

# Exploring Cognitive Dissonance on Social Media

Jie Bai<sup>1,2</sup>

Qingchao Kong<sup>1,2</sup>

Lijing Li<sup>1,2</sup>

Lei Wang<sup>1,2\*</sup>

Daniel Zeng<sup>1,2,3</sup>

<sup>1</sup> The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China

<sup>2</sup> Shenzhen Artificial Intelligence and Data Science Institute (Longhua), Shenzhen, China

<sup>3</sup> University of Chinese Academy of Sciences, China

{baijie2013, qingchao.kong, linjing.li, l.wang, dajun.zeng}@ia.ac.cn

**Abstract**—Cognitive dissonance is a ubiquitous phenomenon which can be applied in various fields potentially. In this paper, we study cognitive dissonance through empirical analysis on social media platforms. Our study focuses on a recent “reversal event” - a topic or event experiencing a reversed development trend because of the new facts. Through statistical analysis and semantic analysis based methods, we found that (1) after the event is revised, the performance of the original followers were abnormal, which is consistent with the existence of cognitive dissonance; (2) the followers’ attitude afterwards usually tended to maintain their previous behaviors. This research provides a primary building block towards the mental inference based behavior prediction for social media users, which is of great value for security related research issues.

**Keywords**—cognitive dissonance, empirical analysis, social media users, mental inference

## I. INTRODUCTION

Cognitive dissonance theory, proposed by Festinger in 1957 [1], is a classical theory in social cognitive psychology. It describes the unbalanced state of mind when individual faced with inconsistent cognitions, such as receiving new information, being forced to do something reluctantly, or making decision from multiple candidates. Because of the inner pressure of being mentally consistency, one usually tend to change the unbalanced state, i.e., “discrepancy reduction”, such as changing the original cognition, involving new factors to balance existing cognitions, or trivializing the cognitions. The cognitive dissonance theory is widely used to model human psychological changes and explain human behaviors. Exploring cognitive dissonance of social media users could provide insights on causes and potential responses of the users to unexpected events or emergencies, which is particularly beneficial to the government and enterprises in security related situations.

The cognitive dissonance theory has been extensively studied from both theoretical research and practical application [2]. A series of works has been proposed to extend and clarify the cognitive dissonance theory. The self-consistency model [3] claimed that the cognitive dissonance happens when inconsistent cognitions are threatening one’s sense of self. For example, the decision makers experience cognitive dissonance when their behaviors conflict with the belief that they are rational [4]. The action-based model [5] explains that the motivation of reducing dissonance is to remove the conflicting action tendencies. While the theory has been applied in various scenes including fake news propagation [6] and decision makings [7], existing literatures usually study cognitive

dissonance through artificial experiments or limited amount of samples. The human behaviors in real, intricate environments needs to be further studied.

In this paper, we conducted a pilot study on cognitive dissonance among social media users. We investigate a typical “reversal event”, a hot topic discussed on Sina Weibo during the period from Jan. 2019 to Mar. 2019, about a famous “postdoc” actor who actually did not meet the standard of a doctoral degree. We noticed that many users discussed about the actor and his “doctor” or “postdoc” identity, both before and after the truth revealed. However, different users held different opinions, some people still expressed their support standpoint even after the truth revealing.

To validate the cognitive dissonance theory to the users who witnessed this reversal event, and discover when and why the cognitive dissonance might happen, we conducted statistical and semantic analysis: for the statistical analysis, we found that the behaviors of the origin followers were significantly abnormal because of the cognitive dissonance; for the semantic analysis, we further investigated the response of the followers, founding that the followers’ attitude afterwards usually tended to maintain their previous behaviors from multiple perspectives. Through the above empirical analysis, we proved the existence of cognitive dissonance on social media users, explored the insights of the state of mind in different user groups, and constructed a primary dataset for mental inference based stance prediction.

## II. DATA COLLECTION AND STATISTICAL ANALYSIS

### A. Data Collection and Preparation

We collected the users (masked by user ID) and their discussion (user posts along with the timestamp) about the targeted topic on Sina Weibo. Specifically, we collected the posts which mentioned the actor as well as his related information, such as “doctor”, “postdoc” or his new productions. The “postdoc” identity of the actor attracted wide attention since Jan. 2019, when he posted his postdoc offer on Sina Weibo; then he was discovered not deserved his doctor degree on around 12:00 am, Feb. 6th, 2019. In the following of this paper, the actor is denoted as “Z”, and the time point of 12:00 am, Feb. 6th, 2019 is denoted as “ $t^*$ ”. Considering the fact of time and data balance, we limited the time of the collected posts to the period “from Jan. 1st, 2019 to Mar. 15th, 2019”. As a result, we acquired 94,548 posts published by 63,528 users, where 2,394 users participating in the discussion before  $t^*$  and 61,417 afterwards.

\* Corresponding author

## B. Statistical Analysis

Through the initial statistics, we found that 283 users had posts both before and after  $t^*$ , i.e., the users who posted before  $t^*$  mostly did not participate in the discussion afterwards. This is inconsistent with the intuition that the followers should have paid more attention to Z when he attracted wider discussion. To analyze the behavior of the followers, their post frequency histograms are plotted in Fig. 1.

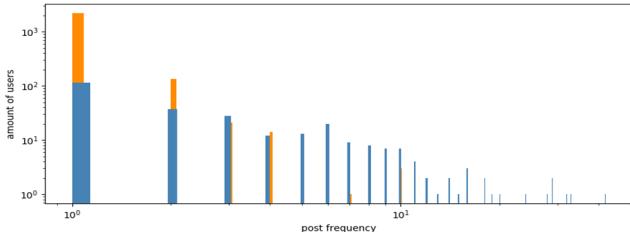


Fig. 1 Comparison of the initial followers' post frequencies. The orange and blue bars represents the post frequencies before and after  $t^*$  separately.

From Fig. 1 we can see that the followers' post frequencies before and after  $t^*$  obey the power law distribution approximately, however, not the same power law distribution. A reasonable assumption is that after  $t^*$ , the followers experienced cognitive dissonance, which make their behavior abnormal. To validate this assumption, i.e., the followers' behavior are abnormal, we conducted a statistical hypothesis test following the method proposed by Bessi [8], which is used to evaluate the difference between two discrete power-law distributions.

For a discrete power-law sample set,  $X = \{x \mid p(x) \propto x^{-\alpha}\}$ , its log-likelihood function is:

$$l(\alpha, x_{\min}) = n \ln \alpha + n \alpha \ln x_{\min} + (\alpha + 1) \sum_{i=1}^n \ln x_i, \quad (1)$$

where  $x_{\min}$  is the minimum of  $X$ , and  $\alpha$  is estimated through:

$$\hat{\alpha} \approx 1 + n \left[ \sum_{i=1}^n \ln \frac{x_i}{x_{\min}} - 0.5 \right]. \quad (2)$$

If there are two discrete power-law samples,  $X_1$  and  $X_2$ , their difference  $\Delta$  could be estimated based on the log-likelihood:

$$\Delta = 2 \times [l(H_1 \mid X_1) + l(H_1 \mid X_2) - l(H_0 \mid X_1 \cup X_2)], \quad (3)$$

where  $l(H_0 \mid X_1 \cup X_2)$  is the null model of the hypothesis test, and  $l(H_1 \mid X_1) + l(H_1 \mid X_2)$  is the alternative model.

Let  $X_1$  and  $X_2$  be the post frequencies of the followers before and after  $t^*$  respectively. The statistical results are:  $\hat{\alpha}_1 = 2.32$ ,  $\hat{\alpha}_2 = 1.61$ ,  $\Delta = 109.70$  with  $p$  value  $< 0.001$ . Based on the results, the followers' behavior are significantly different after  $t^*$ .

However, only the difference cannot prove the behavior abnormality. We further constructed a control group as comparisons, i.e., we conducted the same hypothesis test on the users who participating in the discussion after  $t^*$ . For the most of these people, Z was rather a hot topic than an idle, so there is less possibility that they had cognitive dissonance when referring Z. The statistical results on the control group are:  $\hat{\alpha}_1 = 2.08$ ,  $\hat{\alpha}_2 = 2.11$ ,  $\Delta = 0.07$  with  $p$  value  $< 0.001$ . The scatter

diagrams as well as the corresponding fitting results are plotted in Fig. 2. The results show that the general users perform similarly whenever before or after  $t^*$ , whereas the followers of Z performed different significantly.

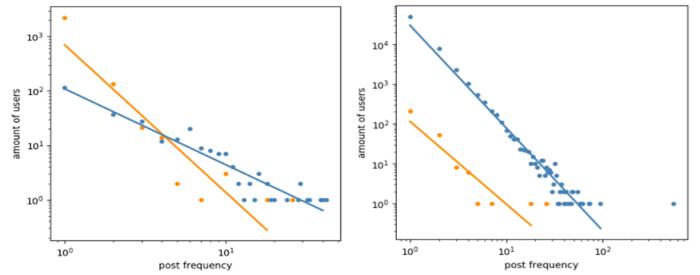


Fig. 2 Statistical data and the fitting results on post frequencies of the followers (left) and the general users (right).

Based on the analysis above, the conclusion can be drawn that the followers' behavior were abnormal after  $t^*$ , i.e., they experienced cognitive dissonance after the reversion of the topic.

## III. SEMANTIC ANALYSIS

In the semantic analysis part, we explored the insights of the followers' cognitive dissonance through their opinions. During the analysis, a dataset for mental inference based stance prediction is also accumulated for future works.



Fig. 3 Word clouds extract from the contents posted by  $U_R$  (left) and  $U_B$  (right). The time range of the contents are Jan. 1st, 2019 to  $t^*$ , i.e., before the topic reversion. Some specific names of actors are removed for better observation.

Firstly, we compared the difference between the followers who posted after  $t^*$  with whose who didn't. For convenience, in the rest of this paper, the group of followers who posted both before and after  $t^*$  is denoted as  $U_R$ , while the other group who posed only before  $t^*$  is denoted as  $U_B$ . The frequent words extracted from their post contents before  $t^*$  are shown in Fig. 3. By observing the contents that they posted before  $t^*$ , we discovered that these two groups of users, who performed differently afterwards, concerned different aspects initially. More specifically, before  $t^*$ , the users in  $U_R$  usually concerned more about the postdoc or doctor identity of Z, whereas the users in  $U_B$  usually concerned about Z as an actor. This observation explains the motivation of the followers' behavior in the future, and reveals the insights of their cognitive dissonance as well, i.e., the followers who concerned the doctor identity of Z tend to post more frequently, while those who concerned Z's acting skill tend to keep silence.

It is worth noticing that even in  $U_R$  the users' response were various. From the viewpoint of cognitive dissonance theory,

the users' response can be illustrated by Fig. 4. Here U denote a user who experienced cognitive dissonance, D denote the postdoc or doctor identity of Z, and "+" or "-" means the positive or negative relationship.

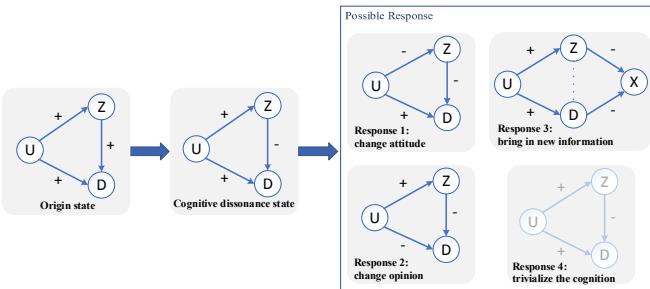


Fig. 4 The evolution of the followers' state of mind according to the cognitive dissonance theory.

Based on the cognitive dissonance theory, we further grouped the users in  $U_R$  according to their response and attitude after  $t^*$ . The grouping results are listed in Table 1, where the groups 1-4 are corresponding to the response 1-4 in Fig. 4, and the group 5 is corresponding to the cognitive dissonance state.

TABLE I. GROUPING RESULTS OF  $U_R$

Index	Response	Attitude	# of users	Representative opinions
1	change attitude	oppose	19	"A person without moral integrity cannot be a good actor."
2	change opinion	support	49	"I like him because of his acting skill, instead of his doctor degree."
3	bring in new information	blame other people	36	"Those people are criticizing because of envy." "The problem is the criterion of his graduate school."
4	trivialize the cognition	neutral	177	Forwarding the related news with no personal attitude explicitly.
5	cognition dissonance state	confused	2	"What should we do to help you?"

According to the attitude of different groups, we had the following findings:

- The users in group 4, which takes the large proportion of  $U_R$ , were just forwarding the related news without personal attitude. We further found that many users in group 4 were agency accounts interested in the latest news, so there was no cognitive dissonance for them; about 1/4 users in group 4 were personal accounts who reduced the cognitive dissonance through trivialization the cognition (by citing the news to express their neutral attitude).
- Among the other users than group 4, most of the users (group 2 & 3) were supportive to Z from multiple perspectives. Specifically, the users in group 2 supported Z by reducing the value of his doctor identity; whereas the users in group 3 put their focus on other people or objects, such as the critics or the graduate school of Z.
- A small part of users (group 1) expressed their criticism to Z explicitly. We noticed that these part of users' comments before  $t^*$  were more calm and objective compared with the other followers.
- There are also individual users that expressed their confusion and sadness, which we believed that they were in the cognition dissonance state. The sentiment of the posts from these users were tend to supportive to Z.

Throughout the analysis, compared with the total number of the followers, there were only a small part of users who participated in the discussion after  $t^*$ , and these users originally paid more attention to the doctor identity of Z than his acting skill; among these users, many chose to be neutral or still supportive to Z, while only a small part of users changed their attitude. All of these conclusions indicate that the users' attitude afterwards usually tended to maintain their previous behaviors. With the conclusions above, the users' stance can be inferred by combining the cognitive dissonance theory and the users' previous behavior as indicator. Besides, we can also use these analysis results to infer the state of mind of the users who keep silence after  $t^*$ .

#### IV. CONCLUSION

In this paper, we investigated the cognitive dissonance phenomenon happened among the social media users, based on a typical reversal event happened recently on Sina Weibo. By conducting statistical analysis and semantic analysis, we validated the existence of cognitive dissonance, and explored the insights of cognitive dissonance for different user groups. Besides, we accumulated a primary dataset on mental inference based stance prediction. Our future work based on this paper includes: (1) in-depth mental and behavior analysis on social media users, (2) mental inference based stance prediction under multi-stage topics, and (3) intervention study of false information diffusion.

#### ACKNOWLEDGMENT

This work was supported in part by the National Key R&D Program of China under Grant 2016QY02D0305 and 2017YFC0820105, the National Natural Science Foundation of China under Grants 71621002, 71702181, the Key Research Program of the Chinese Academy of Sciences under Grant ZDRW-XH-2017-3, as well as the Early Career Development Award of SKLMCCS under Grant 20190204.

#### REFERENCES

- [1] Festinger L. A theory of cognitive dissonance[M]. Stanford university press, 1957.
- [2] Hinojosa A S, Gardner W L, Walker H J, et al. A review of cognitive dissonance theory in management research: Opportunities for further development[J]. Journal of Management, 2017, 43(1): 170-199.
- [3] Aronson E. The theory of cognitive dissonance: A current perspective[M]//Advances in experimental social psychology. Academic Press, 1969, 4: 1-34.
- [4] Sleesman D J, Conlon D E, McNamara G, et al. Cleaning up the big muddy: A meta-analytic review of the determinants of escalation of commitment[J]. Academy of Management Journal, 2012, 55(3): 541-562.
- [5] Harmon - Jones E, Amadio D M, Harmon - Jones C. Action - based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict[J]. Advances in experimental social psychology, 2009, 41: 119-166.
- [6] Nerren J B. Civic Engagement, Fake News and the Path Forward[J]. Journalism, 2018, 8(2): 51-55.
- [7] Acharya A, Blackwell M, Sen M. Explaining preferences from behavior: A cognitive dissonance approach[J]. The Journal of Politics, 2018, 80(2): 400-411.
- [8] Bessi A. Two samples test for discrete power-law distributions[J]. arXiv preprint arXiv:1503.00643, 2015.