Research Article

The limitations of the backfire effect

Research and Politics
July-September 2017: 1–5

DOI: 10.1177/2053168017716547
journals.sagepub.com/home/rap

Kathryn Haglin

Abstract

Nyhan and Reifler (2010, 2015) document a "backfire effect," wherein attempts to correct factual misperceptions increase the prevalence of false beliefs. These results are widely cited both in and outside of political science. In this research note, I report the results of a replication of Nyhan and Reifler's (2015) flu vaccine study that was embedded in a larger study about flu vaccines. The backfire effect was not replicated in my experiment. The main replication result suggests the need for additional studies to verify the backfire effect and identify conditions under which it occurs.

Keywords

Vaccine, replication, backfire effect, misperception

Introduction

Political scientists are increasingly aware of the effect of misperceptions on behavioral intentions and attitudes. While Delli Carpini and Keeter (1996) famously note that most citizens do not have much factual knowledge about politics, Kuklinski et al. (2000) point out the differences between being uninformed and misinformed. When a person is misinformed, false, misleading, or unsubstantiated information can create the basis for their policy preferences. Further, the sources used to obtain the misinformation are often directly related to a person's political preferences.

While there is some evidence that providing relevant facts has the ability to change people's issue opinions (Gilens, 2001; Kuklinski et al., 2000), information is often received in a much noisier environment. Other studies have found that individuals are often resistant to evidence that contradicts their opinions (Redlawsk, 2002; Taber and Lodge, 2006). The literature, however, remains unsettled as to exactly when and how misperceptions can be corrected. In addition, the role of the "backfire effect," where corrective information can actually make false beliefs more prevalent, in these processes remains unclear. For example, Weeks and Garrett (2014) do not find evidence for the backfire effect in a study about correcting rumors in the 2008 presidential campaign. Similarly, Ecker et al.'s (2014) study of racial attitudes finds those attitudes do not change the effectiveness of discounting information. Looking at similar attitudes, Garrett et al. (2013) find no evidence of these backfire effects in a study about a proposed Islamic cultural center in New York City. By contrast, Nyhan and Reifler (2010, 2015) find evidence for a backfire effect in a vaccines context as well as in the case of being correctly informed about the presence of weapons of mass destruction in Iraq.

This research note reports a replication of Nyhan and Reifler's (2015) flu vaccines study embedded within a larger experimental study of flu vaccine intentions and attitudes. Data generated in the experiment do not replicate the backfire effect or the finding that corrections reduce misperceptions about vaccine safety. This suggests that more work is needed to validate the backfire effect, establishing the conditions under which it occurs and the size of its effect.

Methods

Data collection

A replication of Nyhan and Reifler's (2015) flu vaccine study was embedded in a larger experimental study of flu vaccine intentions and attitudes.[1] The replication followed their published methods and procedures. The data were collected using Amazon's Mechanical Turk platform and due to the nature of the larger study, the worker pool was

Texas A&M University College Station, USA

Corresponding author:
Kathryn Haglin, Texas A&M University College Station, 2010 Allen Building, 4348 TAMU, College Station, TX 77843, USA.
Email: hagl0098@tamu.edu

Table 1. Characteristics of respondents in Nyhan and Reifler (N + R) (2015) and replication sample (Replication) by (%).

	Correction (N + R (2015))	Correction (Replication)	Total (N + R (2015))	Total (Replication)
Age				
18–29	18	39	21	40
30–44	29	39	24	42
45–59	29	18	28	15
60+	24	3	26	3
Gender				
Male	49	45	48	49
Female	51	55	52	50
Education				
High school or less	39	6	40	8
Some college	32	24	34	27
Bachelor's degree/college graduate	18	35	17	35
Race/ethnicity				
White	76	77	71	70
Black	9	9	12	8
Hispanic	8	11	10	13
Other	7	3	6	7
Concern about side effects				
Not at all concerned	13	23	14	25
Not too concerned	33	33	31	33
Somewhat concerned	34	19	32	22
Very concerned	11	21	13	11
Extremely concerned	9	3	11	8
Number of observations	341	66	1000	474

limited to those with Internet Protocol addresses located in Texas. The study was fielded beginning in March 2016 and ending in May 2016. Respondents were adults located in Texas who were told they would be answering questions meant to elicit their opinions on important political issues of today. Data were collected from 525 respondents, and after dropping respondents found to not be located in Texas, the final sample used in the analysis totaled 474.

Study design

The replication was embedded in a larger study examining social identity appeals in public health messaging. Respondents who were randomly assigned to the replication condition, "Correction," received the same text from Nyhan and Reifler's (2015) protocol, which was taken nearly verbatim from the US Centers for Disease Control and Prevention website. This text told respondents that people cannot contract the flu from the flu shot or live virus nasal spray.

As noted by Nyhan and Reifler (2015), responses to vaccine information might vary based on one's pre-existing attitudes towards vaccines. Since it was not possible to accurately measure prior vaccine receipt as part of the study, I measured respondents' general concerns about vaccine safety and side effects. Specifically, I asked, "In general, how concerned are you about serious side effects from vaccines?" prior to the interventions. This was measured on a five point scale ranging from "not at all concerned" to "extremely concerned."

Outcome measures

After the experimental intervention, I measured the effects of each treatment on respondents' misperceptions about the flu vaccine (You can get the flu from the seasonal flu vaccine); feelings about vaccine safety (Just based on what you know, how safe do you believe the seasonal flu vaccine, meaning the flu vaccine available every year, is generally for most people to take?); and intent to get vaccinated in the future (How likely is it that you will get a flu vaccine for the seasonal flu during future flu seasons?). These measures were taken from Nyhan and Reifler (2015) and the full text and scale of each measure can be found in the online supplementary materials.

Results

Table 1 summarizes the characteristics of the respondents in the full Mechanical Turk sample and compares the demographics of respondents in Nyhan and Reifler's sample drawn from a YouGov/Polimetrix panel and the

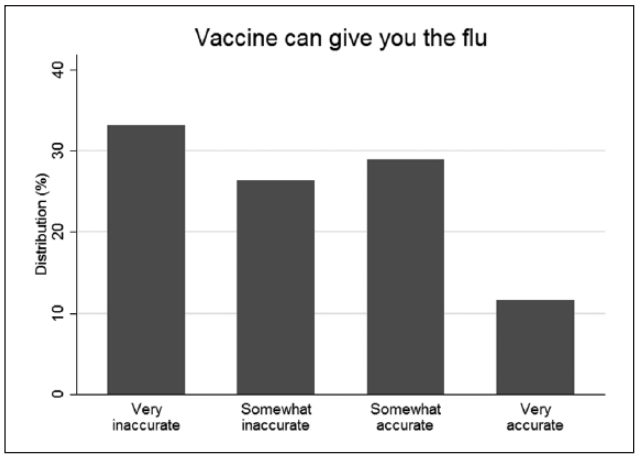

Figure 1. Distribution of vaccine misperception measure in the replication sample.

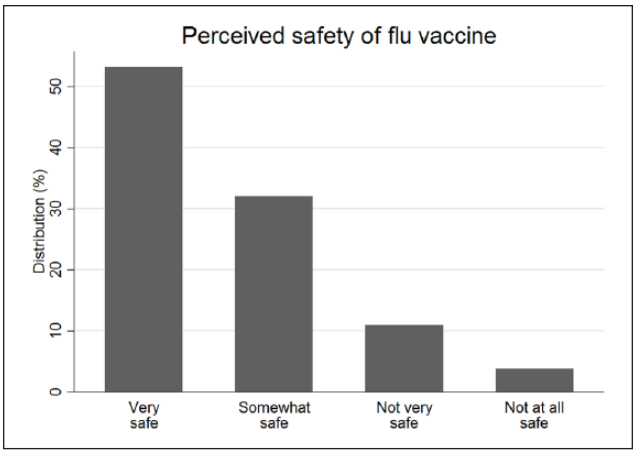

Figure 2. Distribution of vaccine safety measure in the replication sample.

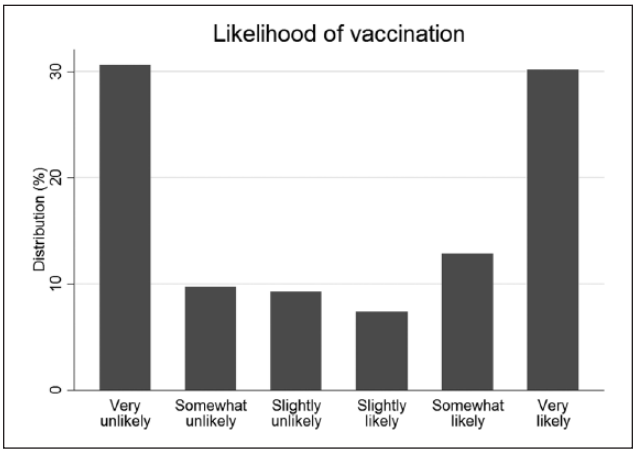

Figure 3. Distribution of intent to vaccinate measure in the replication sample.

replication sample. As expected in samples from Mechanical Turk, participants are not fully representative of the national adult population. The replication sample is younger, with only 3% of the participants in the 60+ category and approximately twice as many 18–29 year olds. The Mechanical Turk sample is also more educated than the original sample, with most of the participants having some college education or a college degree. However, both samples have similar racial compositions and gender distributions.[2]

The samples have a few differences in participants' level of concern about vaccine side effects. The Mechanical Turk participants are less concerned about side effects, with a 10% difference in the "not at all concerned" category between the samples. In the correction treatments, the replication sample was 10% higher in the "very concerned" category, while in the overall sample, the original study had slightly more "very concerned" participants. Finally, while neither study had a large number of "extremely concerned" individuals, the original sample had more of these participants, with a 6% difference between the samples in the correction treatment. Approximately 19% of the replication sample and 24% of the original sample fall into the high concern category, defined as saying you are either "very" or "extremely" concerned about vaccine side effects. Given this discussion, the relative similarity in these distributions gives us confidence that we can make similar inferences about high concern and low concern individuals.

Figures 1–3 summarize the distribution of responses to the three outcome variables of interest across all conditions in the replication study: the misperception that the flu vaccine can give you the flu, perceptions of the vaccine's safety, and self-reported intent to vaccinate in the next flu season. The results indicate that roughly 40% of the respondents believe that the myth that the flu vaccine can give you the flu is "somewhat accurate" (29%) or "very accurate" (12%). Yet far fewer believe the flu vaccine is unsafe, with only a total of 15% saying they consider the vaccine "not very safe" or "not at all safe." Figure 3 shows that the distribution of self-reported intentions to vaccinate is bimodal. 31% of respondents say it is unlikely they will get a flu vaccine in the next flu season, while 30% say it is very likely they will get a vaccine. The remaining 39% of the respondents were not as certain and approximately evenly distributed about the remaining response options.

Comparing these distributions with those from the original study, we can see that they are very similar to one another.[3] For the vaccine misperception measure, each of the bars is within roughly 5% of the corresponding bar in Nyhan and Reifler's distribution. The vaccine safety measure distributions are both compressed towards the "very safe" option, with the Mechanical Turk respondents having a larger proportion saying the vaccine is very safe. Similarly, the intent to vaccinate measure has the same bimodal distribution as the original study found.

Table 2. Replication of correction treatment on "vaccine can give flu"; Nyhan and Reifler (2015) (N + R) and replication sample (Replication).

Treatment	All respondents		Low concern		High concern	
	N + R	Replication	N + R	Replication	N + R	Replication
Correction	–0.39**	–0.79**	–0.34**	–0.87**	–0.49*	–0.87**
% confidence interval	[–0.65, –0.12]	[–1.10, –0.49]	[–0.64, –0.04]	[–1.22, –0.51]	[–1.02, 0.04]	[–1.46, –0.28]
Sample size	995	474	769	379	226	95

Notes: ** indicates significance at $p < 0.05$.
*indicates significance at $p < 0.10$.

Table 3. Replication of correction treatment on vaccine safety; Nyhan and Reifler (2015) (N + R) and replication sample (Replication).

Treatment	All respondents		Low concern		High concern	
	N + R	Replication	N + R	Replication	N + R	Replication
Correction	–0.31**	0.01	–0.33**	0.02	–0.14	–0.32
95% confidence interval	[–0.57, –0.05]	[–0.29, 0.30]	[–0.65, –0.02]	[–0.34, 0.38]	[–0.62, 0.34]	[–0.89, 0.24]
Sample size	997	474	772	379	225	95

Notes: ** indicates significance at $p < 0.05$.

Table 4. Replication of correction treatment on intent to vaccinate; Nyhan and Reifler (2015) (N + R) and replication sample (Replication).

Treatment	All respondents		Low concern		High concern	
	N + R	Replication	N + R	Replication	N + R	Replication
Correction	0.03	–0.18	0.13	–0.19	–0.49**	–0.08
% confidence interval	[–0.22, 0.28]	[–0.47, 0.10]	[–0.17, 0.42]	[–0.51, 0.14]	[–0.97, –0.02]	[–0.68, 0.53]
Sample size	998	474	772	379	226	95

Notes: ** indicates significance at $p < 0.05$.

Results: ordered probit models

The replication experiment's data show that the correction treatment is effective in reducing misperceptions about the flu vaccine. However, it fails to replicate the backfire effect identified by Nyhan and Reifler (2015). Tables 2–4 report these results across all three outcome variables (intent to vaccinate, safety, and vaccine misperception). The original models in Nyhan and Reifler (2015) include a variable for a "Danger" treatment used in their study; the replication study did not include such a treatment and thus it is not in the models shown here.[4] Higher values indicate more negative views of the flu vaccine; therefore, negative coefficients indicate that the intervention reduced false beliefs.

Table 2 shows the replication of the vaccine misperception results. Like Nyhan and Reifler, I find that the correction treatment was able to significantly reduce the false belief that the flu vaccine can give you the flu. This result holds in the full sample, as well as for both high and low concern individuals. However, Table 3 reports a failure to replicate these findings for the vaccine safety variable. While Nyhan and Reifler find that the correction reduced the belief that the flu vaccine is unsafe, this result is not found in the replication results. The replication coefficients are positive in the full sample and low concern groups and the high concern coefficient is negatively signed. This indicates that the backfire effect is not present in the full sample and low concern groups and that while there is some evidence of a backfire effect for high concern individuals, like the original study, the finding is not statistically significant.

Finally, Table 4 shows the effects of the correction intervention on individuals self-reported intent to vaccinate. Here, Nyhan and Reifler find a "backfire" effect amongst those in the high concern group, with the correction leading to a decrease in the likelihood to vaccinate among those most concerned about side effects. My results do not replicate this finding. While the coefficient is negatively signed, the confidence interval is centered on zero. This precise zero indicates that the correction treatment in the

replication study had no effect on the high concern individuals' behavioral intentions.

Discussion and conclusions

The results of this replication confirm that corrective information seeking to debunk myths about the flu vaccine affect beliefs about vaccine safety and intentions differently. However, I fail to replicate the "backfire effect" reported by Nyhan and Reifler (2015). These findings suggest that more work is needed to validate the backfire effect and the conditions under which it occurs. While my study is not a direct replication in that I do not use a national sample, my findings contribute to the literature on the backfire effect by replicating an existing experimental procedure as part of a larger study on a different population. While my study population presents its own unique challenges, the unsuccessful replication of the Nyhan and Reifler (2015) corrective information experiment shows that the primary result might be context-dependent and indicates the need for additional research to identify conditions when it occurs, when it does not, and which individuals are most strongly affected.

The consequences of misinformation can be costly as the major parties in the United States stake out positions on many areas of scientific research. In an increasingly polarized discussion space filled with a plethora of information sources, citizens can easily fall victim to misinformation. Research in political science that addresses how to establish best practices in not only correcting misperceptions, but also how to best shape political debates about science, will be increasingly important as these discussions unfold.

Declaration of Conflicting Interest

The author declares that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Supplementary Materials

The supplementary files are available at: http://journals.sagepub.com/doi/suppl/10.1177/2053168017716547. The replication files are available at: http://dx.doi.org/10.7910/DVN/1RD4HI

Notes

1. Nyhan and Reifler first identified the backfire effect in their 2010 paper; however, this paper does not replicate any portion of the 2010 article.
2. Additional information about the distribution of party identification, ideology, and economic views in the Mechanical Turk sample can be found in the Online Supplementary Materials.
3. The distributions from Nyhan and Reifler (2015) can be found in the Online Supplementary Materials.
4. In Nyhan and Reifler (2015), the "Danger" treatment used text from the US Centers for Disease Control and Prevention website that informed respondents that the flu is contagious, gave a list of signs and symptoms, and discussed the serious risks the flu poses.

Carnegie Corporation of New York Grant

This publication was made possible (in part) by a grant from Carnegie Corporation of New York. The statements made and views expressed are solely the responsibility of the author.

References

Delli Carpini MX and Keeter S (1997) *What Americans Know About politics and Why it Matters*. New Haven, CT: Yale University Press,.

Ecker UKH, Lewandowsky S, Fenton O, et al. (2014) Do people keep believing because they want to? Preexisting attitudes and the continued influence of misinformation. *Memory & cognition* 42(2): 292–304.

Garrett RK, Nisbet EC and Lynch EK (2013) Undermining the corrective effects of mediabased political fact checking? The role of contextual cues and nave theory. *Journal of Communication* 63(4): 617–637.

Gilens M (2001) Political ignorance and collective policy preferences. *American Political Science Review* 95(2): 379–396.

Kuklinski JH, Quirk PJ, Jerit J, et al. (2000) Misinformation and the currency of democratic citizenship. *The Journal of Politics* 62(3): 790–816.

Nyhan B and Reifler J (2010) When corrections fail: The persistence of political misperceptions. *Political Behavior* 32(2): 303–330.

Nyhan B and Reifler J (2015) Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine* 33(3): 459–464.

Redlawsk DP (2002) Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *Journal of Politics* 64(4): 1021–1044.

Taber CS and Lodge M (2006) Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50(3): 755–769.

Weeks BE and Garrett RK (2014) Electoral consequences of political rumors: Motivated reasoning, candidate rumors, and vote choice during the 2008 US presidential election. *International Journal of Public Opinion Research* 26(4): 401–422.