Erratum for "On quantifying schematicity of future narratives"

Isaac Kinley[1,*] Donna Rose Addis[1,2], Reece P Roberts[3,4], Samuel Fynes-Clinton[1], Yang Xu[2]

1 Rotman Research Institute, Baycrest Academy of Research and Education
2 Department of Psychology, University of Toronto
3 School of Psychology, Faculty of Science, University of Auckland
4 Centre for Brain Research, University of Auckland
5 Department of Computer Science, Cognitive Science Program, University of Toronto
*** correseponding@author.mail**

In Kinley et al. (2024), we incorrectly reported the processing steps used for experiment 1. While the paper is accurate with respect to experiment 2, a file organization issue led to an earlier processing pipeline being used to derive experiment 1 results. First, preprocessing of narratives in experiment 1 involved:

1. Expanding contractions
2. Removing punctuation
3. Whitespace tokenization
4. Removal of stopwords defined by the Natural Language Toolkit (NLTK; Bird et al., 2009)

Second, to derive a word association graph based on the *Small World of Words* norms (De Deyne et al., 2019), weights between pairs of words (which were not lemmatized) were computed as:

$$W_{\text{word1,word2}} = \frac{p(\text{word1}|\text{word2}) + p(\text{word2}|\text{word1})}{2}$$

I.e., given the probability of *word1* as a response to the prompt *word2* and its converse, we computed the average rather than the maximum.

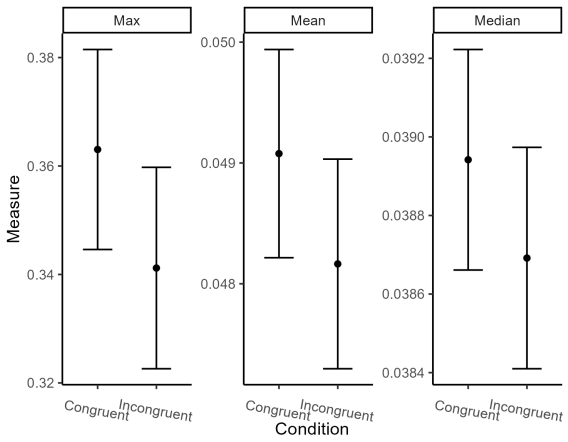

Figure 1. Corrected version of figure 3, top panel, in Kinley et al. (2024).

The statistics given in the Kinley et al. (2024) are thus based on computing associativity measures following this pipeline for experiment 1 and following the reported pipeline for experiment 2. When we apply the reported pipeline to experiment 1, p values for the mixed effects models (Singmann et al., 2024) comparing mean, median, and max associativity between the congruent and incongruent conditions were in fact .395, .337, and .353. This is reflected in figure 1, a corrected version of the top panel of figure 3 from Kinley et al. (2024).

The individual-wise results shown in figure 4 in Kinley et al. (2024) remain unchanged. That is, averaging the mean associativity measure within each condition for each participant, the same number of participants had on average higher mean associativity in

the congruent versus incongruent condition in experiment 1 for both processing pipelines.

In the final mixed effects model, combining results from both experiments to model mean associativity with fixed effects of experiment and experimental condition and a random intercept for each participant, the effect of experimental condition remained significant ($p = 0.01$). However, mean associativity was no longer higher for experiment 2 vs 1 ($p = 0.21$)—this result had been an artifact of the differing preprocessing steps between the two conditions.

With these updated results, it is necessary to revise our discussion somewhat. On incongruent trials in experiment 1, participants imagined events involving a location, person, and object each drawn from a different "sphere" of their life. In contrast, on incongruent trials in experiment 2, participants imagined events involving locations and pairs of objects that had been judged unlikely to co-occur in a large online norming study. Thus, whereas experiment 2 was based on in/congruence of stimuli with culturally shared schemas, experiment 1 was based on in/congruence with idiosyncratic schemas based on individuals' particular experiences. Word associations would be expected to be most informative with respect to culturally shared schemas, and we expressed some surprise in Kinley et al. (2024) that they were apparently informative with respect to idiosyncratic schemas. With the updated results, we can see that word associations are less informative with respect to idiosyncratic schemas after all.

References

Bird, S., Klein, E., and Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.".

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., and Storms, G. (2019). The "small world of words" english word association norms for over 12,000 cue words. *Behavior research methods*, 51:987–1006.

Kinley, I., Addis, D. R., Fynes-Clinton, S., Roberts, R., and Xu, Y. (2024). On quantifying schematicity of future narratives. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 46.

Singmann, H., Bolker, B., Westfall, J., Aust, F., and Ben-Shachar, M. S. (2024). *afex: Analysis of Factorial Experiments*. R package version 1.3-1.