Speaker reliability: calibrating confidence with evidence

Mélinda Pozzi & Diana Mazzarella

Cognitive Science Center, University of Neuchâtel, Switzerland

Overconfidence is typically damaging for one's reputation as a reliable source of information. Research in psychology shows that, when deciding whether to trust a speaker, addressees do not exclusively rely on the speaker's confidence ("confidence heuristics"), but consider, whenever possible, whether the speaker's degree of confidence matches with the accuracy of their claim. As a result, a confident speaker whose messages turn out to be false will typically lose their credibility (Tenney et al., 2007; 2008; 2011; Vullioud et al., 2017).

Crucially, though, preliminary findings from Tenney et al. (2008) indicate that an overconfident speaker does not suffer any reputational costs if their mistake is taken to be *justified*. This suggests that the speaker's perceived reliability as a source of information depends on whether their confidence matches with the quality of the evidence at their disposal ("confidence-evidence calibration"). If this is the case, then, even an accurate informant should lose their credibility if the evidence available to them does not warrant the degree of confidence expressed (bad confidence-evidence calibration).

The present study has two aims. First, replicating Tenney et al. (2008) results showing that overconfidence does not backfire if inaccuracy is justified by strong evidence: an inaccurate confident speaker who communicates false information that is justified by strong evidence does not lose their credibility (hypothesis 1 – experiment 1). Second, investigating whether confidence can backfire if accuracy is not justified by enough evidence (the speaker is accurate "by chance"): an accurate confident speaker who makes a claim that is not supported by enough evidence will lose their credibility (hypothesis 2 – experiment 2). Our study is pre-registered here: https://osf.io/fbv8g/?view_only=8d90bab9d82a43e1a7928e4de4aca7ef

We conducted two online experiments in which participants were presented with two testimonies concerning a car accident, judged the credibility of the two witnesses (one confident and one unconfident male witness), and were asked to choose which of the two depositions they believe. The material was adapted from Tenney et al. (2008, Experiment 2). In experiment 1, both witnesses were inaccurate but were justified by strong evidence. In experiment 2, both witnesses were accurate but had weak evidence. In both experiments, we measured participants' credibility judgments (on a scale from 1 to 6) and believability choices (*Who do you believe?*) at three distinct times: (1) participants have no information about accuracy and strength of evidence, (2) participants get feedback about accuracy, (3) participants get feedback about evidence.

The first experiment (N = 108) replicated Tenney et al. (2008) results. There was a significant interaction effect of confidence and time on credibility ($F(2, 321) = 93.018$, $p < .001$), and believability changed over time as predicted ($\chi^2(2) = 9.663$, $p = 0.008$). In the absence of any information about accuracy and strength of evidence (Time 1), the confident witness was rated as more credible and was more likely to be believed than the unconfident witness. At Time 2, when both witnesses turned out to be inaccurate, the confident witness lost his credibility to the benefit of the unconfident witness. At Time 3, when the inaccuracy was found to be justified by strong evidence, the confident witness' credibility was restored. The second experiment ($N = 109$) supported our second hypothesis. There was a significant interaction effect of confidence and time on credibility ($F(2, 324) = 35.115$, $p < .001$), and believability changed over time as predicted ($\chi^2(2) = 45.942$, $p < 0.001$). In the absence of any information about accuracy and strength of evidence (Time 1), the confident witness was rated as more credible and was more likely to be believed than the unconfident witness. At Time 2, when both witnesses turned out to be accurate, the confident witness kept his credibility. At Time 3, when the testimony of the witnesses was found to be warranted by weak evidence, the confident witness suffered a reputational loss.

This study shows that reputation management in communication depends on how well the speaker's confidence is calibrated to her evidential basis.

References

Tenney, E. R., MacCoun, R. J., Spellman, B. A., & Hastie, R. (2007). Calibration trumps confidence as a basis for witness credibility. *Psychological Science, 18*(1), 46–50. http://dx.doi.org/10.1111/j.1467-9280.2007.01847.x

Tenney, E. R., Spellman, B. A., & MacCoun, R. J. (2008). The benefits of knowing what you know (and what you don't): How calibration affects credibility. *Journal of Experimental Social Psychology, 44*(5), 1368–1375. http://dx.doi.org/10.1016/j.jesp.2008.04.006

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental psychology, 47*(4), 1065–1077. http://dx.doi.org/10.1037/a0023273

Vullioud, C., Clément, F., Scott-Phillips, T., & Mercier, H. (2017). Confidence as an expression of commitment: Why misplaced expressions of confidence backfire. *Evolution and Human Behavior, 38*(1), 9–17. http://dx.doi.org/10.1016/j.evolhumbehav.2016.06.002
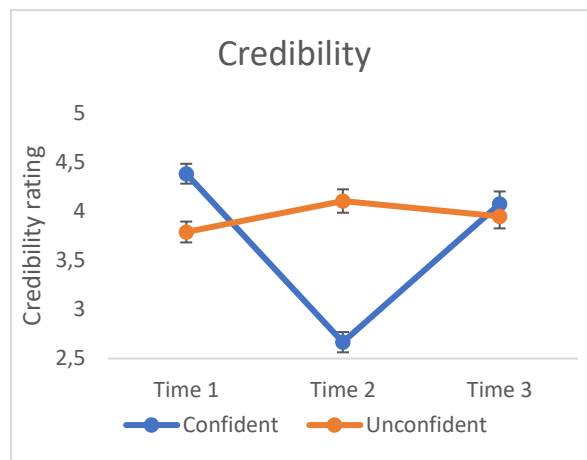
Figures

Experiment 1 - hypothesis 1

Figure 1. Credibility scores of the confident (blue) and unconfident (orange) witness, on a scale from 1 "not credible" to 6 "credible".
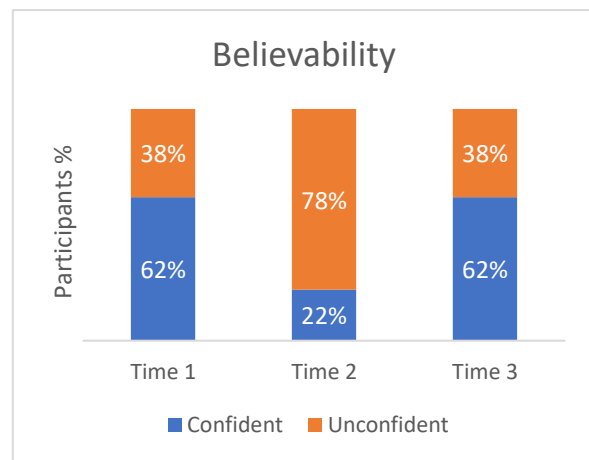

Figure 2. Percentage of participants who believed the confident (blue) or unconfident (orange) witness.
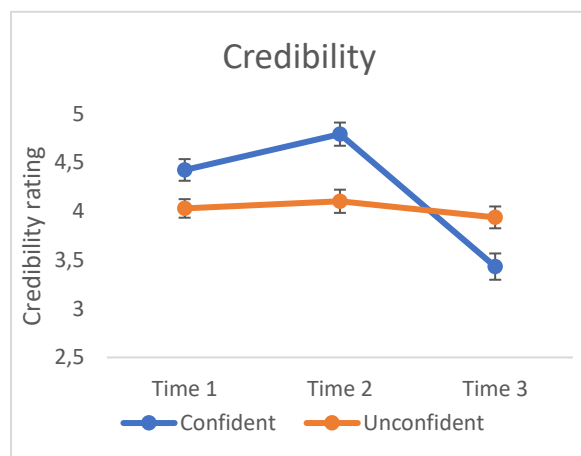
Experiment 2 – hypothesis 2

Figure 3. Credibility scores of the confident (blue) and unconfident (orange) witness, on a scale from 1 "not credible" to 6 "credible".
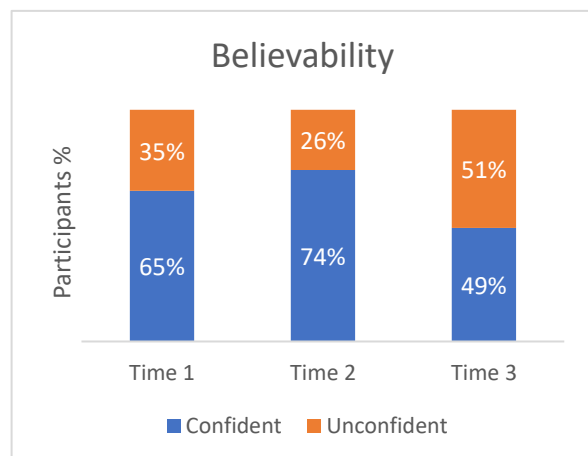

Figure 4. Percentage of participants who believed the confident (blue) or unconfident (orange) witness.