The "Cry Wolf" Effect and Weather-Related Decision Making

Susan JOSLYN and Jared LECLERC
University of Washington, Seattle, WA, USA

ABSTRACT

Introduction: Poor compliance with weather warnings might be due partly to residents' prior experience with false alarms, leading them to distrust future warnings. **Method**: In a computer-based task, participants used forecasts to make weather-related decisions. Some participants were given explicit advice from an automated decision aid, comparable to weather warnings. We systematically manipulated the false alarm rate and whether the forecasts included uncertainty estimates. **Results and discussion**: Lower false alarms increased participants' decision quality and compliance, but the effect was small. Adding an uncertainty estimate to the forecast was more beneficial. These results carry important implications for weather risk communication.

KEYWORDS

Decision making; earth and atmospheric sciences; risk communication; decision support aid.

INTRODUCTION

General public end-users have vast experience with weather predictions and their subsequent outcomes: daily weather forecasts are readily available and widely used. Evidence suggests that, as a result, end-users have garnered a fairly sophisticated understanding of the relative accuracy of such forecasts and the uncertainty involved (Joslyn & Savelli, 2010). This is one domain in which most of us have at least some degree of expertise. Nonetheless, many people make what appear to be remarkably poor decisions when it comes to severe weather situations. Despite recent improvements in lead time and weather forecast accuracy, weather-related injury and death remain a serious problem. There is growing consensus that public response to warning forecasts, or lack thereof, is a significant contributing factor (e.g., Nagele & Trainor, 2012; Riad et al., 1999).

There are numerous possible reasons that people ignore weather warnings. Extreme weather events require vulnerable residents to make a choice between two general courses of action: 1) take precautionary action, e.g., evacuation, which is often regarded as costly because of the inconvenience, the risk of looting, etc. (Blendon, 2008; Cutter & Smith, 2009 ; Smith & McCarty, 2009); or 2) take no action and risk a severe loss, e.g., injury or death. In situations such as this, in which both alternatives involve potential losses, research suggests that people tend to be "risk seeking" (Kahneman & Tversky, 1979), preferring to take the risk rather than incur the cost of precautionary action. Furthermore, survey research suggests that people understand that extreme events are rare (Joslyn & Savelli, 2010) so they might underestimate the risk of an actual threat.

In addition, people might distrust weather warnings (Dow & Cutter, 1998; Morss & Hayden, 2010). Prior experience with false alarms, warnings that were perceived as unnecessary, may lead to a reluctance to comply with future alarms, known as the "cry wolf" effect (Breznitz, 1985). Because of the high potential loss that can result from extreme weather events, as well as the need to prepare several days in advance, weather warnings are often given when the probability of the event at any given location is low. Thus, the predominant error for warning forecasts tends to be a false alarm (Joslyn & Savelli, 2010; Barnes et al., 2007). While there has long been concern about the impact of false alarms on trust in weather warnings, the psychological effects remain unclear (Barnes et al., 2007). Among survey studies that investigate the impact of false alarms in natural settings, there is some evidence suggesting that people are fairly tolerant of such errors (e.g., Baker, 1991), while other evidence suggests classic false alarm effects (e.g., Atwood & Major, 1998). However, in survey studies, there are multiple uncontrolled variables, not the least of which is respondents' prior exposure to false alarms. There is some evidence that the cry wolf effect may not be apparent after a single false alarm but arises only after several false alarm experiences (Carsell, 2001). For that reason, an experimental approach may be required to obtain direct evidence for false alarm effects.

The experiment reported here is the first direct experimental evidence of which we are aware for a cry wolf effect in weather-related decision making. It was obtained by systematically manipulating the false alarm rate in a winter weather task to determine the impact on trust and compliance. The practical motivation for this research was to reproduce the effect in the laboratory and then explore factors to attenuate it. For instance, there may be a rate of false alarms that is tolerable to users and does not reduce trust or compliance (Roulston & Smith, 2004).

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

On the other hand, there may be ways to communicate warnings to reduce negative effects of false alarms and increase compliance so that warning thresholds need not be altered. Here we tested warnings that included probabilistic uncertainty estimates. Warnings that acknowledge the uncertainty up front may be judged less wrong when they fail to verify. Indeed, there is a growing literature that suggests that people have more trust in forecasts and make better decisions when uncertainty estimates are included (Joslyn et al., 2007; Nadav-Greenberg & Joslyn, 2009; Roulston et al., 2006). Thus, the research reported here also sought to explore whether uncertainty forecasts attenuated decrease in decision quality due to false alarms.

METHOD

Participants performed a computer-based task in which they made a series of decisions about applying salt brine treatment to the roads to prevent icy driving conditions the following morning. Participants were told that treatment should be applied beforehand if sub-freezing temperatures were anticipated that night. On each of 60 trials, simulating two winter months, participants chose between spending $1,000 to apply the salt treatment or risking a $5,000 penalty if they chose to withhold treatment and a freezing temperature was observed. They were given a monthly budget of $35,000 and paid cash commensurate with ending balance. Thus, the goal of the task was to maximize the budget. The decisions, which were a simplified version of the real-world task, were based on overnight low temperature forecasts and the advice of a computerized "decision support aid." After each trial, participants were informed of the observed temperature and any budget adjustments. The 60 forecasts and observed temperatures included a realistic range of temperatures for Washington State, where the experiment was conducted, as well as reliable probabilities where applicable.

The economically rational strategy in this situation (Murphy, 1977) was to apply salt treatment whenever the probability of freezing was 20% or higher. That is the point at which the expected value of the penalty (Bernoulli, 1954), obtained by weighting the $5,000 penalty by the probability that it will be incurred, and the cost of salting are equivalent: $5,000 (possible penalty) * 20% (probability of freezing) = $1,000 (cost).

On each trial, some participants had the advice of the "decision support aid" (DSA), which was described as taking into account the forecasted temperature and uncertainty, as well as the cost of salting and the penalty for not salting. In fact, it employed a rule based on the probability of freezing. In one condition, it was the economically rational rule, recommending salting whenever the probability of freezing was 20% or higher. Thus, following the advice of the decision support aid in this condition led to optimal performance, although, as with warnings for actual severe weather, it exposed participants to many false alarms, i.e., trials on which salting was recommended but freezing temperatures were not observed. Our own past research using this experimental paradigm and threshold suggested that people become reluctant to follow the advice after a few trials (Joslyn & LeClerc, 2012). Roulston & Smith (2004) suggested that raising the threshold to a level greater than the economically rational threshold might increase user compliance by lowering the false alarm rate (which, however, also increases the miss rate). Building upon this idea, we systematically manipulated the false alarm rate by creating three conditions in which the thresholds for recommending salt were 10%, 20% and 30% chance of freezing. The lower the threshold, the greater the proportion of false alarms and the smaller the proportion of misses. It is important to note, however, that the forecasted temperature, observed temperature and the economically rational threshold for salting (20% chance of freezing) were identical in all three conditions. The only thing that changed was the probability of freezing at which the decision support aid advised applying treatment.

In addition, in order to determine whether adding an uncertainty estimate to the forecast attenuates the cry wolf effect, we manipulated forecast format. One group of participants received the nighttime low temperature forecast and decision advice. Another group received the nighttime low temperature forecast, decision advice, and the probability of freezing. A final group received only the nighttime low temperature forecast and served as an overall control condition. Thus, there were 7 between-groups conditions.

Although this experimental task differs in many respects from real-world decisions, it shares several critical features with weather warning situations. Participants make decisions with outcomes that affect them directly, represented in the task as cash rewards. The decision task involves a choice between expending resources for protection and risking a potentially costly loss, as is the case when preparing for real-world weather events. Moreover, because of the high cost-loss ratio, precautionary action is required at a fairly low threshold, also true of many warning forecasts. Finally, it is a task in which a common error is risk seeking, mirroring the non-compliance with weather warnings.

RESULTS

We explored two dependent measures: mean expected value of participant decisions and compliance with the advice. Expected value was assessed by assigning each decision a theoretical value. A decision to salt was assigned the cost of salting, -$1,000. A decision against salting was assigned the penalty (-$5,000) multiplied by the probability of freezing on that trial (the likelihood that penalty would be incurred). A mean expected decision value was calculated for each participant. Compliance was summarized as a ratio of the number of times that the participant followed the advice provided and the control baseline rate (the number of advice-compatible decisions when no advice was provided), with higher values indicating greater compliance.

The false alarm rate had a significant effect on both expected value and compliance, suggesting a cry wolf effect. Participants experiencing the highest false alarm rate (DSA 10% condition) had significantly lower mean expected value than those with the lowest false alarm rate (DSA 30% condition) $t(64.66) = 2.56$, $p = .01$ (see table below). Moreover, compliance with the advice decreased as false alarms increased. Participants in the DSA 10% condition were significantly less compliant than participants in the DSA 30% condition, $t(69) = 2.25$, $p = .03$. However, performance and compliance were similar between participants in the DSA 20% and DSA 30% conditions, suggesting that lowering the false alarm rate slightly from the optimal rate is not effective.

Experimental condition	Mean expected decision value	Compliance ratio
DSA 10%	-$1,078.12	1.01
DSA 20%	-$1,049.55	1.10
DSA 20% + Freeze Probability	-$988.04	1.20
DSA 30%	-$1,033.36	1.12

In fact, the most beneficial manipulation was adding an uncertainty estimate to the forecast, which dramatically increased both compliance and mean expected decision value. Participants with forecasts that included both decision advice and the freeze probability had significantly higher mean expected decision values than participants with the advice alone in both the DSA 20% condition, $t(55.13) = 4.02$, $p < .001$, and the DSA 30% condition, $t(71) = 3.52$, $p = .001$. The same pattern of results was found for compliance.

CONCLUSION

This experiment provides evidence for both the cry wolf effect and the benefit of uncertainty forecasts. First, the decision advice with the highest false alarm rate led to worse decisions and lower compliance than did the advice with the lowest false alarm rate. However, a slight reduction of false alarms from the economically optimal threshold (DSA 20% condition) to the 30% threshold did not have a significant effect. Additional research is necessary to determine whether an even greater reduction in false alarms would be effective or whether the corresponding increase in miss rate would counteract any benefits.

Second, and even more importantly, advice that included probabilistic forecasts led to better decisions and higher compliance than advice alone. As noted in earlier studies (e.g., Joslyn & LeClerc, 2012), inclusion of uncertainty estimates might make forecasts seem more plausible, compensating any negative impact of false alarms, and allow individuals to tailor their decisions to their own risk tolerances.

Thus, despite the longstanding concern about cry wolf effects in warning forecasts, reducing false alarms to any practical level may not be effective in increasing compliance. This evidence suggests that a much more effective solution is to leave the false alarm rate where it is and provide users with an explicit uncertainty estimate.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation under Grant No. 1023354.

REFERENCES

Atwood, L. E., & Major , A. M. (1988). Exploring the “cry wolf” hypothesis. *International Journal of Mass Emergencies and Disasters,* 16, 279–302.

Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters,* 9, 287–310.

Barnes, L. R., Gruntfest, E. C., Hayden, M. H., Schultz, D. M., & Benight, C. (2007). False alarms and close calls: A conceptual model of warning accuracy. *Weather and Forecasting,* 22, 1140 –1147.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22, 23-36.

Blendon, R. J. (2008). High-risk area hurricane survey. [Available at www.hsph.harvard.edu/hurricane/topline.doc.]

Breznitz, S. (1985). False alarms: Their effects on fear and adjustment. *Issues in Mental Health Nursing,* 7, 335-348.

Carsell, K. (2001). Impacts of a false alarm : The January 29, 2000 Ventura, California experience. M.S. thesis. Dept. Of Geography and Environmental Studies, University of Colorado at Colorado Springs, 86 pp. [Available from Dept. Of Geography and Environmental Studies, University of Colorado at Colorado Springs, P.O. Box 7150, Colorado Springs, CO 80933.]

Cutter, S. L., & Smith, M. M. (2009). Fleeing from the hurricane’s wrath: Evacuation and the two Americas. *Environment,* 51, 26 -36.

Dow, K., & Cutter, S. L. (1998). Cry wolf: Repeat response to hurricane evacuation orders. *Coastal Management,* 26, 237–252.

Joslyn, S., & LeClerc, J. (2012). Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of Experimental Psychology: Applied,* 18, 126-140.

Joslyn, S., Pak, K., Jones, D., Pyles, J., & Hunt, E. (2007). The Effect of probabilistic information on threshold forecasts. *Weather and Forecasting,* 22, 804–812.

Joslyn, S., & Savelli, S. (2010). Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteorological Applications,* 17, 180–195.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica,* 47, 263-292.

Morss, R. E., & Hayden, M. H. (2010). Storm surge and "certain death": Interviews with Texas coastal residents following Hurricane Ike. *Weather, Climate, & Society*, 2, 174-189.

Murphy, A. H. (1977). The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Monthly Weather Review*, 105 (7), 803-816.

Nadav-Greenberg, L., & Joslyn, S. (2009). Uncertainty forecasts improve decision-making among non-experts. *Journal of Cognitive Decision Making,* 2, 24-47.

Nagele, D. E., & Trainor, J. E. (2012). Geographic specificity, tornadoes, and protective action. *Weather, Climate, & Society*, 4, 145–155.

Riad, J. K., Norris, F. H., & Ruback, R. B. (1999). Predicting evacuation in two major disasters: Risk perception, social influence, and access to resources. *Journal of Applied Social Psychology,* 29, 918-934.

Roulston, M. S., Bolton, G. E., Kleit, A. N., & Sears-Collins, A. L. (2006). A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting,* 21, 116-122.

Roulston, M. S., & Smith, L. A. (2004). The boy who cried wolf revisited: The impact of false alarm intolerance on cost–loss scenarios. *Weather and Forecasting,* 19, 391–397.

Smith, S. K., & McCarty, C. (2009). Fleeing the storm(s): An examination of evacuation behavior during Florida's 2004 hurricane season. *Demography,* 46, 127-145.