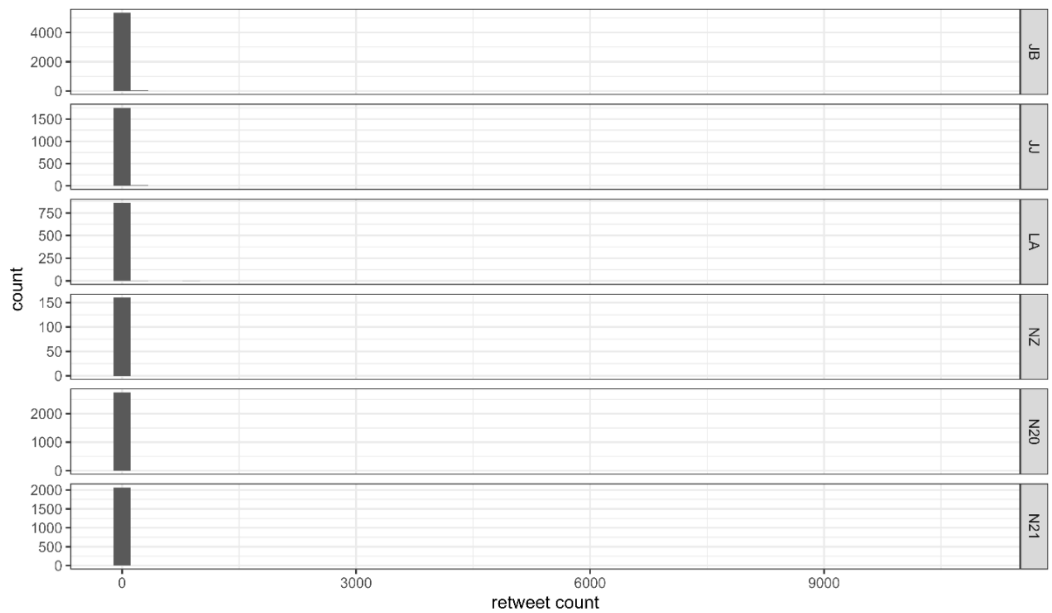
Supplementary Information

Lean-back and lean-forward online behaviors: the role of emotions in passive versus proactive information diffusion of social media content.

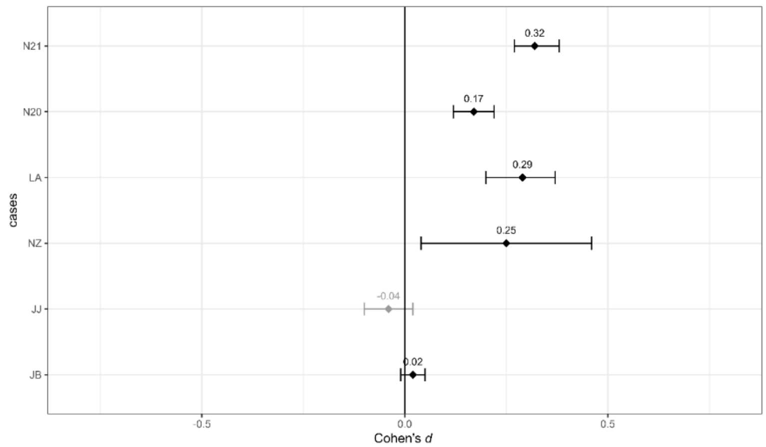
Section 1: Robustness check of comparisons

In our data collection, after eliminating duplicates, we classified tweets in lean-back and lean-forward behaviors based on the information obtained from the Twitter API, which makes an explicit reference if the tweet is a retweet (lean-back) or a quote (lean-forward). Quote tweets are tweets that refer to previous content within the platform. Therefore, nothing prevents a quote tweet from being retweeted later within the information cascade associated with the topic. When a quote tweet is retweeted later, that information will appear with both behaviors in our dataset, making the data non-completely independent. In principle, a double classification of a tweet is correct, considering the definition of lean behaviors in the manuscript, but it generates some duplicated information. We took the following steps to address the issue described and test the robustness of our results. First, we noted that quote tweets behave like regular tweets in the case of retweeting following a power law distribution (Lu et al., 2014). The power law distribution of the number of retweets in the case of quote tweets in Figure S1 indicates that a small number of quote tweets are retweeted; therefore, the duplicated information is expected not to change significantly the results of the analysis.

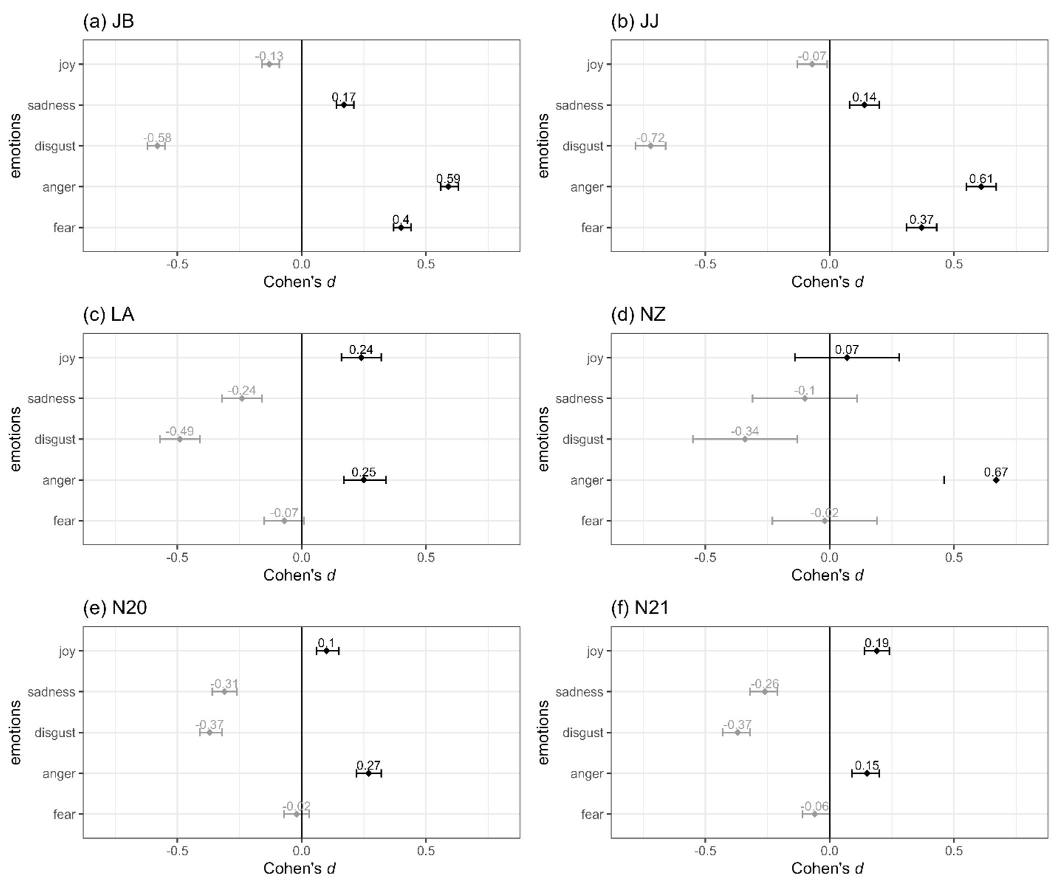


SF1. Distribution of the number of retweets counts for quote tweets for the six cases studied.

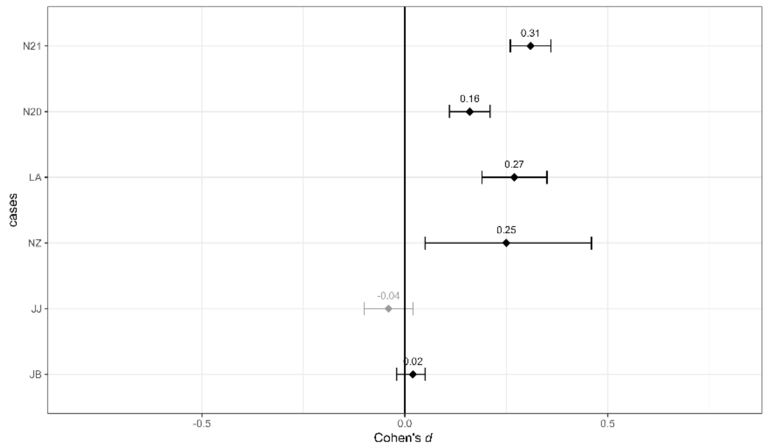
Second, to verify that duplicated information does not significantly change the results of the effect sizes reported for RQ1 and RQ2, we replicated our analysis with three variations of the datasets to eliminate the non-independency of the data: (1) we deleted the duplicated information from both lean behaviors, (2) we deleted the duplicated information from lean-forward, and (3) we deleted the duplicated information from lean-back behavior. The results of these three variations consistently show that the differences between lean behaviors with portions of non-independent data do not change significantly when that non-independence is removed. An explanation for these results is that most quote tweets are not retweeted (see SF1); therefore, the duplication of information was small in the complete dataset, and it seems negligible for the statistical inference about lean-forward and lean-back behaviors. Figures SF2 and SF3 present the results of RQ1 and RQ2 with modification (1) of the dataset, Figures SF4 and SF5 present the results of RQ1 and RQ2 with modification (2), and Figures SF6 and SF7 present the results of RQ1 and RQ2 with the modification (3),



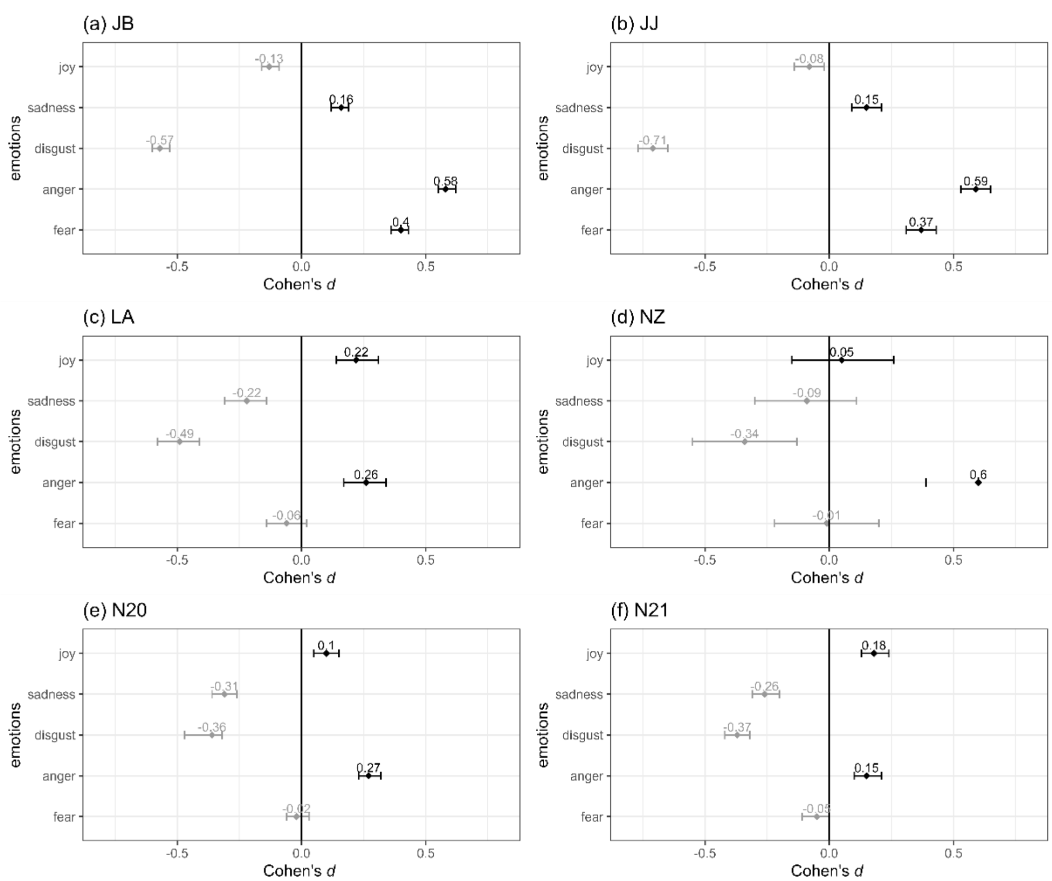
SF2. Cohen’s d effect sizes and confidence intervals comparing mean valence level for all cases. Dataset with duplicated information eliminated from both lean behaviors, modification (1).



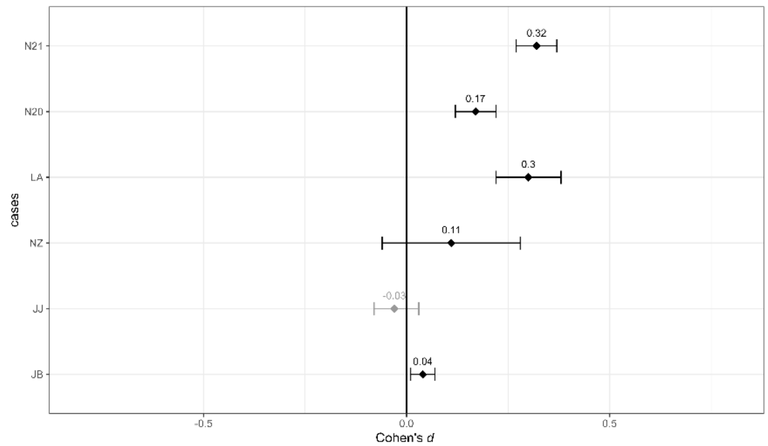
SF3a-e. Cohen’s d effect sizes and confidence intervals comparing mean emotions level. JB & JJ: political; LA & NZ: earthquakes; N20 & N21: sports. Dataset with duplicated information eliminated from both lean behaviors. Modification (1).



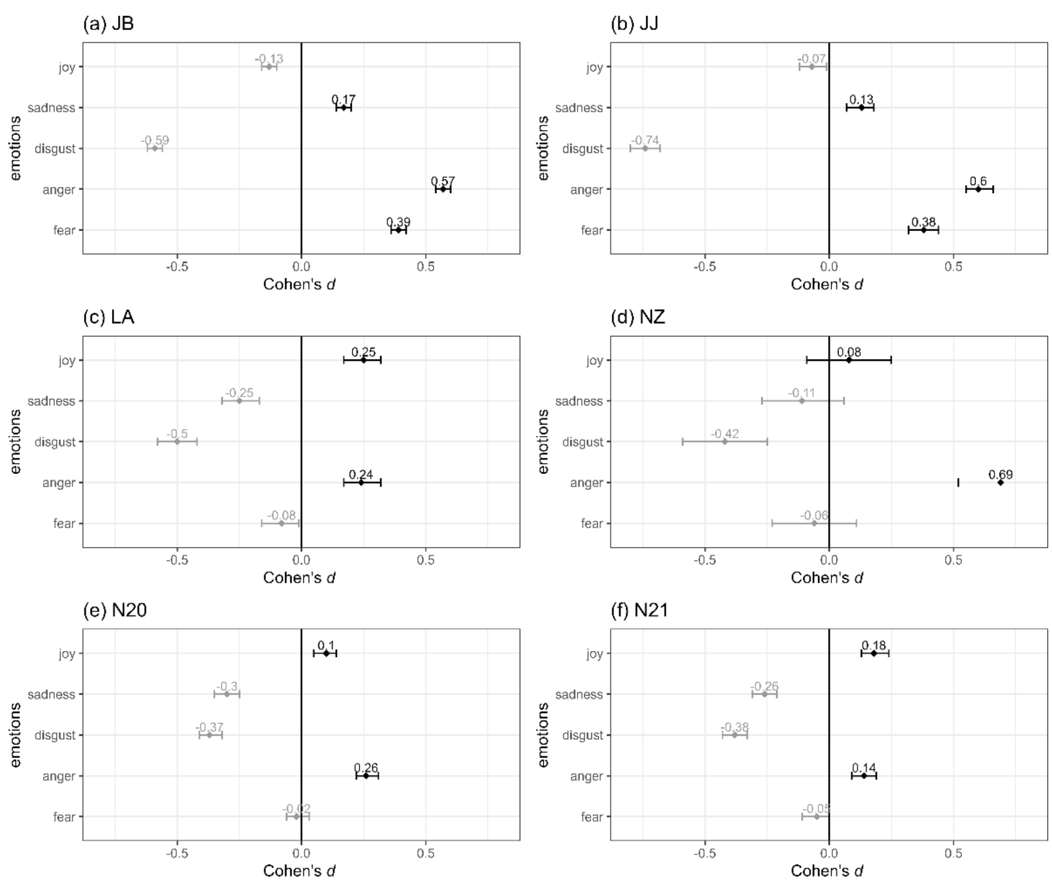
SF4. Cohen’s d effect sizes and confidence intervals comparing mean valence level for all cases. Dataset with duplicated information eliminated from lean-forward behavior. Modification (2).



SF5a-e. Cohen’s d effect sizes and confidence intervals comparing mean emotions level. JB & JJ: political; LA & NZ: earthquakes; N20 & N21: sports. Dataset with duplicated information eliminated from lean-forward behavior. Modification (2).



SF6. Cohen’s d effect sizes and confidence intervals comparing mean valence level for all cases. Dataset with duplicated information eliminated from lean-back behavior. Modification (3).



SF7a-e. Cohen’s d effect sizes and confidence intervals comparing mean emotions level. JB & JJ: political; LA & NZ: earthquakes; N20 & N21: sports. Dataset with duplicated information eliminated from lean-forward behavior. Modification (3).

Section 2: Familywise correction of SCM estimates

Considering the number of estimates in the SCM, we conducted a familywise correction for the significance α level to control Type I error in our analysis. According to Smith and Cribbie (2013), an adjusted Bonferroni procedure that accounts for the degree of correlation between parameters is recommended instead of the classic Bonferroni correction when using structural models. The traditional Bonferroni correction has been shown to be overly conservative, making it highly unlikely that researchers uncover all of the true relations in the model (Cribbie, 2007). Smith and Cribbie (2013) showed that the adjusted Bonferroni method "produces per-parameter higher power rates than the original Bonferroni procedure without inflating familywise error"(p.79). The adjusted Bonferroni method is defined according to Equation S1:

$$\alpha_{per\ test} = \frac{\alpha_{familywise}}{k^{1-\sqrt{\overline{|r_j|}}}} \qquad \text{Equation S1}$$

Here, α represents the significance level, k the number of null hypothesis tests performed, and $\overline{|r_j|}$ the average absolute correlation of parameter j with other parameters in the model. Using R, we implemented Equation S1. According to the results presented in Table S1, we opted for using a significance level of 0.008 to assess individual parameter estimates in the results of this work.

Table S1.

Significance level by estimates for an $\alpha_{familywise}$ level of .05 in the SCM.

	$\alpha_{adjusted}$ politics	$\alpha_{adjusted}$ earthquake	$\alpha_{adjusted}$ sports
fear → M	.009	.010	.008
ang → M	.009	.009	.009
dis → M	.011	.009	.009
sad → M	.009	.010	.010
joy → M	.011	.011	.012

M → fear	.009	.010	.008
M → ang	.009	.009	.009
M → dis	.010	.009	.008
M → sad	.009	.009	.009
M → joy	.010	.011	.010

Note. A → B denotes the effect of A on B

References

Cribbie, R. A. (2007). Multiplicity Control in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(1), 98–112. https://doi.org/10.1080/10705510709336738

Lu, Y., Zhang, P., Cao, Y., Hu, Y., & Guo, L. (2014). On the Frequency Distribution of Retweets. *Procedia Computer Science*, *31*, 747–753. https://doi.org/10.1016/j.procs.2014.05.323

Smith, C. E., & Cribbie, R. A. (2013). Multiplicity Control in Structural Equation Modeling: Incorporating Parameter Dependencies. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(1), 79–85. https://doi.org/10.1080/10705511.2013.742385