Communicating Warnings: Does Color-Coding Help?

Gala Gulacsik, Susan L. Joslyn, & John Robinson
University of Washington

Poster 3065

Introduction

- Currently, severe weather risk is communicated using "Watches" and "Warnings," although their effectiveness is debated.
- Research suggests that including explicit numeric probabilities in forecasts improve people's understanding of risk as well as the quality of their decisions [1].
- However, in a dynamic situation with multiple updates, including probabilities with each may overwhelm users ability to process the information.
- In many applied contexts, color-coded risk information is promoted as a simpler approach despite minimal evidence supporting this claim.

Research Questions

- Do people make better decisions with event likelihood information compared to the conventional Watch & Warning forecasts?
- Does the expression of likelihood make a difference to understanding (e.g. numerical, color-coded).
- Do people trust one format more than another?

Method

Task

- Participants (N=268) experienced 40 virtual storms (trials) that could produce tornadoes. Wind speeds of all storms were 73-112 mph.
- Participants received 7 sequential forecasts per trial. At each of the 7 decision points, they chose to *wait* for more information or to make a final decision for that trial – to *take shelter* or *not take shelter*. They learned whether a tornado hit at the end of the trial (based on a experiment by Schwartz and Howell, 1985).
- Participants earned extra credit and a cash reward for performance.

Cost-Loss Structure

Starting balance of 24,000 points. Participants were to minimize costs and losses.

Decision	Cost
Wait	Decision points 1-3: no cost Decision points 4-7: 20-points per wait decision
Take Shelter	$Shelter\ Cost = 300 + [3 * decision\ point^2]$
Not Take Shelter	No cost

1500-point penalty if a tornado hit and the participant chose to *not take shelter*

Conditions & Stimuli

Participants were randomly assigned to one of 3 between groups conditions.

1) Watch & Warning	2) Color	3) Probability	Optimal Decision
no watch or warning	green	< 13%	Not shelter
watch	yellow	≥ 13% and ≤ 24%	Wait
warning	orange	> 24%	Take Shelter

Storm Location Grid with Forecast Stimuli and Optimal Decision

Longitude (Decision Point); Latitude

Latitude	1	2	3	4	5	6	7	8
1				no w/w green 6%	no w/w green 3%	no w/w green 0%	no w/w green 0%	no tornado hit
2			watch yellow 13%	no w/w green 12%	no w/w green 11%	no w/w green 9%	no w/w green 0%	no tornado hit
3		watch yellow 19%	watch yellow 20%	watch yellow 21%	watch yellow 23%	watch yellow 24%	warning orange 30%	no tornado hit
4	watch yellow 20%	watch yellow 21%	watch yellow 23%	warning orange 25%	warning orange 29%	warning orange 34%	warning orange 40%	tornado hit
5		watch yellow 19%	watch yellow 20%	watch yellow 21%	watch yellow 23%	watch yellow 24%	warning orange 30%	no tornado hit
6			watch yellow 13%	no w/w green 12%	no w/w green 11%	no w/w green 9%	no w/w green 0%	no tornado hit
7				no w/w green 6%	no w/w green 3%	no w/w green 0%	no w/w green 0%	no tornado hit

Optimal Decision Key.
- wait
- take shelter
- not take shelter

Independent Variables

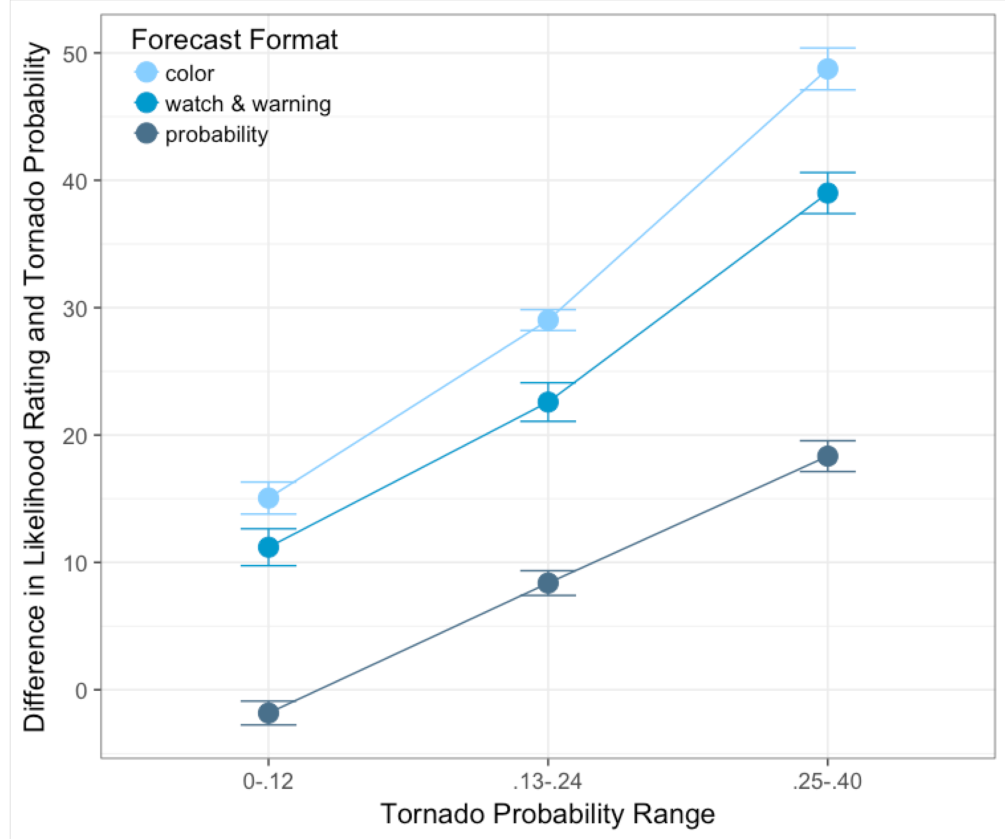
Between subjects: Forecast format (Watch & Warning; Color; Probability)

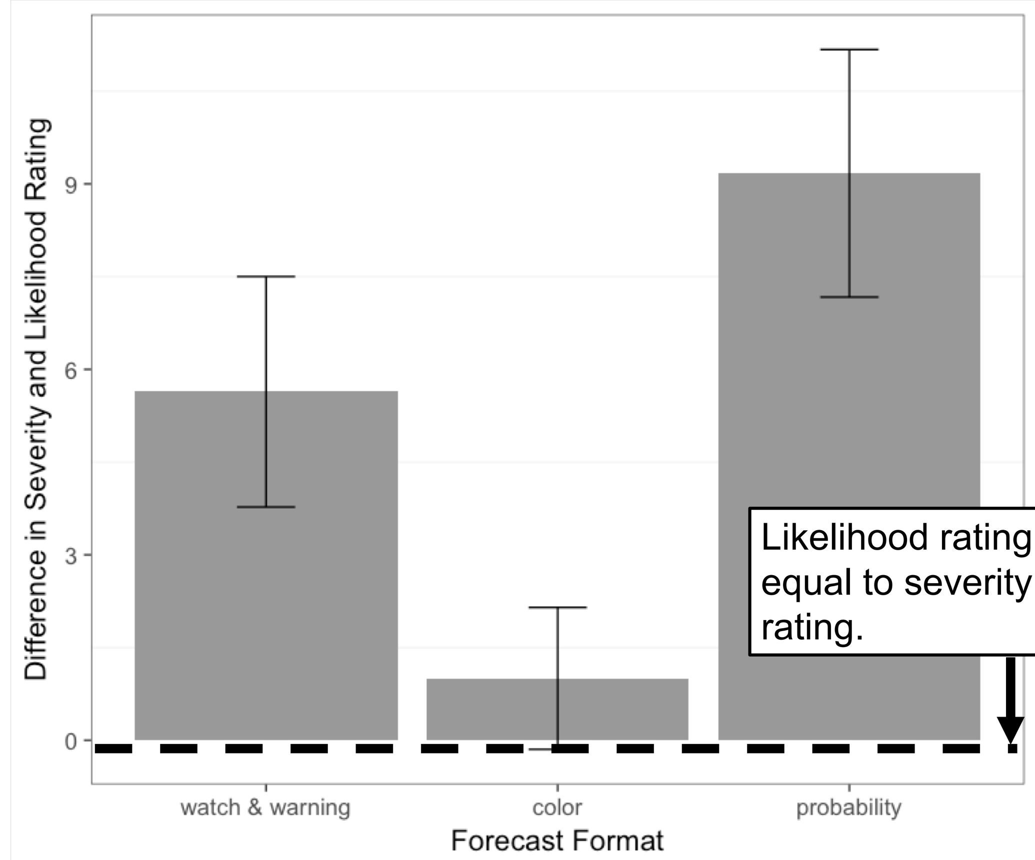
Within subjects: Tornado Probability Range

Dependent Variables

- Likelihood Rating: Impossible ——— Certain
- Severity Rating: Not severe ——— Very severe
- Trust Rating: Not at all ——— Completely
- Decision Quality: Expected Loss ((Cost of Shelter or Penalty) x Pr(Hitting Home))

Results: Understanding





Likelihood Rating

- Main Effect: Likelihood difference was greatest in Color and Watch & Warning forecasts, and lowest in Probability, $F(2,765) = 324.88$, $p<.001$, $\eta^2=.85$.
- Main Effect: As tornado probability increased so did the error in likelihood rating, $F(2,765) = 241.29$, $p <.001$, $\eta^2=.63$.
- Interaction: Likelihood difference was greater in the higher tornado probability for Color and Watch & Warning than for Probability forecasts $F(4,765)=7.29$, $p<.001$.

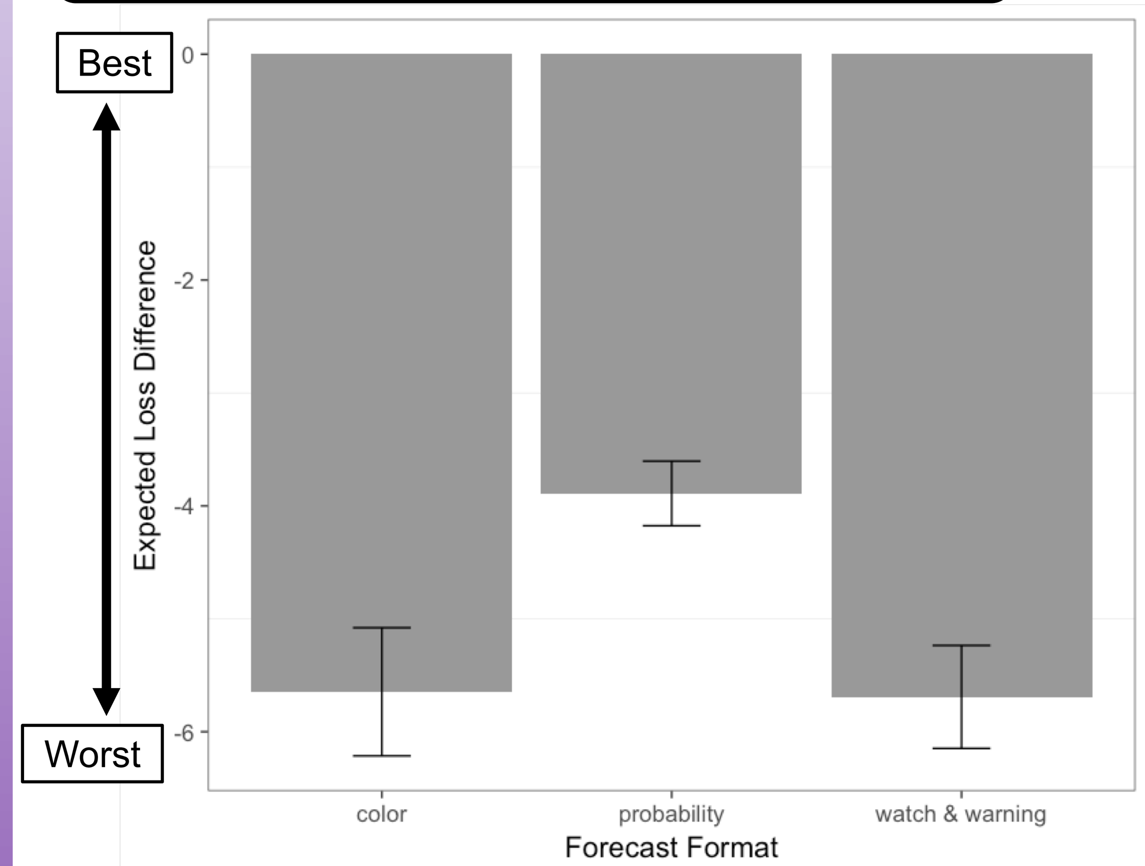
Mistaking Likelihood for Severity

- Main Effect: Participants with color-coded forecasts showed the least difference in likelihood and severity ratings suggesting that they were mistaking likelihood for severity, and those with probability showed the most difference, $F(2,255)= 15.23$ $p < .001$, $\eta^2=.12$.

Results: Trust

- Trust in Probability *(M=48.65, SD=18.49)* > Color *(M=40.16, SD=18.61)*, $p<.01$, but was not significantly different from the Watch & Warning forecasts *(M=45.94, SD=19.11)*.

Results: Decision Quality



Expected Loss (EL) Difference Formula

$$EL_{difference} = \frac{\sum_{i=1}^{n} EL_{min} - EL_{decision}}{n}$$

EL_{min} = Expected loss of the optimal decision.
$EL_{decision}$ = Expected loss of the participant's decision.
n = Number of decisions up to the final decision.

- EL differences of decisions with Probability forecasts *(M = -3.89, SD = 2.71)* were the smallest, while Color *(M = -5.65, SD = 5.08)*, and Watch & Warning *(M = -5.69, SD = 4.27)* showed the greatest EL difference, $F(2,255) = 5.53$, $p < .001$, $\eta^2=.04$.
- EL difference of Probability < Color, $p<.01$, and watch & warning forecasts, $p < .01$.

Conclusions

- Despite the expected heavy cognitive load of numeric probabilities, participants made better decisions, understood the forecasts best, and trusted, and did so with multiple forecasts updates.
- Decision quality and trust in the forecast was lowest for watch & warning and color-coding decisions.

References

1. Joslyn, S. L., & LeClerc, J. E. (2013). Decisions with uncertainty: The glass half full. *Current Directions in Psychological Science, 22*(4), 308-315.
2. Schwartz, D. R., & Howell, W. C. (1985). Optional stopping performance under graphic and numeric CRT formatting. *Human Factors, 27*(4), 433-444.

This material is based upon work supported by the NSF under grant number 1559126. For more information contact Gala Gulacsik, gala@uw.edu.