

Metaphors Affect Reasoning: Measuring Effects of Metaphor in a Dynamic Opinion Landscape

Paul H. Thibodeau (paul.thibodeau@oberlin.edu)

Peace O. Iyiewaure (peace.iyiewaure@oberlin.edu)

Oberlin College Department of Psychology
120 W. Lorain St; Oberlin, OH 44074 USA

Lera Boroditsky (lera@ucsd.edu)

UCSD Department of Cognitive Science
9500 Gilman Dr.; La Jolla, CA 92093 USA

Abstract

Metaphors pervade discussions of critical issues, making up as much as 10-20% of natural discourse. Recent work has suggested that these conventional and systematic metaphors influence the way people reason about the issues they describe. For instance, Thibodeau & Boroditsky (2011, 2013) found that people were more likely to want to fight back against a crime *beast* by increasing the police force but more likely to want to diagnose and treat a crime *virus* through social reform. Here, we report two norming studies and two experiments that reveal a shift in the overall landscape of opinion on the topic of crime. Importantly, we find that the metaphors continue to have an influence on people's reasoning about crime. Our results and analyses highlight the importance of up-to-date opinion norms and carefully controlled materials in metaphor research.

Keywords: metaphor; analogy; framing; reasoning

Introduction

Metaphors pervade discussions of critical issues, making up as much as 10-20% of natural discourse (Steen et al., 2010). Many metaphors are conventional and systematic (Lakoff & Johnson, 1980). For example, we commonly use terms that are associated with disease to talk about crime, as in "Crime *plagues* and *infects* cities." Further, most complex social issues are talked about using more than one system of metaphors. For example, in addition to disease metaphors we can also talk about crime as a beast or wild animal, as in "Crime *attacks* and *preys on* cities."

Recent work (e.g., Hauser & Schwarz, 2014; Landau, Sullivan, & Greenberg, 2009; Thibodeau & Boroditsky, 2011, 2013) has suggested that these metaphors are more than simply colorful ways of talking. Using different metaphors leads people to reason differently about social issues and follow different paths of inference. For example, when people saw a report that described crime as a *beast*, they were more likely to want to fight back by increasing the police force. When they read that crime was a *virus* they preferred diagnosing and treating the problem through social reform (Thibodeau & Boroditsky, 2011, 2013).

Of course, metaphors about social issues exist in a dynamic public discourse. As topics gain and lose popularity, get reframed in the news cycle, and are reconsidered in light of new events, public opinion and public engagement on different issues can shift. In this

paper, we revisit earlier studies (conducted from 2008 to 2011), replicate the results and compare our new data with both the older data sets and with other recently collected samples (Steen, Reijniere, & Burgers, 2014).

We find that while the overall landscape of opinion on the topic of crime has shifted, metaphorical frames continue to have an influence on people's reasoning about crime. In two experiments and two norming studies we examine some of the methodological considerations for research on metaphor as it exists in the dynamic landscape of public discourse. We address some common misunderstandings and review some inferential limitations of the available methods. Our results and analyses highlight the importance of up-to-date opinion norms and carefully controlled materials in metaphor research.

Experiment 1

Participants

Data from 650 participants were collected on Mechanical Turk using three exclusion criteria: we restricted our sample to participants living in the US, who were at least 18 years old, with a performance rating of at least 90%. We paid participants \$1.00.

Data from 84 (13%) participants were excluded because they submitted an incorrect completion code or because they reported having completed a similar study in the past. We also excluded data from participants who reported that English was not their first language, from participants who reported that they did not live in the US, and from participants who responded extremely quickly or slowly (i.e. faster than 10 seconds or slower than 300 seconds). This left data from 526 participants for analysis.

The average age of participants was 32.2 ($sd = 10.3$). Slightly more than half were male (56%), and 38%, 46%, and 16% identified as Democrat, Independent, and Republican, respectively.

Materials & Design

Participants were randomly assigned to one of two framing conditions. In one, crime was described as a *virus* ravaging a city; in the other crime was described as a *beast* ravaging a city. The remainder of the report was identical across conditions.

Crime is a {beast/virus} ravaging the city of Addison. Five years ago Addison was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the city's defense systems have weakened, and the city has succumbed to crime. Today, there are more than 55,000 criminal incidents a year - up by more than 10,000 per year. There is a worry that if the city does not regain its strength soon, even more serious problems may start to develop.

After reading the report participants were asked what they thought Addison should do to address the problem by rank ordering the response options listed below.

1. Increase street patrols that look for criminals.
2. Increase prison sentences for convicted offenders.
3. Reform education practices and create after school programs.
4. Expand economic welfare programs and create jobs.
5. Develop neighborhood watch programs and do more community outreach.

The amount of time that participants spent reading and ranking the responses was recorded. Participants were also asked a set of background demographic questions, including their age, sex, educational background, first language, geographic location, and political affiliation.

Norming Study: Participants, Materials & Design

A separate group of 250 participants were asked to complete two tasks¹. In one, they were asked to rate the degree to which each of the five policy responses emphasized enforcement versus reform using a 101-point scale (-50 = completely reform-oriented; 50 = completely enforcement-oriented).

Participants were also asked to match policy approaches to metaphors. They were told that two politicians were using different metaphors (i.e. *virus* or *beast*) to support different policy interventions for a crime problem. Participants were presented with the five policies listed above and asked to match each one to *virus* or to *beast*.

The order of the policy options in both of the norming tasks was randomized. The matching task always preceded the rating task.

Norming Results

Ratings Analyses in prior work (Steen et al., 2014; Thibodeau & Boroditsky, 2011, 2013), coded the policy responses into two categories: those that were enforcement-oriented and those that were reform-oriented. To establish an empirical basis for this coding Thibodeau & Boroditsky (2013) collected ratings of the five policy approaches along a dimension of enforcement versus reform. In that study, the

“patrols” ($M = 87.21, sd = 13.5$), “prison” ($M = 85.11, sd = 22.37$), and “neighborhood watch” ($M = 58.69, sd = 25.77$) options were viewed as enforcement-oriented (i.e. above the midpoint, 50, of a scale that ranged from 0, very reform oriented to 100, very enforcement-oriented). The other two response options – “education” ($M = 17.14, sd = 27.13$) and “economy” ($M = 20.82, sd = 30.93$) – were, on the other hand, viewed as reform-oriented. However, Thibodeau & Boroditsky (2013) noted that the “neighborhood watch” “option was not rated as extreme as ‘street patrols’ or ‘prison sentences’ suggest[ing] that it may represent a more balanced approach” (p. 4).

There are several important differences in the sociopolitical context today (i.e. in 2014) relative to when the initial experiments and norming studies were conducted (i.e. between 2008 and 2011). For instance, an economic crisis was at its peak when the initial samples were collected. In addition, the initial studies pre-dated salient incidents related to race, policing, and social justice that were major news issues in 2013 and 2014. Have people's interpretations of the issues and stimuli changed?

We found that the options that had been rated as strongly enforcement-oriented in the past were rated as strongly enforcement-oriented by the current sample: “patrols” ($M = 86.901, sd = 14.965$), and “prison” ($M = 87.479, sd = 18.351$). We also found that the options that had been rated as strongly reform-oriented in the past were rated as strongly reform-oriented by the current sample: “economy” ($M = 14.752, sd = 21.522$) and “education” ($M = 15.372, sd = 21.578$).

The “neighborhood watch” option, however, was rated as more reform-oriented ($M = 36.736, sd = 28.556$) in the current study, significantly below the midpoint of the scale, $t[241] = 7.226, p < .001$, and significantly different from how it was rated in earlier studies, $t[46.936; \text{variances not assumed to be equal}; n_1 = 35, n_2 = 242] = 4.643, p < .001$. Note that ratings for the other four policy options did not change significantly over this time (all $ps > .2$).

Matching Although the distinction between enforcement and reform is an important one, it may not be the best way to think about the relationship between the metaphors and interventions. As we have argued, our theory is that the frames should make people more likely to select policies that are congruent with the entailments of the metaphors.

Figure 1 illustrates how participants viewed the relationships between the policies and metaphors in the current study². Five separate repeated measure logistic regressions revealed that participants were more likely to match the “economy”, $\chi^2(1) = 38.666, p < .001 (B = -1.694, SE = .304, p < .001)$, “education”, $\chi^2(1) = 27.07, p < .001 (B = -1.469, SE = .308, p < .001)$, and “neighborhood watch”, $\chi^2(1) = 22.429, p < .001 (B = -1.262, SE = .284, p < .001)$

² Note that a similar norming study was reported by Thibodeau & Boroditsky (2013); however, the “neighborhood watch” option was not included in the response set because we had not included it as a response option in most of our experiments.

¹ Data from 8 participants were excluded because they submitted an incorrect completion code.

policies to the *virus* frame; participants were more likely to match policies that emphasized “patrols”, $\chi^2(1) = 37.085, p < .001$ ($B = 1.269, SE = .217, p < .001$), and “prison” sentences, $\chi^2(1) = 48.529, p < .001$ ($B = 1.562, SE = .241, p < .001$) to the *beast* frame.

This analysis suggests that policies that address the economy, educational system, and neighborhood watches are viewed not only as more reform-oriented but should be coded as congruent with the *virus* frame. In contrast, policies that emphasize increasing police patrols and extending prison sentences should be coded as congruent with the *beast* frame.

In several analyses below, we report the results of two methods for coding the policies: one that is consistent with the older set of norms (collected in 2011, reported in Thibodeau & Boroditsky, 2013), in which the “neighborhood watch” option is categorized as consistent with the *beast* metaphor; and one that is consistent with the results of the more recent norming study, in which the “neighborhood watch option” is categorized as consistent with the *virus* metaphor.

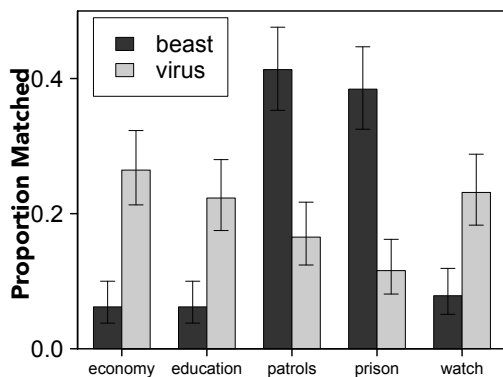


Figure 1. The proportion of policy responses that were matched to the *beast* and *virus* metaphors. Error bars reflect 95% confidence intervals.

Reanalyzing Previously Published Data

In addition to the new data collected in Experiment 1, we also analyzed data from a very similar study conducted by another laboratory in 2014 (Steen et al. 2014; Experiment 4³). These data were collected close in time to our new sample and norming studies (i.e. these data were collected in August of 2014; the norming study and current experiments were conducted in December of 2014).

Modeling Data

Many of the analyses that we report involve fitting logistic regression models with several predictors. The primary benefit of fitting logistic regression models is that we can include demographic characteristics as covariates. This is important as Thibodeau & Boroditsky (2011, 2013) found

³ Experiment 4 was the only study with a sufficiently large sample to be able to detect an effect of metaphorical frame.

that people with, for instance, different political affiliations tended to think differently about crime and were affected differently by the frames. To find the best models for the data, we utilized a stepwise model selection algorithm from the MASS library in R (Ripley et al., 2014). This algorithm takes a maximally parameterized model and tests alternatives that include subsets of predictor variables by comparing AIC values (by both pairing down from the maximally parameterized one and working up from the minimally parameterized one; Jaeger, 2008; Venables & Ripley, 2002).

For consistency, we fit the same maximally parameterized model in every analysis, which included tests of main effects by frame (*beast* or *virus*), time spent reading the report, political affiliation (Democrat, Republican, Independent), age, sex, and education and interactions between the frame and covariates (e.g., the frame and political affiliation). As a result, in most cases the initial model included 14 parameters. We report the best fitting model with an index of fit (AIC) as well as the AIC for the maximally and minimally parameterized models.

Results

We first present analyses based on the full distribution of responses. We then analyze the data as dichotomously coded according to old opinion norms (collected in 2011), and new opinion norms (collected in 2014).

The Full Distribution of Responses We found that the metaphor frames affected peoples’ responses on data pooled from Experiment 4 of Steen et al.,(2014) and data from the current Experiment 1, $\chi^2(4, N=876) = 16.346, p = .003$, Cohen’s $w = .14$. Analyzed separately, we found a marginal effect in Experiment 4, $\chi^2(4, N=350) = 8.609, p = .072$, Cohen’s $w = .16$, and a significant effect in our Experiment 1, $\chi^2(4, N=526) = 13.075, p = .011$, Cohen’s $w = .16$ (see Figure 2).

To determine which response option(s) drove the omnibus effect, we conducted five post-hoc chi-square tests of independence ($\alpha = .01$): one for each response option by constructing five separate 2 (frame: *virus* or *beast*) by 2 (option chosen: yes or no) frequency tables. This analysis revealed that the effect was driven by the “neighborhood watch” option, $\chi^2(1, N=526) = 13.728, p < .001$.

That is, people who read that crime was a *virus* were more likely to endorse the proposal to “Develop neighborhood watch programs and do more community outreach,” than people who read that crime was a *beast*. Although the “neighborhood watch” option was coded as enforcement-oriented in prior work, and therefore congruent with the *beast* frame, as we described above, it is currently viewed as more reform-oriented and consistent with the *virus* frame.

In our earlier samples that included the “neighborhood watch” option (i.e. Thibodeau & Boroditsky, 2013, Experiments 3 and 4), the overall distribution of choices differed from that of the more recent samples (see Figure 2), $\chi^2(4, N = 1,400) = 61.281, p < .001$. Bonferroni-corrected (α

= .01) post-hoc tests revealed that, in the past, people were more likely to endorse policies that emphasized the “economy” (32.3% vs. 21.1%), $\chi^2(1) = 24.626, p < .001$, and “prison sentences” (13.5% vs. 6.9%), $\chi^2(1) = 19.255, p < .001$; they were less likely to endorse policy approaches that emphasized “neighborhood watches” (15.8% vs. 26.2%), $\chi^2(1) = 21.646, p < .001$. There was not a significant difference in the proportion of people who endorsed “education” (17.7% vs. 19.5%), $\chi^2(1) = .645, p = .422$, or “street patrols” (20.6% vs. 26.3%), $\chi^2(1) = 6.194, p = .013$.

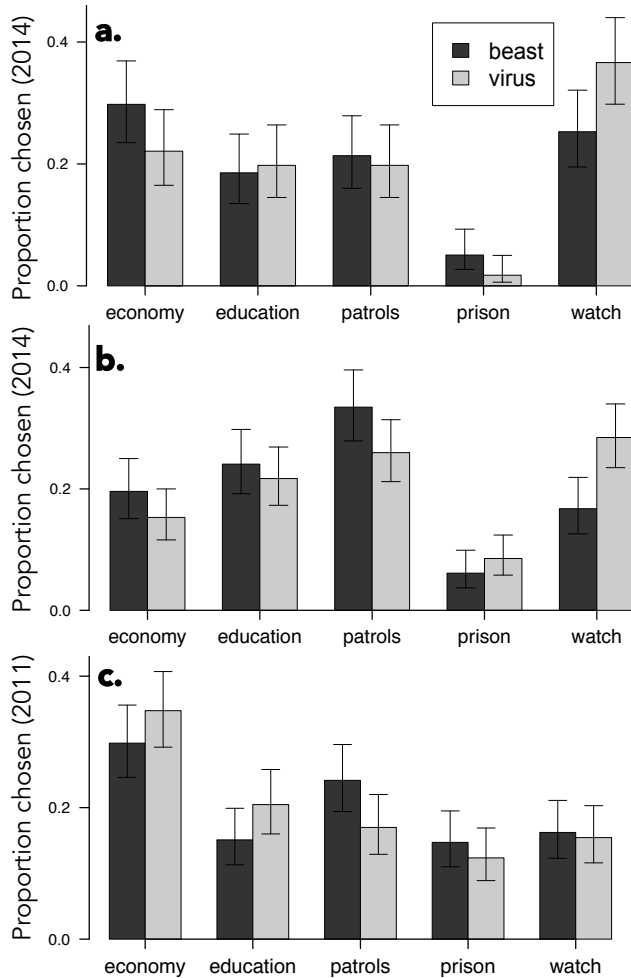


Figure 2. The distribution of policy choices by metaphor frame for three samples: a. Experiment 4 of Steen et al. (2014); b. the current Experiment 1; and c. data pooled from Thibodeau & Boroditsky (2013) Experiments 3 and 4. Error bars denote 95% confidence intervals.

This difference suggests that there may have been a cultural shift in how people conceptualize aspects of crime (including neighborhood watch programs, prison, and policing), the economy, and education over the past six years.

To understand how these cultural shifts may have affected the interpretation of results reported in Steen et al., (2014)

as either enforcement- or reform-oriented we dichotomized the responses according to both sets of opinion norms.

Coded Responses When data from Steen et al., (2014) and our Experiment 1 were coded according to 2011 opinion norms, with the “neighborhood watch” option categorized as enforcement-oriented and consistent with the *beast* metaphor, we did not find the predicted effect of the metaphor frame ($AIC = 1164.2$; $AIC_{Max} = 1192.6$; $AIC_{Min} = 1194.6$).

In fact, on this coding scheme, people who read that crime was a *virus* appear more enforcement-oriented ($B = 1.031, SE = .478, p = .031$), as do older participants ($B = 0.042, SE = .011, p < .001$), and Republicans (who appear more enforcement-oriented than Democrats, $B = .791, SE = .220, p < .001$, and Independents, $B = .537, SE = .218, p = .014$).

That is, the apparent effect of the metaphor frame was exactly the reverse of what we found in prior work. To examine whether this apparent discrepancy was due to an outdated coding scheme, we examined the same data but this time coding the responses according to new norming data ($AIC = 1060.9$; $AIC_{Max} = 1074.0$; $AIC_{Min} = 1096.7$).

This model revealed an effect of the metaphor frame in the predicted direction. People who read that crime was a *virus* were more likely to prefer a policy response that was consistent with the *virus* metaphor (i.e. to be more reform-oriented), $B = -.474, SE = .231, p = .041$: 34.0% of participants who read that crime was a *beast* chose an enforcement-oriented response compared to 29.6% of participants who read that crime was a *virus*. Participants were 4.9 and 5.1 percentage points more likely to choose an enforcement-oriented response in Steen et al., (2014)’s Experiment 4 and our Experiment 1, respectively.

In addition, Republicans were more likely to be enforcement-oriented than Independents, $B = .888, SE = .208, p < .001$, and Democrats, $B = 1.008, SE = .213, p < .001$. On average, Republicans choose an enforcement-oriented response 50.7% of the time; Independents and Democrats choose an enforcement-oriented response 29.6% and 26.6% of the time, respectively.

Separate analyses on the data from Experiment 4 of Steen et al., (2014) and the current Experiment 1 reveal consistent results. Both data sets show the predicted effect of metaphorical frame on people’s policy preferences, replicating prior findings (Thibodeau & Boroditsky, 2013). The apparent discrepancy reported in Steen et al., (2014) was introduced because the coding scheme relied on outdated opinion norms.

Experiment 2

To further investigate the reliability of the framing effect and to ensure that it could be elicited in the absence of the relatively ambiguous “neighborhood watch” option, we conducted a follow-up experiment in which participants were asked to choose between two policy approaches.

Participants

Data from 650 participants were collected on Mechanical Turk, using the same exclusion criteria as in Experiment 1. Data from 109 participants were excluded for one of the reasons listed in Experiment 1, leaving data from 521 participants for analysis.

The average age of participants was 32.6 ($sd = 10.9$). Slightly more than half were male (53%), and 41%, 43%, and 16% identified as Democrat, Independent, and Republican, respectively.

Materials & Design

The materials and design for Experiment 2 were identical to those of Experiment 1, except that participants were asked to choose between two policy responses: “patrols” and “education”. We chose these two policy options because they clearly differed along the two relevant dimensions: the “patrols” option was viewed as more consistent with the *beast* frame and more enforcement-oriented; the “education” option was viewed as more consistent with the *virus* frame and more reform-oriented. In addition, both of these response options were widely endorsed in previous experiments, unlike, for instance, the “prison” option.

Results

In this case, we found that 54.4% of policy preferences were congruent with the metaphor frame (i.e. on coding the “patrols” option as congruent with the *beast* metaphor and the “education” option as congruent with *virus*), $\chi^2(1, N = 521) = 4.240, p = .039, w = .09$. Republicans were more enforcement-oriented than Democrats, $B = .921, SE = .282, p = .001$, and Independents were more enforcement-oriented than Democrats, $B = .538, SE = .197, p = .006$ (AIC = 702.5; AIC_{Max} = 714.0 AIC_{Min} = 712.2).

That is, 61.0% of participants who read that crime was a *beast* preferred a policy that emphasized increasing “patrols” over reforming “education.” In contrast, 53.4% of participants who read that crime was a *virus* preferred the enforcement-oriented policy. For context, 70.7% of Republicans, 61.1% of Independents, and 48.8% of Democrats preferred the enforcement-oriented approach.

General Discussion

In summary, these analyses reveal a reliable effect of a metaphor frame on a policy judgment: on both coded and raw (i.e. uncoded) data. In Experiment 1 people who read that crime was a *virus* were more likely to view the “neighborhood watch” option as the most effective policy approach compared to people who read that crime was a *beast*. In Experiment 2, we found that the metaphor framing effect did not rely solely on the presence of this response option.

One additional conclusion that can be drawn from these experiments is that it can be difficult to study framing in a dynamic real-world context. The issue of crime is consistently a focal issue in the media, although the way it is

covered changes in predictable and unpredictable ways depending on the media outlet and as a result of salient events and political cycles (e.g., Barak, 1995). These variables can have profound effects on the way people think about crime (e.g., Iyengar, 1994). As a result, it is important to consider the possibility that variability in the cultural context may affect the way that metaphors influence reasoning for real-world issues by, for instance, using within-sample norming studies to gauge the degree to which people view policies as consistent with frames.

Can we tell which of the two metaphors (or both) are doing the work?

The results so far inform us that there is a psychological difference between the two metaphorical frames. However they do not tell us *how* the metaphors are shaping people’s views relative to what they would have thought without a metaphor. Are both metaphors equally contributing to the separation of opinion in the two conditions, or is just one of the metaphors doing all the work? It is also possible, for example, that both *beast* and *virus* metaphors make people more enforcement-oriented than they would have been without the metaphors, but the effect of the *beast* metaphor is stronger. Likewise, it is possible that both the *beast* and *virus* metaphors make people more reform-oriented than they would have been without the metaphors, but the effect of the *virus* metaphor is stronger.

These questions raise an interesting conundrum. What is an appropriate baseline that we can use to measure against? One possibility is to ask people for a policy preference before exposing them to the manipulation (i.e. to include a pre-test; e.g., as in Steen et al., 2014’s Experiments 1 and 2). There are several problems with this approach. For instance, if the pre- and post-test measures are taken close to each other in time, with participants being asked the same questions twice, the design becomes vulnerable to strategic decisions on the part of the participants (i.e. participants may feel anchored to their initial judgment or pressured to change their preference).

A second possible approach to gauging a baseline is to compare the metaphor framing conditions to a “neutral” or “non-metaphorical” framing condition. However, this path too is problematic: How would one establish that any alternative framing is indeed “neutral” with respect to the two metaphorical conditions?

For example, suppose we replace “Crime is a *beast/virus* ravaging the city of Addison” with “Crime is a *problem* ravaging the city of Addison” (as in Steen et al., 2014). There are two concerns.

1) How do we know that *problem* is not in fact more similar to one of the two metaphors? For example, if we were to find that responses to the *beast* frame are the same as to the *problem* frame, are we licensed to conclude that it is the *virus* frame that affects people’s thinking? Clearly not. All we would know is that the *beast* frame and the *problem* frame are similar, and that the *virus* frame is different from both of them. Since there is no external standard by which

we could judge the *problem* frame as neutral, comparing it to the other two conditions doesn't get us any closer to answering which of the frames may or may not have more of an effect on reasoning.

2) Linguistic stimuli are complex and people's responses to linguistic stimuli depend strongly on many properties such as frequency, vividness, conventionality, emotional valence, arousal and so on. Replacing *beast* or *virus* with *problem* introduces a host of uncontrolled changes along these dimensions. The same issue holds if one simply removes the words *beast* or *virus* to make "Crime is ravaging the city of Addison." That is, replacing or removing the key metaphorical nouns does not make a neutral stimulus, it makes a *different* stimulus.

To illustrate the difficulty of constructing an appropriate "neutral" comparison frame, we collected data in a second norming study. Participants in this task (N = 248) were asked to judge the severity, conventionality, and metaphoricality of three "Crime is a ..." frames: *virus*, *beast*, and *problem*.

We found that people rated the *problem* frame not only as less metaphorical than the *virus* or *beast* frames (paired sample t-tests revealed differences between the metaphor and *problem* frames, $ps < .001$), but also as connoting a less severe crime problem ($ps < .001$) and differing in conventionality ($ps < .001$) (see Table 1).

Table 1 Mean ratings of three "Crime is a ..." frames (*sd*)

	Severity	Metaphoricity	Conventionality
<i>Beast</i>	79.6 (17.0)	88.7 (13.9)	33.4 (24.1)
<i>Virus</i>	78.0 (15.9)	86.6 (16.0)	40.1 (26.1)
<i>Problem</i>	58.1 (26.8)	22.2 (25.5)	88.4 (15.3)

These ratings suggest that a *problem* frame differs from the *virus* and *beast* frames in a variety of ways. It is not only less metaphorical than the other two. But it connotes a less severe instance of crime and is viewed as a more conventional expression.

In contrast, the *beast* and *virus* frames are rated as connoting a similarly severe instance of crime, and are viewed as similarly conventional expressions.

As a result, we can be confident that differences in participants' policy preferences elicited by the *virus* and *beast* frames do not stem from differences in the degree to which the frames are metaphorical or connote a severe instance of crime. However, a *problem* frame fails to control for these important linguistic variables.

Conclusion

With two norming studies and two experiments, we further explored a metaphor framing effect in an important and dynamic real world context: crime. The two norming studies highlight critical factors for researchers to keep in mind when conducting framing studies in real world contexts. The first reveals the importance of using up-to-date norming data; the second illustrates limitations of comparing metaphorical to non-metaphorical frames.

The two experiments replicate prior work (Thibodeau & Boroditsky, 2011, 2013) and pinpoint the cause of a reported null effect (Steen et al., 2014). The first experiment shows that people are more likely to pursue enforcement-oriented policy interventions when crime is framed as a *beast* compared to when crime is framed as a *virus*. However, this experiment included a relatively ambiguous policy option – "neighborhood watches." In a follow-up experiment that excluded this option, we again found a reliable metaphor framing effect, showing that the effect of the metaphors did not depend on the presence of an ambiguous policy response.

In sum, the results confirm that natural language metaphors can affect the way we reason about complex problems.

References

- Barak G (Ed.) (1995) *Media, process, and the social construction of crime: Studies in newsmaking criminology* (Vol. 10). Taylor & Francis.
- Hauser, D. J., & Schwarz, N. (2014). The War on Prevention Bellicose Cancer Metaphors Hurt (Some) Prevention Intentions. *Personality and Social Psychology Bulletin*, 0146167214557006.
- Iyengar S (1994) *Is anyone responsible?: How television frames political issues*. University of Chicago Press.
- Jaeger FT (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434-446.
- Lakoff, G., & Johnson, M. (1980). *Metaphor we live by*. Chicago: University of Chicago Press.
- Landau, M.J., Sullivan, D., Greenberg, J. (2009). Evidence that self-relevant motives and metaphoric framing interact to influence political and social attitudes. *Psychological Science*, 20, 1421-1427.
- Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., Firth, D. (2014). Package "MASS": Support functions and datasets for Venables and Ripley's MASS.
- Steen, G.J., Dorst, A., Herrmann, B., Kall, A., Krennmayr, T., Pasma, T. (2010). *A method for linguistic metaphor identification*. Amsterdam, The Netherlands: Benjamins.
- Steen GJ, Reijniere WG, Burgers C (2014) When Do Natural Language Metaphors Influence Reasoning? A Follow-Up Study to Thibodeau and Boroditsky (2013). *Plos One* 9(12): e113536.
- Thibodeau, P.H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE*, 6(2): e16782.
- Thibodeau, P.H., & Boroditsky, L. (2013). Natural language metaphors covertly influence reasoning. *PLoS ONE*, 8(1): e52961.
- Venables, W.N., Ripley, B.D. (2002). *Modern applied statistics with S*. Springer.