BCS Final Paper

Tyler Trine

1) What is your "elevator pitch" for your target study (the summary of the importance, method, and main findings of your target study that you could say in a 45 second elevator ride ~ 100 words). Your audience is Neil Degrasse Tyson (high science literacy, non-specialist).

The brain is a prediction engine; it continuously integrates its inputs to form expectations for the future. An open cognitive question, however, concerns the adaptive capacity of this engine. Once it learns that a space is predictable, is this knowledge fixed? Or can it detect when its predictive assumptions no longer apply, and override them? Take, for example, the use of adjectives in English. We usually say, "pick up the small cup" if there are two or more differently-sized cups. Let's say there are just two. Someone attempting to fulfill the request to "pick up the small cup" will accurately predict the object of the request (i.e., the 'small cup') after hearing only the word 'small'. This has been empirically demonstrated by tracking the eye movements of subjects. Now let's add a twist: Imagine the speaker uses adjectives incorrectly; that she always calls the cup small, even when there is no other cup present. How do listeners react? Do they continue to predict the noun after hearing the adjective? Or do they wait, because they notice that such a prediction is, in general, invalid for this particular speaker? Grodner and Sedivy's 2011 paper finds the latter. This result suggests that the brain updates the way in which it predicts the future, even as it updates the predictions themselves.

2) Looking at the results section of your target experiment, describe which reported effects from the target experiment you were able to replicate and which effects did not replicate. For all effects, qualitatively compare the directions of the effects and the magnitudes of the effects. BE CLEAR ABOUT WHAT EACH EFFECT IS MEASURING AND COMPARING, AND WHAT IT MEANS.

The original study found, in line with prior research, that:
1) Subjects correctly predicted the target earlier when the trial had a contrast pair than when it did not.

It also found that:
2) This effect ceased to be significant when the speaker ceased to be reliable.

Further, it found:
1~2) There is no significant interaction between these two factors.

Our results confirmed 1) and 1~2), but not 2). Actually, we found a (not statistically significant) reversal in the direction of the target results for 2). In other words, the subject looked at the target even earlier when the speaker was unreliable. Our results, then, strengthen the case that the brain performs predictive inference, but they do not strengthen the case that this inference is adaptable to unexpected contexts. As I will discuss in the next prompt, there are a number of shortcomings in both the target study and our replication which may contribute to this failure to replicate.

3) Do you think any of the failures to replicate (that you described in section 2) indicate that the original study reported a false effect? Why or why not. BE SPECIFIC. If instead there were few or no failures to replicate describe why you think your experiment is so robust.

As is often the case, it's hard to say. In short, not necessarily. The study has so many moving parts; one cannot possibly hope to control for all confounds. Here is our methodological analysis, both for the original study and for our replication.

Flaws in our design

- Perception of speaker unreliability. As we learned in our follow-up survey, subjects weren't too convinced that the speaker was unreliable. They may have anticipated a phonological speech impairment from our instructions.
- Image set. We used non-normalized clipart which was appropriated ad-hoc from the interned. Further, the way we generated contrastive pairs involves distortions.
- Noise. The experiment's location and equipment switched midway through. There were also minor inconsistencies with instruction delivery at first.

Flaws in their analysis

- Confounds. Their study has another within-subject factor: Scalar vs. material modifiers. They do not report the significance of this interaction.

Divergence from original study

- Them: Physical objects. Us: Computer screen.
- Them: Live speaker. Us: Recorded audio.
- Them: Task size of several sq. feet. Us: ~1 sq. foot.

4) Provide a list of your group members, including yourself, and your subjective estimate of each person's contribution to the research project (summing to 100% including yourself). Give your best subjective estimate based both on the amount of time and ideas each person contributed.

Bethany - 33.3%
Anaclare - 33.3%
Tyler - 33.3%