., Caren, N., Chenoweth, E., Heaney, M. T., Leung, T., ... & Pressman, J. (2019). The science of contemporary street protest: New efforts in the United States. *Science advances*, 5(10), eaaw5461.
- Fitzsimmons, G. M., & Kay, A. C. (2004). Language and interpersonal cognition: Causal effects of variations in pronoun usage on perceptions of closeness. *Personality and Social Psychology Bulletin*, 30, 547-557.
- Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language style matching as a predictor of social dynamics in small groups. *Communication Research*, 37(1), 3-19.
- Google (2017). *Google Trends*. Retrieved from <https://trends.google.com/trends>

- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive science*, *17*(3), 311-347.
- Gortner, E. M., & Pennebaker, J. W. (2003). The archival anatomy of a disaster: Media coverage and community-wide health effects of the Texas A&M bonfire tragedy. *Journal of Social and Clinical Psychology*, *22*(5), 580-603.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, 424-438.
- Hamamura, T., & Xu, Y. (2015). Changes in Chinese culture as examined through changes in personal pronoun usage. *Journal of Cross-Cultural Psychology*, *46*(7), 930-941.
- Hasan-Aslih, S., Shuman, E., Goldenberg, A., Pliskin, R., van Zomeren, M., & Halperin, E. (2020). The Quest for Hope: Disadvantaged Group Members Can Fulfill Their Desire to Feel Hope, but Only When They Believe in Their Power. *Social Psychological and Personality Science*, 1948550619898321.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29-29.
- Hinde, J., & Demétrio, C. G. (1998). Overdispersion: models and estimation. *Computational Statistics and Data Analysis*, *27*(2), 151-170.
- Jun, S. P., Yoo, H. S., & Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological forecasting and social change*, *130*, 69-87.
- Kawakami, K., & Dion, K. L. (1995). Social identity and affect as determinants of collective action: Toward an integration of relative deprivation and social identity theories. *Theory & Psychology*, *5*(4), 551-577.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of econometrics*, *54*(1-3), 159-178.

- Lütkepohl, H. (2013). *Introduction to multiple time series analysis*. Springer Science & Business Media.
- Mahfud, Y., & Adam-Troian, J. (2019). “Macron demission!”: Loss of significance generates violent extremism for the Yellow Vests through feelings of anomia. *Group Processes & Intergroup Relations*, 1368430219880954.
- Marzouki, Y., & Oullier, O. (2014). Internet search volume as a proxy approach to the Virtual Collective Consciousness. In A. D. Cheok (Ed.), *Hyperconnectivity and the future of Internet communication* (p. 30-58). Balti, Moldova: Lambert.
- Mavrotas, G., & Kelly, R. (2001). Old wine in new bottles: Testing causality between savings and growth. *The Manchester School*, 69, 97-105.
- Mixon, J. (2010). GRETL: an econometrics package for teaching and research. *Managerial Finance*.
- Mooijman, M., Hoover, J., Lin, Y., Ji, H., & Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. *Nature human behaviour*, 2(6), 389-396.
- Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, 3(3), 221-229.
- O'Hara, R., & Kotze, J. (2010). Do not log-transform count data. *Nature Precedings*, 1-1.
- Oyserman, D. (2016). What does a priming perspective reveal about culture: Culture-as-situated cognition. *Current Opinion in Psychology*, 12, 94-99.
- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence?. *Perspectives on Psychological Science*, 7(6), 528-530.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of personality and social psychology*, 77(6), 1296.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1), 547-577.

- Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific reports*, 3, 1684.
- Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1933), 5707-5719.
- Team, R. C. (2013). R: A language and environment for statistical computing.
- Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331-363.
- Santos, H. C., Varnum, M. E., & Grossmann, I. (2017). Global increases in individualism. *Psychological science*, 28(9), 1228-1239.
- Schielezeth, H., & Nakagawa, S. (2013). Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution*, 4(1), 14-24.
- Seider, B. H., Hirschberger, G., Nelson, K. L., & Levenson, R. W. (2009). We can work it out: Age differences in relational pronouns, physiology, and behavior in marital conflict. *Psychology and aging*, 24(3), 604.
- Shi, J., Hao, Z., Saeri, A. K., & Cui, L. (2015). The dual-pathway model of collective action: Impacts of types of collective action and social identity. *Group Processes & Intergroup Relations*, 18(1), 45-65.
- Shuman, E., Cohen-Chen, S., Hirsch-Hoefler, S., & Halperin, E. (2016). Explaining normative versus nonnormative action: The role of implicit theories. *Political Psychology*, 37(6), 835-852.
- Simmons, R. A., Gordon, P. C., & Chambless, D. L. (2005). Pronouns in marital interaction: What do “you” and “I” say about marital health?. *Psychological science*, 16(12), 932-936.
- Simon, B., & Klandermans, B. (2001). Politicized collective identity: A social psychological analysis. *American psychologist*, 56(4), 319.

- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, 1-48.
- Steege, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702-712.
- Steffens, N. K., & Haslam, S. A. (2013). Power through 'us': Leaders' use of we-referencing language predicts election victory. *PloS one*, 8(10), e77952.
- Stroebe, W. (2016). Are most published social psychological findings false?. *Journal of Experimental Social Psychology*, 66, 134-144.
- Subašić, E., Schmitt, M. T., & Reynolds, K. J. (2011). Are we all in this together? Co-victimization, inclusive social identity and collective action in solidarity with the disadvantaged. *British Journal of Social Psychology*, 50(4), 707-725.
- Tausch, N., Becker, J. C., Spears, R., Christ, O., Saab, R., Singh, P., & Siddiqui, R. N. (2011). Explaining radical group behavior: Developing emotion and efficacy routes to normative and nonnormative collective action. *Journal of personality and social psychology*, 101(1), 129.
- Tausch, N., Saguy, T., & Bryson, J. (2015). How does intergroup contact affect social change? Its impact on collective action and individual mobility intentions among members of a disadvantaged group. *Journal of Social Issues*, 71(3), 536-553.
- Templeton, A., Drury, J., & Philippides, A. (2015). From mindless masses to small groups: conceptualizing collective behavior in crowd modeling. *Review of General Psychology*, 19(3), 215-229.
- Thomas, E. F., Mavor, K. I., & McGarty, C. (2012). Social identities facilitate and encapsulate action-relevant constructs: A test of the social identity model of collective action. *Group Processes & Intergroup Relations*, 15(1), 75-88.
- Thomas, E. F., Zubielevitch, E., Sibley, C. G., & Osborne, D. (2020). Testing the social identity model of collective action longitudinally and across structurally

- disadvantaged and advantaged groups. *Personality and Social Psychology Bulletin*, 46(6), 823-838.
- Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of econometrics*, 66(1-2), 225-250.
- Twenge, J. M., Campbell, W. K., & Gentile, B. (2012). Increases in individualistic words and phrases in American books, 1960–2008. *PloS one*, 7(7), e40181.
- Ufkes, E. G., Calcagno, J., Glasford, D. E., & Dovidio, J. F. (2016). Understanding how common ingroup identity undermines collective action among disadvantaged-group members. *Journal of experimental social psychology*, 63, 26-35.
- Van den Bos, K. (2018). *Why people radicalize: How unfairness judgments are used to fuel radical beliefs, extremist behaviors, and terrorism*. Oxford University Press.
- Van den Bos, K. (2020). Unfairness and radicalization. *Annual review of psychology*, 71, 563-588.
- Van Zomeren, M. (2013). Four core social–psychological motivations to undertake collective action. *Social and Personality Psychology Compass*, 7(6), 378-388.
- Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. *Psychological bulletin*, 134(4), 504.
- Van Zomeren, M., Postmes, T., & Spears, R. (2012). On conviction's collective consequences: Integrating moral conviction with the social identity model of collective action. *British Journal of Social Psychology*, 51(1), 52-71.