The Impact of Inconsistent Forecasts on User Trust

Jessica N. Burgeno & Susan L. Joslyn
University of Washington

Introduction

- The National Weather Service believes forecast consistency (ex: forecast provided on day 1 the same as day 2) is important for user trust(1).
- Because weather models are constantly updating, growing more accurate on average (2, 3), preserving consistency can be at a cost to accuracy.
- While the negative effect of inaccuracy on trust is well supported (4-6), there's little to no support for the anticipated negative effect of sequential inconsistency on trust;
- However, consistency among multiple simultaneous advisors has been found to enhance confidence in decisions made based on their advice (7).

Research Questions

- Does inconsistency impact user trust?
- How does it relate to the already established impact of inaccuracy on trust?
- To what degree are participants influenced by earlier forecasts when they are inconsistent?

Method

Task:

- Undergraduate participants (N=162) made several school closure decisions based on snow accumulation forecasts made 1-and 2-days prior to the expected snow storm.
- Participants earned a cash reward commensurate with performance, and course credit.

Instructions: expect ≥ 6 inches of snow accumulation → **Close**
expect < 6 inches → **Stay Open**

Cost Structure:

Decision	Observed Accumulation: < 6" snow	Observed Accumulation: =>6" snow
Open	0pts	6pts
Close	2pts	2pts

- Starting balance: 120pts
- $1/4pts over 72pts

Independent Variables

Within subjects

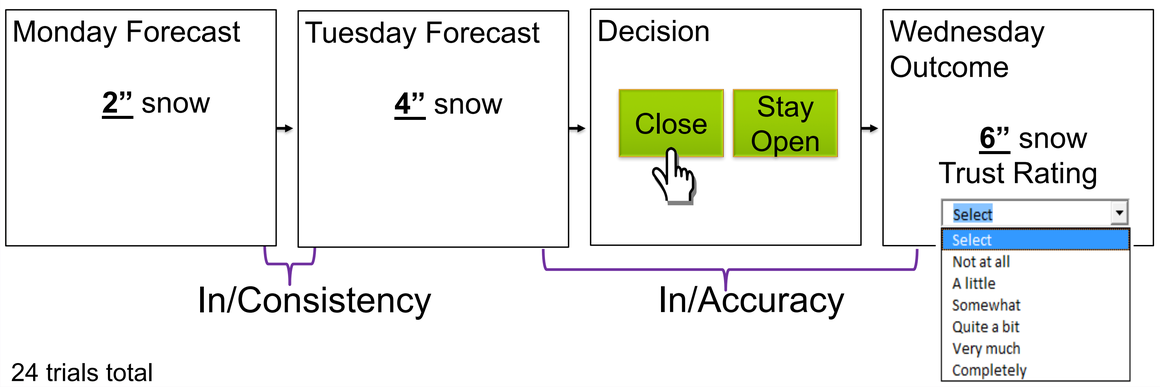
- Consistency: Consistent: Day 1= Day 2 snow forecast (in inches)
  Inconsistent: Day 1≠Day 2 snow forecast (difference = 2 in.)
- Accuracy: Accurate: Day 2 snow forecast = Observed snow (in inches)
  Inaccurate: Day 2 snow forecast ≠ Observed snow (difference = 2 in.)

Forecast Type was also manipulated between subjects in a previous experiment (Deterministic, Probabilistic), however there was no effect perhaps because forecasts were not reliable.

Dependent Variables

- Trust (6-point scale: "Not at all" to "Completely")
- Snow Accumulation Estimates (in inches)

Trial Events



24 trials total

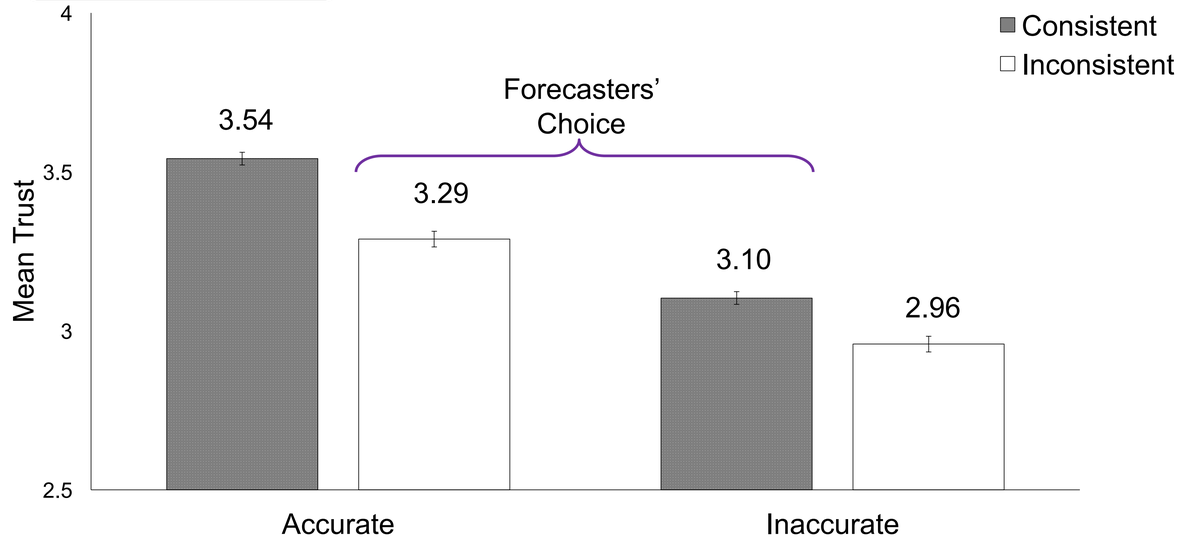
Method (Continued)

Forecast Data

	Accurate			Inaccurate		
	Forecasts		Observed	Forecasts		Observed
	Day 1	Day 2	Outcome	Day 1	Day 2	Outcome
Consistent	4	4	4	4	4	6
	5	5	5	5	5	7
	6	6	6	6	6	4
	7	7	7	7	7	5
Inconsistent	4	6	6	3	5	7
	5	7	7	2	4	6
	6	4	4	9	7	5
	7	5	5	8	6	4

2x2 within subjects design; the magnitude of all inconsistencies and inaccuracies was 2 inches

Results: Trust



Mean trust rating by trial type. Error bars show 95%. CI.

- Trust in Consistent (M=3.32) > Inconsistent forecasts (M=3.12), *F(1, 160)=12.59, p=.001*
- Trust in Accurate (M=3.41) > Inaccurate forecasts (M=3.03), *F(1,160)=52.18, p<.001*
- Effect of Inconsistency, η_p^2***=.07*** < Effect of Inaccuracy, η_p^2***=.25***
- Consistency x Accuracy interaction failed to reach sig., *F(1,160)=1.94, p=.17,* η_p^2*=.01*
- Forecasters' Choice: Trust in Accurate Inconsistent > Inaccurate Consistent trials, t(161)=2.35, p=.02

Results: Weighting Forecasts

Research question: Arguably users should ignore Forecast 1 as updated Forecast 2 replaces it. Do they?

- A 2 factor regression model explained 56% of variance in snow accumulation estimates, *F(2,1293)=825.36, p<.001,* R^2*=.56.*
- Weighting of Day 2, ***β=.71*** *(p<.001)* > Day 1 forecast, ***β=.08*** *(p<.001)*

Conclusions

- Forecast inconsistency reduces trust in forecasts but not to the extent that inaccuracy does in this experimental paradigm.
- To preserve trust, meteorologists should prioritize accuracy over the maintenance of consistency.
- People demonstrated much greater weighting of more recent forecasts, suggesting they may understand it is more accurate.

References

1.National Oceanographic and Atmospheric Administration. (2016). Risk communication and behavior… **2.**Lazo, J. K., Morss, R. E., & Demuth, J. L. (2009). Bulletin of the American Meteorological Society, 90(6), 785. **3.**Wilson, L. J., & Giles, A. (2013). *Meteorological Applications*, 20(2), 206. **4.**Gupta, N., Bisantz, A. M., & Singh, T. (2001). In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 45(23), 1699. **5.**Kahn, B. E., & Luce, M. F. (2003). *Marketing Science*, 22(3), 393. **6.**Joslyn, S. L., & LeClerc, J. E. (2012). *Journal of Experimental Psychology: Applied*, 18(1), 126. **7.**Budescu, D. V. (2005). In Information Sampling and Adaptive Cognition, 327.

This material is based upon work supported by the NSF under grant number 1559126. For further information please contact Jessica Burgeno at jburgeno@uw.edu.