Reporting model results

Judith Degen

May 24, 2021

1   General notes on reporting results

When reporting results, here's a good set of steps to follow:

1. Start by pointing to a relevant visualization of the data and reminding the reader of the questions to be addressed with the data analysis.

2. Report the analysis you conducted (e.g., mixed effects logistic or linear regression) and the model specification: what was the outcome variable? What were the fixed and random effects? How were the fixed effects coded?

3. Report the model coefficients and direction of effects.

4. Optional: if you don't have a dedicated Discussion section, include brief discussion of results in the Results section.

The next two sections contain examples of reporting mixed effects logistic regression (Section 2) and mixed effects linear regression (Section 3). The prose in Section 2 is taken (and slightly modified) from Degen, Hawkins, Graf, Kreiss, and Goodman (2020). The prose in Section 3 is taken (and slightly modified) from Degen and Tonhauser (to appear).

2 Example: reporting mixed effects logistic regression results

Proportions of redundant *color-and-size* utterances are shown in Figure 1 alongside model predictions (to be explained further in Section 5. There are three main questions of interest: first, do we replicate the color/size asymmetry in probability of redundant adjective use? Second, do we replicate the previously established effect of increased redundant color use with increasing scene variation? Third, is there an effect of scene variation on redundant size use and if so, is it smaller compared to that on color use, as is predicted under asymmetric semantic values for color and size adjectives?
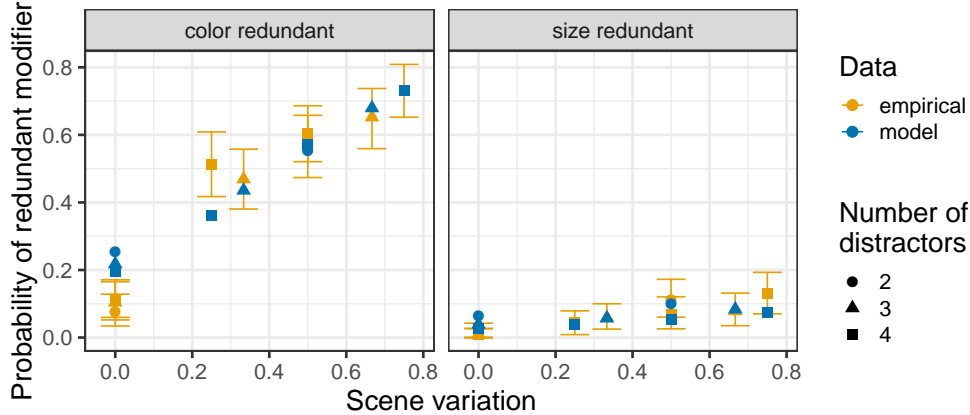

Figure 1: Empirical redundant utterance proportions (blue) alongside point-wise maximum a posteriori (MAP) estimates of the RSA model's posterior predictives for redundant utterance probability (red) as a function of scene variation in the color redundant (left) and size redundant (right) condition. Here and in all following plots, error bars indicate 95% bootstrapped confidence intervals.

We addressed all of these questions by conducting a single mixed effects logistic regression analysis predicting redundant over minimal adjective use from fixed effects of sufficient property (color vs. size, centered, size positive), scene variation (proportion of distractors that does not share the insufficient property value with the target, centered), and the interaction between the two.[1] The model included the maximal random effects structure justified by the design:[2] by-participant and by-item random intercepts as well as by-participant random slopes for the fixed effects of sufficient property and scene variation.

We observed a main effect of sufficient property, such that speakers were more likely to redundantly use color than size adjectives ($\beta = 3.54$, $SE = .22$, $p < .0001$), replicating the much-documented color-size asymmetry. We further observed a main effect of scene variation, such that redundant adjective use increased with increasing scene variation ($\beta = 4.62$, $SE = .38$, $p < .0001$). Finally, we also observed a significant interaction between sufficient property and scene variation ($\beta = 2.26$, $SE = .74$, $p < .003$). Simple effects analysis revealed that the interaction was driven by the scene variation effect being smaller in the *color-sufficient* condition ($\beta = 3.49$, $SE = .65$, $p < .0001$) than in the *size-sufficient* condition ($\beta = 5.75$, $SE = .38$, $p < .0001$), as predicted if size modifiers are noisier than color modifiers.

---

[1] All mixed effects analyses reported in this paper were conducted with the `lme4` package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2017).

[2] Alternatively: "that allowed the model to converge"

3 Example: reporting mixed effects linear regression results

Figure 2 shows the mean prior probabilities of the 20 contents by fact. To assess whether fact type affected prior content probability, we conducted a mixed-effects linear regression predicting slider rating from dummy-coded fact type (reference level: 'lower probability') and random by-item and by-participant intercepts and slopes for fact type.[3] Each content's mean prior probability was rated as higher when it was presented with its higher probability fact than when it was presented with its lower probability fact ($\beta = 0.45$, $SE = 0.01$, $t = 31.12$, $p < .0001$). This suggests that the manipulation of the prior probability of the 20 contents was successful.
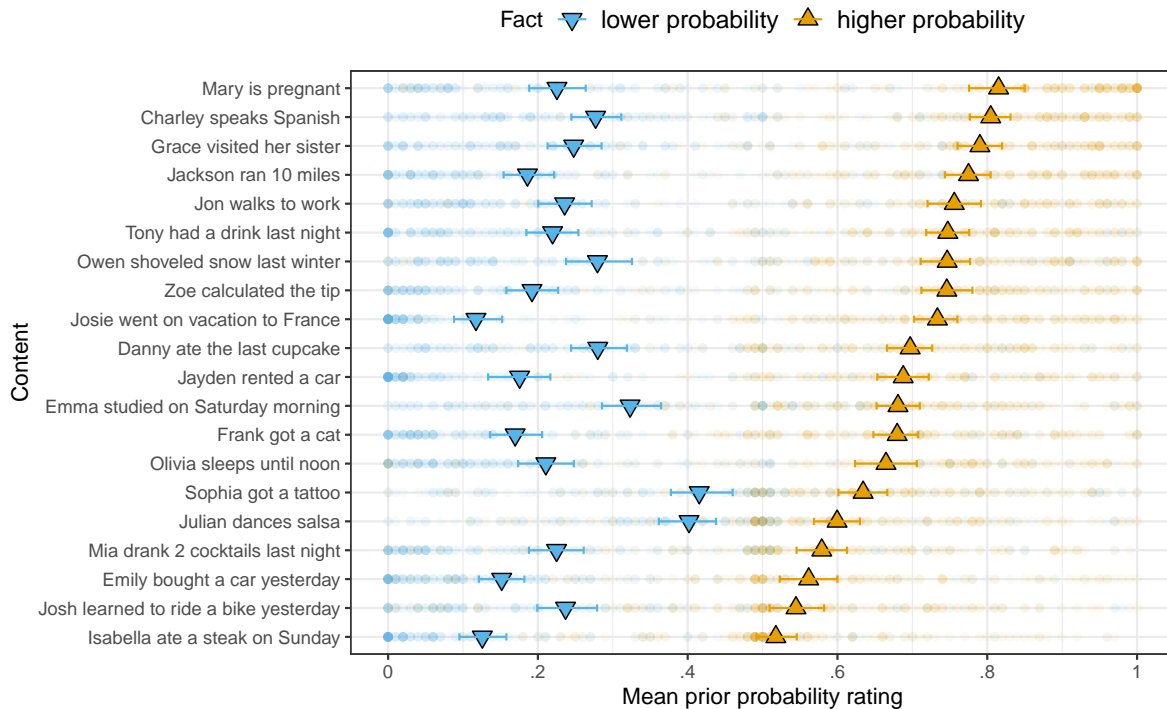

Figure 2: Mean prior probability by content and fact in Exp. 1. Error bars indicate 95% bootstrapped confidence intervals. Transparent dots indicate individual participant ratings.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Degen, J., Hawkins, R. X. D., Graf, C., Kreiss, E., & Goodman, N. D. (2020). When redundancy is useful: A Bayesian approach to 'overinformative' referring expressions. *Psychological Review*, *127*, 591621. doi: 10.1037/rev0000186

Degen, J., & Tonhauser, J. (to appear). Prior beliefs modulate projection. *Open Mind: Discoveries in Cognitive Science*.

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria.

---

[3]All analyses were conducted in R (R Core Team, 2017) using the `lme4` package (Bates et al., 2015).